Cochrane Library

Advances in Evidence Synthesis

special issue

Research abstracts 2020

cochranelibrary.com

Advances in Evidence Synthesis: special issue

The abstracts in this Supplement to the *Cochrane Database of Systematic Reviews* were originally submitted for presentation at the Cochrane Colloquium that was planned to take place in Toronto, Canada, on 4-7 October 2020.

In May 2020, as a result of the continuining spread and impact of COVID-19, Cochrane's Governing Board and Cochrane Canada agreed to cancel the Toronto Colloquium and hold it instead in **September 2022**, with the **Global Evidence Summit** already scheduled for 2021. Despite the cancellation, the review process for the submitted abstracts continued, and this Supplement includes accepted abstracts of research that would have been presented as oral or poster presentations. Note that a number of the abstracts refer to ongoing work, with results planned to be presented at the Colloquium or at a future date.

Published September 2020

This Supplement can be cited as:
Advances in Evidence Synthesis: special issue. *Cochrane Database of Systematic Reviews* 2020;(9 Suppl 1).
https://doi.org/10.1002/14651858.CD202001

Abstracts can be cited as:
[Authors]. [Abstract title]. Advances in Evidence Synthesis: special issue. *Cochrane Database of Systematic Reviews* 2020;(9 Suppl 1):[abstract page number] https://doi.org/10.1002/14651858.CD202001

Contents

Scientific committees and reviewers

Scientific committee

Holger Schünneman (Chair)	John Lavis	Nancy Santesso
Elie Akl	Eva Madrid	Danielle Schlosser
Jeffrey Braithwaite	Joerg Meerpohl	Mark Skinner
Claudia Emerson	Jordi Pardo Pardo	Karla Soares-Weiser
Miloslav Klugar	Ludovic Reveiz	Peter Tugwell
Andreas Laupacis	Craig W Robbins	Alexis Turgeon

Abstract sub-committee

Grigorios Leontiadis (Co-Chair)	Solange Durão	Liliya Eugenevna Ziganshina
Nancy Santesso (Co-Chair)	Giordano Perez-Gaxiola	

Abstract reviewers

Sue Brennan	Susan MacDonald	Pankaj B Shah
Anna Chaimani	Martin Meremikwu	Ian Shemilt
Andrea Cipriani	Kaelan Moat	Siddharth Singh
Luis Colunga	Carl Moons	Xin Sun
Emmanuel Effa	Hassan Murad	Yemisi Takwoingi
Maicon Falavigna	Adriani Nikolakopoulou	Virginia F Moça Trevisani
Ronald Flumingnan	Anna Noel-Storr	Marialena Trivella
Gerald Gartlehner	Jane Noyes	Apostolos Tsapas
Su Golder	Chavdar Pavlov	Frances Tse
Claudia Granados	Rapat Pittayanon	David Tunnicliffe
Takeshi Kanno	Livia Puljak	Charles Wiysonge
Tamara Kredo	Maria Ximena Rojas Reyes	Taryn Young
Mariska Leeflang	Anke Rohwer	Cathy Yuan
Carol Lefebvre	Meghan Sebastianski	

PRIORITY SETTING

A snapshot of the review prioritization work conducted by the Gynaecological, Neuro-oncology and Orphan Cancers Group since 2007

Quinn G[1], Jess C[1], Grant R[1], Tomlinson E[2], Harrison T[1], Platt J[1]

[1] Gynaecological, Neuro-oncology and Orphan Cancers, UK; [2] Cancer Network, UK

Background: Setting systematic review topic priorities is vital for Cochrane Review Groups. It helps to ensure Cochrane Review evidence is relevant to end-users such as clinicians, consumers, healthcare professionals and policy makers. Since 2007, the Gynaecological, Neuro-oncology and Orphan Cancers (GNOC) Group have been involved in five priority setting exercises.

Objectives: To prioritise systematic review topics, including updates of reviews and new review titles, to identify the top priority topics in the areas of gynae-oncology and neuro-oncology.

Methods: We have designed and conducted three in-house prioritization exercises. In 2007, we engaged with stakeholders to prioritise the top 20 topics in gynae-oncology. We asked stakeholders to share via email their top five priority topics, creating a list of 75 topics which were then reduced to 20 at a face-to-face stakeholder workshop held at King's Fund in London. We matched titles with new author teams and provided them with review support from a Statistical Editor/Methodologist. In 2011, we used a similar process but replaced the workshop with a spreadsheet of suggested topics for stakeholders to score in order of priority. From this we prioritized a further 16 new gynae-oncology topics. In 2013, we completed our first Neuro-oncology exercise with two leading Neuro-oncology societies. Using similar methods, we worked with these organizations to prioritise topics and recruit new author teams. Between 2013 and 2015 we were involved in two James Lind Alliance Priority Setting Partnerships, the first focusing on brain and spinal cord tumours and the second on womb cancer. We identified priority topics in both areas, which have been, or are being published as Cochrane Reviews. Alongside this, we have been formulating the best approach to prioritise review updates, analysing the impact and usage of our reviews and producing a tracking system to include this data.

Results: Since 2007 (as a result of prioritization projects), in the area of gynae-oncology, we have published 37 new priority reviews and 22 priority update reviews, with 4 new reviews in development. In neuro-oncology, we have published 7 new priority reviews and we have 3 new priority protocols and 3 new reviews in progress.

Conclusions: By setting priorities we have developed a balanced systematic review portfolio meeting the needs of our stakeholders. We will further this work in 2020 by engaging with stakeholders again to prioritise new reviews and review updates in the area of gynae-oncology.

Patient or healthcare consumer involvement: Our priority setting work is relevant to patients and consumers as it ensures their input as stakeholders is implemented from the start of review production, resulting in published reviews which answer the questions that matter most to them.

Approach and lessons learned from a priority-setting exercise conducted by the Cochrane Breast Cancer Group

Willson M[1], Tomlinson E[2], Wilcken N[1], Goodwin A[1]

[1] Cochrane Breast Cancer Group, The University of Sydney, Australia; [2] Cochrane Cancer Network, UK

Background: Historically, the Cochrane Breast Cancer Group typically registered new topics and updated reviews based on feedback from its international editorial board and funders. In 2019, to comply with new guidance from Cochrane and to continue to meet the needs of its readers, the Breast Cancer Group conducted a formal priority-setting exercise for new and existing topics. As the methods for priority-setting exercises are diverse, determining and executing the right approach was not necessarily a straightforward process.

Objectives: To develop and apply an approach for conducting a priority-setting exercise for Cochrane breast cancer review topics.

Methods: We developed an approach that could be implemented with limited resources, would comply with Cochrane's mandatory standards for priority setting, would facilitate engagement from a wide group of individuals and organizations within the breast cancer community and would allow decisive feedback on the topics that the Cochrane Breast Cancer Group should focus on in 2020.

Results: We used a hybrid approach involving Cochrane's Updating Classification System, citation metrics, conference abstracts and editorial board feedback to develop a preliminary list of 25 breast cancer review topics. These 25 topics were circulated to the breast cancer community in the form of a priority-setting survey and respondents were asked to rank their top 10 topics. This process resulted in nearly 200 responses, 100 complete responses and a clear ranking of breast cancer topics for development or updating. Underlying this process, there were multiple decision points (i.e. the who, what and how) to consider before starting, during and after the priority-setting exercise. The time required for completing this exercise was longer than expected, new tools/resources were found to assist with the process, and warning signs noted if the same process would be repeated.

Conclusions: The priority-setting approach led to the development of a clear set of priority topics for the Cochrane Breast Cancer Group. Careful consideration and clarity of the intended outcome from this exercise is needed to assist with planning the right approach.

Patient or healthcare consumer involvement: Participants of the healthcare system were provided feedback on the approach taken (as part of their role on the editorial board) and responded to the priority-setting survey. The Cochrane Breast Cancer Group is very grateful for the continued support it receives from healthcare consumers during topic selection, the peer-review process and dissemination activities. This project will help to ensure the Group's work continues to be relevant and useful to patients and consumers.

Convening a stakeholder group to prioritise topics across the whole scope of a large review group

Dennett E[1], Stovold E[1], Fortescue R[1]

[1] Cochrane Airways, St George's University of London, UK

Background: Cochrane Airways works with authors (typically healthcare professionals and researchers) to produce systematic reviews on asthma, chronic obstructive pulmonary disease (COPD), bronchiectasis, sleep apnoea, cough, interstitial and other lung diseases. We have published more than 370 reviews and have over 40 protocols in progress. Cochrane Airways has always experienced a tension in dealing with the many different topics in our scope. Asthma and COPD affect a lot of people around the world and cost health services a lot of money. We have done priority-setting work in asthma and COPD previously. But the rarer diseases in our scope are very important for the people living with them and those that care for them, and deserve to be prioritized. While we would like to have enough time to prioritise each section of our scope in detail, it is not practical with our current resources, hence our decision to conduct a "whole-of-scope" prioritization exercise. We also wanted to make our process more transparent and bring in stakeholder voices.

Methods: We convened a group of stakeholders, the Cochrane Airways Priority-Setting Group (CAPSG), to prioritise research questions that will be developed into a series of Cochrane Reviews. We planned to generate research questions in two ways: 1) a one-off survey in 2019; and 2) a rolling priority setting process. We posted and promoted a survey on social media and by email to ask patients, carers, researchers and healthcare professionals for their most important questions about respiratory health. We called the survey "your lungs, your questions".

Results: We received 147 responses to the survey. We cleaned up the responses by removing the uncertainties that were unclear, splitting some longer answers into separate uncertainties and redrafting some for clarity. This resulted in 100 unique uncertainties. The CAPSG comprised 12 people including people living with a chronic lung disease, researchers and health professionals. The CAPSG attended three online meetings and ranked the 100 uncertainties in two rounds. The final outcome was a ranked list of 12 uncertainties covering a range of airways diseases.

Future work: We will carry out further scoping work to develop the uncertainties into questions suitable for Cochrane Reviews. Through the rolling priority-setting program we aim to select priority Cochrane updates and reviews from ongoing work such as literature surveillance, the most highly cited or accessed reviews, reviews identified by guideline groups and review proposals submitted by prospective Cochrane authors. We expect to be able to present the first round of this project at the 2020 Cochrane Colloquium.

Patient or healthcare consumer involvement: Sixty-eight per cent of respondents to the survey identified primarily as a patient and 21% identified as a carer. Two members of the CAPSG identified as patients, one as both a family member and researcher, and at least one of the researchers live with a respiratory condition.

Engaging patients in priority setting: a British case study

Smart P[1]

[1] Cochrane Skin Group, UK

Background: This abstract is based on the presumption that it is important that research on the effects of treatments do not overlook the shared interests of patients, carers and clinicians, otherwise it is probable that many areas of potentially important research will be neglected. Using a case study, the abstract will examine the system of Priority Setting Partnerships (PSPs) developed by the UK James Lind Alliance (JLA) in partnership with research institutions to ensure that programmes focus on the priorities set in partnership with patients and clinicians.

Objectives:

1) To outline the UK system of PSPs established to ensure health care systematic research is focused on an agreed set of research objectives, taking account of the views of all relevant parties.
2) To summarise the benefits and challenges of the PSP process, using the Cellulitis PSP as a case study.
3) To assess briefly with participants the extent to which the UK system might be adapted for other countries/cultures.

Methods:

1) A brief introduction to the process adopted by PSPs and increasing use of PSPs by health research institutions;
2) Using a case study, explain how the process works in practice;
3) A brief review of how the outcomes of the PSP are informing research in practice;
4) A brief guided discussion on how a PSP system might work in the countries/cultures represented at the presentation.

Results: The presentation will include an overview of how the process works, using the outcomes from the Cellulitis PSP as a practical example, including an assessment of the period over which the PSP was active and the resources involved. It will seek to quantify the benefits that have accrued from the PSP.

Conclusions: The outcomes of the PSP are now being used to guide a number of research projects, at the Centre for Evidence Based Dermatology (CEBD), University of Nottingham, which was the lead body for the Cellulitis PSP, and other institutions. The PSP process meets the requirements of priority setting in the UK. Are these benefits transferable to other environments?

Patient or healthcare consumer involvement: The submission relates to abstract topic categories consumer involvement and priority setting. The membership of the PSP steering group included four patient representatives (including the author), alongside six clinicians and four support staff from JLA/CEBD. At the final stage, the number of patients was increased, to ensure an even wider representation of consumer views.

How does the current burden of disease ranking (from the Global Burden of Disease project) compare to that of topics in Cochrane and Prospero's SRs registered in the last five years by region?

Bardach A[1], Glujovsky D[1], Ciapponi A[1]

[1] Instituto de Efectividad Clínica y Sanitaria, Argentina

Background: Only a fraction of the world's resources for health research and development (R&D) are spent on high burden health problems, which represent almost all of the world's preventable mortality. This landscape study will explore the degree of correlation between burden of disease in different parts of the world with the distribution of topics addressed in systematic reviews registered in the two main databases. We hypothesize that systematic reviews may capture only a small proportion of all incident evidence synthesis research, and even this could vary substantially by region.

Objectives: To compare the ranking and frequency of conditions that produce a greater burden of disease according to the Institute for Health Metrics and Evaluation (IHME) Global Burden of Disease (GBD) with the ranking of topics of systematic reviews in the Cochrane Library and PROSPERO databases, in the last five years. We will also explore sources of funding and the evolution of trends.

Methods: We will manually review the rankings of different health causes, diseases or risk factors, and the primary funding source by examining the registered record in the Cochrane Library and PROSPERO, and will compare it to the ranking of burden of disease in terms of Disability Adjusted Life Years and deaths, from the IHME's GBD project according to International Classification of Diseases (ICD-10) codes, in all World Health Organization world regions.

Results: We are still working on this study. Results will be shown at the Colloquium.

Conclusions: Will be shown at the Colloquium.

Patient or healthcare consumer involvement: Patients will benefit if current research reflects priorities given by the high burden diseases, especially patients who are more vulnerable.

Obesity gap analysis and prioritization project

Richardson R[1], Schoonees A[2], Bero L[3], Bickerdike L[1], Chong LY[4], Durao S[5], Metzendorf M[6], Naude C[5], Richter B[6], Wolfenden L[7], Farquhar C[8]

[1] Cochrane Abdomen and Endocrine Network, UK; [2] Department of Global Health, Stellenbosch University, South Africa; [3] Cochrane Public Health and Health Systems Network, Australia; [4] Cochrane Public Health and Health Systems Network, UK; [5] Cochrane Nutrition Field, South Africa; [6] Cochrane Metabolic and Endocrine Disorders Group, Germany; [7] Cochrane Public Health Group, Australia; [8] Cochrane Abdomen and Endocrine Network, New Zealand

Background: Obesity is a significant public health problem and is a risk factor for various diseases with considerable health impacts and costs. The Cochrane Abdomen & Endocrine and Public Health & Health Systems Networks and Nutrition Field are undertaking joint prioritization work. This project was funded by the Cochrane Networks Innovations Fund 2019.

Objectives: Our overall aim is to identify the 'top ten' research priorities in the field of obesity that are the most important for stakeholders. Our specific objectives are to identify gaps in the Cochrane Database of Systematic Reviews by comparing existing titles to research needs identified in clinical guidelines, and to consult with stakeholders to prioritise these gaps.

Methods: This is an ongoing project that will be completed before the Colloquium. We conducted comprehensive searches and screened search records to identify relevant Cochrane titles and clinical guidelines using pre-specified eligibility criteria. We data-extracted the Cochrane titles, including their outcomes, and are busy extracting relevant data, recommendations and research needs identified in each guideline. We will produce a 'gap map' by comparing the Cochrane titles with the identified research needs. Finally, we will identify and consult with key stakeholders using survey methodology to identify the key areas where Cochrane evidence is needed. Stakeholders will include consumers, policy-makers and clinicians.

Results: We identified 41 Cochrane titles that relate to obesity. These were produced by six different Cochrane Review Groups. The Metabolic and Endocrine Disorders Group produced the majority, with 22 titles. Participants included pregnant and breastfeeding women, children, adolescents, adults and older people, as well as healthcare professionals. Interventions were mostly delivered at an individual level (n = 31). We have compiled a set of outcomes for obesity from the included Cochrane titles that will be useful in guiding further research on a core set of outcomes. We identified around 20 relevant guidelines, including ones produced by WHO and several countries such as Canada, Korea and Germany.

Conclusions: We will present all our findings at the Colloquium, including the relevant Cochrane titles and their outcomes, the guidelines and research needs identified, the evidence gap map and the results of the prioritization exercise. We anticipate that this work will provide an invaluable basis for Cochrane's future work in this critical area. Our work will also provide a case study in how to carry out a complex multi-component project across several Cochrane entities.

Patient or healthcare consumer involvement: We plan to include healthcare consumers in our prioritization exercise so that we can identify the research needs that are critical to them.

Outcomes of comprehensive prioritization exercises at Cochrane Skin

Scott H[1], Doney E[1], Prescott L[1], Dellavalle R[2], Williams H[1], Boyle R[3]

[1] Cochrane Skin, University of Nottingham, UK; [2] Cochrane Skin, University of Colorado School of Medicine, US; [3] Cochrane Skin, University of Nottingham, Imperial College London, UK

Background: Prioritization is a key activity of Cochrane Review Groups but one that is approached using a variety of different methods. In order to ensure that Cochrane Skin resources are used where impact is greatest, in 2017 we undertook a comprehensive prioritization exercise involving a wide range of stakeholders, and repeated this in 2020.

Objectives: To present outcomes of Cochrane Skin prioritization in 2017 and 2020 and key learning points.

Methods: Cochrane Skin's 2017 and 2020 prioritization exercises involved patients, guideline developers, Cochrane Review authors and editors. We collated responses from these groups and separately reviewed Global Burden of Disease data for skin conditions, and representation of burdensome skin diseases and download and citation metrics within the Cochrane Skin portfolio. Clinical editors were presented with a summary report and asked to rank proposed titles. We advertised prioritized titles and awarded them to author teams through a competitive selection process based on their skills and available resources.

Results: In 2017 we selected seven titles. Three reviews were already ongoing at the time of the prioritization process, of which two were subsequently removed from the author teams due to inadequate progress. One was re-advertised and allocated to a new team, with the protocol due to be published in 2020; one was replaced with a different, related title already under way; and for the third the protocol was published in 2018 and continued with new team member input. Four titles were awarded to new teams and their protocols were published 9, 10, 11 and 12 months after teams were awarded the project. Although we aimed for submission of reviews within a year of protocol publication for the new titles, this was only achieved for one review, published 16 months after protocol publication. One was submitted 19 months after, one over 2 years and one withdrawn from authors due to inadequate progress. In 2020 we received 168 specific title suggestions from 45 stakeholders.

Conclusions: Timely delivery of priority review titles is challenging. Outside of Cochrane Centres and Review Groups, we identified only limited capacity and resource for timely completion of complex systematic reviews. Despite careful consideration of the skills and commitment of the author teams who applied to take on prioritized titles in 2017, and proactive support and monitoring of authors, delays in review production have been unavoidable. For allocation of 2020 titles, we will give teams more time to work together with the editorial team in developing funding applications to support delivery of complex reviews and will establish a memorandum of understanding, with milestones and break clauses, for review teams who take on a prioritized review title.

Patient or healthcare consumer involvement: We included patients and patient representative organizations in the prioritization process, and author teams were asked to include a consumer author and consider including core outcome sets developed with consumer input.

Prioritization methods in the development and adaptation of health practice guidelines: a systematic survey of published guidance

El-Harakeh A[1], Chamseddine F[2], Itani Z[2], El Khoury R[2], Khabsa J[2], Ghaddar F[1], Akl EA[1]

[1] American University of Beirut, Lebanon; [2] Clinical Research Institute, American University of Beirut Medical Center, Lebanon

Background: Large variation exists in the implementation of prioritization exercises in the development of health practice guidelines. This underlines the importance of exploring available prioritization guidance and assess their level of comprehensiveness.

Objectives: To describe prioritization approaches proposed by a representative sample of guideline-producing organizations for the de novo development and adaptation of guidelines in the clinical, public health, and health systems fields.

Methods: We conducted a systematic survey to identify a comprehensive list of guideline-producing organizations and to compile publicly available methodological documents related to their prioritization processes. Teams of two review authors worked independently and in duplicate to complete eligibility assessment and data abstraction. We collected data on the general characteristics of the organizations and on proposed prioritization steps and criteria. We adopted 11 categories of prioritization steps and used a common framework of prioritization criteria that we developed for a recent systematic review on prioritization approaches in the development of health practice guidelines. We consolidated findings in a semi-quantitative and narrative way.

Results: Our final sample consisted of 114 guideline-producing organizations. Most organizations were professional associations (62%), based in North America (45%), and from the clinical field (83%). While 76% of the identified guidance documents focused on prioritization in the de novo development of guidelines, 5% were on adaptation and 3% were on both. 65% of the guidance documents focused on prioritizing guideline topics, 4% on recommendation questions, and 13% on both, topics and questions. Prioritization of performance measures was addressed in 7% of the documents. For prioritization of topics, the most frequently reported steps were the generation of an initial list of topics (63%), mainly through expert opinion (90%), and the use of prioritization criteria (59%).

Conclusions: This survey included prioritization methods that addressed different prioritization steps for guideline development and adaptation. This can guide the work of researchers, funders, and other stakeholders seeking to prioritize guideline topics and questions.

Patient or healthcare consumer involvement: This research highlights the need to involve patients and consumers in the conduct of prioritization for the de novo development and adaptation of health practice guidelines.

Prioritizing topics for new Cochrane systematic reviews in oral health

Littlewood A[1], Glenny A[1], Walsh T[1], Clarkson J[1]

[1] Cochrane Oral Health, UK

Background: Cochrane Oral Health (COH) is undertaking a priority-setting project, in order to ensure that their reviews are clinically important and relevant to consumers. Part of the project is to identify new review topics that may help to fill gaps in the evidence base, and which answer questions that consumers may have about their oral health.

Objectives: To prioritise topics for new systematic reviews in oral health by: establishing developments in the oral health evidence base by using technology (text mining) to explore clinical trials records for new interventions; identifying evidence gaps from guideline documents; and engaging with consumers to establish their priority topics in oral health.

Methods: We undertook three projects: 1) We explored the evidence base by searching for clinical trials in oral health. We used text mining software to pull out interventions from trial registry records to identify any new interventions; 2) We searched three databases for guidelines in key oral health conditions. We extracted recommendations for further research and examined them for evidence gaps that could be filled by a new systematic review; 3) We launched an online survey so that consumers could suggest any questions that they felt were important for COH to answer. We assessed the results of these three projects to make sure that they were relevant to oral health, and suitable for a Cochrane Review. We then mapped them against COH's existing systematic review portfolio and any topics not covered were established as potential new priority titles.

Results: The three projects undertaken to find potential new priority topics for COH identified 215 questions or topic areas that were not covered by existing COH reviews. Of these, 117 (54%) were deemed to be suitable for a COH systematic review by COH's Priority Setting Steering Group. Of the rejected questions, 25 were not related to Oral Health. The remaining 73 covered questions that are not in Cochrane's remit: for example, establishing the prevalence of a condition. Thirteen questions in total were identified by more than one of the priority-setting projects. These will form part of Cochrane Oral Health's final priority-setting phase, where we intend to involve stakeholders and consumers in an online ranking exercise, looking at both new titles and existing Cochrane Reviews.

Conclusions: These three projects combined revealed significant gaps in the evidence that need to be explored further. This methodology could be adopted in other topic areas as an effective priority-setting method.

Patient or healthcare consumer involvement: Consumers were involved in this project by giving their input in the survey, and their assistance will set the agenda of COH for the next five years.

Priority setting for the topic work participation: collaboration project between Cochrane Work and Cochrane Insurance Medicine

Boschman J[1], Verbeek J[1], Ravinskaya M[1], Verbel Facuseh A[2], Friberg E[3], Kunz R[4], Hoving J[1]

[1] Coronel Institute of Occupational Health, Amsterdam UMC, Amsterdam, The Netherlands; [2] Universitätsspital Basel, Basel, Switzerland; [3] Karolinska Institute, Solna, Sweden; [4] University of Basel, Basel, Switzerland

Background: Cochrane Work is part of the Cochrane Public Health network and operates as an independent producer and publisher of systematic reviews evaluating the effectiveness of interventions aiming to improve workers' health and safety. One of Cochrane Work's topics is that of work participation interventions and work participation outcomes and their implementation in practice. In this project Cochrane Work and Cochrane Insurance Medicine work in collaboration to establish a priority list for intervention reviews that focus on work participation.

Objectives: This prioritization process aims to identify relevant topics for five new reviews and priorities for updates on the topic of work participation.

Methods: Our data collection will consist of the following steps to identify potential topics for systematic reviews: 1) Evidence mapping. We will analyse our review portfolio, published systematic reviews 2010 to 2020, published randomized controlled trials (RCTs) 2015 to 2020, and the Cochrane Library. We will search MEDLINE, Embase, PsycINFO and the Cochrane Central Register of Controlled Trials for guidelines, systematic reviews and RCTs addressing questions about work participation (using terms such as work status, work ability, sickness absence, work disability, employability, employment, return to work). We will extract the research questions and data on population, intervention, control and outcomes. We will remove irrelevant questions if current systematic review(s) already exist and new trials are unlikely; if any topic is beyond our scope, or if the topic is unclear or ill-defined. 2) Stakeholder consultation. We will present the evidence map with potentially relevant topics to a broad range of stakeholders such as: employee organizations, patients, employers' organizations, occupational and insurance physicians, guideline developers, researchers/developers of technology, government authorities, funders. In addition, we will include stakeholders in at least three European countries. They will comment on the question(s) that they would like Cochrane Work to answer in terms of burden of disease, importance for patients, costs or cost savings, influence on research or value for guidelines or policies. 3) Steering group decision. Our steering group will use the Center for Systematic Reviews on Health Policy and Systems Research (SPARK) tool to review the appropriateness and feasibility of the topics and generate a list of review priorities of relevance for the theme work participation. We will evaluate the efforts of this evidence mapping process and specify future strategies for priority setting for Cochrane Work, such as a stakeholder consultation in a round-table conference.

Conclusions: Cochrane Work and Cochrane Insurance Medicine will co-operate in this new project to undertake a prioritization process for the theme work participation. In September 2020 we will have identified topics for five new reviews and priorities for updates.

Patient or healthcare consumer involvement: We will involve employees through employee organizations and patients through collaboration with the Dutch Patient Federation.

Priority setting in the Cochrane Back and Neck Group

Furlan A[1], Pardo Pardo J[2], Harbin S[1], Chou R[3]

[1] Cochrane Back and Neck Group, Canada; [2] Cochrane Musculoskeletal Group, Canada; [3] Cochrane Back and Neck Group, USA

Background: Low back pain (LBP) affects 80% of people at some time in their lives. There are many clinical practice guidelines on the management of low back pain produced by various national and international groups. The Cochrane Back and Neck Group (CBN) has an editorial board composed of internationally renowned scientists, clinicians and consumers.

Objectives: To develop a set of priorities for reviews in the CBN.

Methods: In the past three years, CBN conducted priority setting with organizations that develop clinical practice guidelines for LBP. We contacted seven groups (one in the UK, one in the Netherlands, one in the US, one in Canada, and three in Australia) and were able to engage five. We sent a spreadsheet listing 70 published reviews and asked the teams to highlight the ones we should prioritize for update; we also asked teams to propose new titles. Subsequently, the CBN reviewed the top 10 that were recommended by guideline developers, and also the top 10 reviews most accessed and cited in the Cochrane Library in 2017. We selected the reviews that were common across at least two of these lists. Finally, the CBN editorial board decided on broad topic areas and other areas of clinical interest to finalize priorities.

Results: The priority list has been useful in guiding the acceptance or rejection of new titles submitted to CBN. Since 2018, CBN has received five proposals from teams interested in starting a Cochrane Review, and rejected all of them because they did not match a priority set by our group. However, the current list of priorities does not include the opinions of people directly impacted by Cochrane Reviews such as patients with LBP, their caregivers, their families, and frontline clinicians.

Conclusions: Our current process of prioritizing reviews relies on our 2017 list of priorities and the subjective opinions of the members of our editorial board. This list needs to be updated on a regular basis.

Patient or healthcare consumer involvement: There is a need to establish a diverse stakeholder group composed of people living with chronic LBP, patient advocates, frontline clinicians, provincial health authorities, and charity organizations to set priorities for reviews in the CBN.

Priority setting within Cochrane Haematology

Piechotta V[1], Jakob T[2], Tomlinson E[3], Ocheni S[4], Theurich S[5], Estcourt LJ[6], Skoetz N[7]

[1] Cochrane Haematology, Department I of Internal Medicine, Center for Integrated Oncology Aachen, Bonn, Cologne, Dusseldorf, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany; [2] Cochrane Haematology, Department of Internal Medicine, Center for Integrated Oncology Aachen, Bonn, Cologne, Dusseldorf, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany; [3] Cochrane Cancer Network, Royal United Hospital, Bath, UK; [4] Department of Haematology and Immunology, University of Nigeria, Ituku-Ozalla Campus, Enugu, Nigeria; [5] Department of Medicine III, University Hospital LMU, Ludwig-Maximilians-Universität München, Munich, Germany; [6] Cochrane Haematology, Haematology and Transfusion Medicine, NHS Blood and Transplant, and Radcliffe Department of Medicine, University of Oxford, Oxford, UK; [7] Cochrane Cancer, Department I of Internal Medicine, Center for Integrated Oncology Aachen, Bonn, Cologne, Dusseldorf, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany

Background: Priority setting for systematic reviews is just as important as prioritizing primary research. It ensures that systematic reviews address highly relevant research questions and thus meet the needs of those affected. Involving stakeholders, such as patient representatives, healthcare providers and guideline co-ordinators in prioritization processes helps to concentrate limited research resources on answering the important questions and to reduce research waste. In the past, reviews developed within the Cochrane Haematology Group have mainly been projects commissioned by the German government, grants from the National Institute for Health Research, or priorities highlighted by national and international bodies, e.g. the World Health Organization. However, up to now a structured prioritization process was missing.

Objective: To identify five new systematic review questions and 10 existing systematic reviews that need updating in the area of Haematology, considering potential differences between high-income countries and low-and middle-income countries, by implementing a structured prioritization process.

Approach: 1) We formed a priority setting steering group, involving the Group's Co-ordinating Editor, Managing Editors, Clinical Experts from high-income countries and low- and middle-income countries, the Network Senior Editor, and the Network Support Fellow. 2) To assess the relevance of our existing reviews, we analysed citation metrics, guidelines that cited our reviews, and download metrics of all the reviews in our Group's portfolio. We used our findings to develop a list of potential priority topics and reviewed their up-to-dateness. 3) To identify evidence gaps, we reviewed research roadmaps of regional and global advocacy groups and societies, checked the prioritized topics of the James Lind Alliance and consulted our Group's editors, patient representatives, and guideline co-ordinators. 4) Through a stakeholder mapping exercise done by the Network Support Fellow, we identified a broader community of external stakeholders, such as healthcare consumers and providers, and policy makers. We collated both lists in a survey and circulated it to the identified stakeholders. 5) Following the analysis of the survey results, we will announce the prioritized topics on our Group's website and distribute them via Cochrane channels to attract interested author teams. This process will be repeated every three years.

Expected outputs: Cochrane Haematology will publish at least two prioritized review updates and one new priority review per year.

Patient or healthcare consumer involvement: Our group is working with external stakeholders, such as patient representatives, healthcare providers and guideline co-ordinators to prioritise review topics within their existing portfolio and to identify new topics of high relevance. Involving stakeholders ensures that reviews developed within our group are of utmost importance. Working together will lead to the effective dissemination of the evidence and increase impact and understanding by those who need it.

Setting priorities for Cochrane Reviews of Multiple Sclerosis

Celani MG[1], Nonino F[2], Filippini G[3], D'Amico R[4], Baldin E[2], Mahan KM[5]

[1] Cochrane Neurological Sciences Field; Department of Neurophysiopathology, Perugia Hospital, Italy; [2] Affiliate of Cochrane Italy; IRCCS – Istituto delle Scienze Neurologiche di Bologna, Italy; [3] Cochrane Multiple Sclerosis and Rare Diseases of the CNS Group; Fondazione IRCCS Istituto Carlo Besta, Italy; [4] Cochrane Italia Dipartimento di Scienze Mediche e Chirurgiche Materno Infantili e dell'Adulto Università degli Studi di Modena e Reggio Emilia Unità Semplice di Supporto Metodologico Statistico alla Ricerca Clinica Azienda Ospedaliero-Universitaria di Mo, Italy; [5] Cochrane Neurological Sciences Field; Umbria Region Health Authority, Italy

Background: Cochrane Multiple Sclerosis and Rare Diseases of the CNS Group (MSCRG) in collaboration with Cochrane Neurological Sciences Field (CNSF), Cochrane Knowledge Translation Group (KTG) and the Affiliate of Cochrane Italy Geographic Group decided to start an international consultation process aimed at identifying five new topics for upcoming Cochrane Reviews on MS.

Material and methods: The consultation will be performed according to the Guidance Note by the KTG.

Stage 1. An Advisory Group (AG) will be established and will lead the priority setting process. Five core areas of potential uncertainty (diagnosis, health communication and participation, pharmacological treatments, rehabilitation and social support interventions), identified through an appraisal of clinical pathways for MS, will be considered. Specific topics with potential uncertainty will be highlighted within each area, and a tentative list of 10 top research questions will be defined by the AG and ranked according to their answerability and importance.

Stage 2. We will use an international online survey to collect other questions on relevant aspects of MS care. Persons with MS and their carers, clinicians and researchers will be invited to participate in the survey through the help of international patient 'advocacy associations', Cochrane, MSCRG, CNSF and MS scientific societies. CNSF will refine the questions and group them together into the core areas. At the end of this process the AG will identify 10 new research questions.

Stage 3. We will run a second online survey submitting the 20 research questions and ask participants to rank them according to their relevance and expected impact of the review results on clinical practice and benefit to persons with MS.

Stage 4 CNSF will consider the top 10 research questions prioritized through the survey and compare them with the existing MSCRG review portfolio.

Stage 5. In order to rule out the already answered questions we will perform a systematic review of clinical practice guidelines (CPGs) on the diagnosis and treatment of MS published since 2015 in English using AGREE II for quality assessment.

Stage 6. Up to five priority Cochrane Reviews to be newly produced or updated will be identified.

Patient or healthcare consumer involvement: We will involve persons with MS and their carers, clinicians and researchers in the international survey.

Updating gap in Cochrane systematic reviews: preliminary results of an observational pilot study

Antequera Martín A[1], Solà I[2], Bonfill X[3]

[1] Universitat Autònoma de Barcelona, Centre-Biomedical Research Institute Sant Pau, Barcelona, Spain; [2] Iberoamerican Cochrane Centre, Biomedical Research Institute Sant Pau-CIBER of Epidemiology and Public Health (CIBERESP-IIB Sant Pau), Barcelona, Spain; [3] Iberoamerican Cochrane Centre, Biomedical Research Institute Sant Pau-CIBER of Epidemiology and Public Health (CIBERESP-IIB Sant Pau), Universitat Autònoma de Barcelona, Spain

Background: Cochrane systematic reviews (SRs) are the cornerstone of setting clinical recommendations, which requires a rigorous developing and updating process. Although Cochrane's policy defines that Cochrane Reviews should be assessed for updating within two years of publication, previous studies suggest that a minority of Cochrane Reviews were updated in the five years after their publication. Musculoskeletal conditions are the leading contributor to disability worldwide (data from Global Burden Disease 2017 and WHO). The Musculoskeletal Cochrane Review Group determined the priority setting, and for some of their topic areas, including osteoporosis and rheumatoid arthritis, they identified priority topics. We conducted a pilot study for monitoring the impact of updating gaps on osteoporosis and rheumatoid arthritis SRs.

Objectives: We described the updating status for Cochrane SRs of osteoporosis and its relation with the response from non-Cochrane organizations in terms of recency of their reviews.

Methods: Observational pilot study. We searched reviews on the populations of interest in the Cochrane Database of Systematic Reviews and selected those that were out of the updating threshold set in current policies (three years). We matched each Cochrane Review question with a search in PubMed for retrieving more recent SRs using the fields of osteoporosis (according to their most representative subject heading combined with a text term searched at the title field). We globally searched and then limited the search to the core journals set in PubMed and the 15 top journals in the Scimago ranking for Rheumatology. We collected data on the topic, publication year, search strategy period, and the number of inclusions into guidelines. We performed descriptive analyses and compared temporal trends over time.

Results: We identified 219 SRs of interventions within the Musculoskeletal Cochrane Review Group. Almost a quarter of the total (24.59%) was published in or after 2014, just over a third (33.34%) were for osteoarthritis and 6.4% (N = 14) for osteoporosis. Six reviews addressed bisphosphonates therapy, four reviews focused on calcium and vitamin D treatments, and the remainder assessed other interventions. We found 49 non-Cochrane SRs published after the Cochrane SR publication date, ten of them published in leading journals. Regarding changes in the conclusions, 21% of non-Cochrane SRs suggested a shift in the conclusion. Extracting data on rheumatoid arthritis is ongoing.

Conclusion: The updating process for Cochrane SRs of osteoporosis is poor. Osteoporosis Cochrane Reviews are not covering the body of evidence in progress or completed.

Patient or healthcare consumer involvement: None.

Will you fund my idea? A retrospective look at public nominations to the Agency for Healthcare Research and Quality's Evidence-based Practice Center program

Huppert J[1], Banez L[1], Iyer S[1], Lucero J[1], Chang S[1], Chang C[1]

[1] Agency for Healthcare Research and Quality (AHRQ), USA

Background: Since 2007, the Agency for Healthcare Research and Quality (AHRQ) Evidence-based Practice Center (EPC) program has provided systematic reviews (SRs) in response to public nominations. AHRQ prioritizes which new SRs to fund via structured assessment and selection.

Objectives: To improve our prioritization process by evaluating outcomes of nominations from groups using the AHRQ public process.

Methods: We retrospectively analyzed our 2007 to 2017 nomination database for the number and proportion of public nominations received and funded by type of nominator. Due to changes in EPC Program resources, we evaluated changes in two periods (2007 to 2012 vs. 2013 to 2017).

Results: We received 783 public nominations and selected 159 (20%) for funding. Nominators self-identified as Individual (40.4%; researchers, healthcare professionals, patients, consumer advocates); Professional society, (17.4%) and Other stakeholders (42.2%; payer, industry, non-governmental/non-profit, academic). Professional societies contributed 17% of nominations and 38% of funded SRs (Table 1). Over time, the number of nominations decreased, the number and proportion from Professional Societies increased (from 10% to 34%). Those from Others decreased (from 53% to 16%). For Individuals, the absolute numbers decreased, but the proportion increased (from 36% to 50%; P < 0.01 for each comparison). Total funding rates were higher for Professional society (44%) than Other (20.5%) or Individual (9.8%) nominations, (P < 0.01); this pattern held in both periods. For all groups, the most common reason for not funding was overlap with a recent SR (36%). Other reasons varied by group. Professional societies were less likely to submit topics that were outside of the EPC program scope (1%) or not feasible for SR (too few primary studies) (18%). Individuals' nominations were more likely to be outside of the EPC program scope (29%). Other nominations were more likely to be viable for SR but not prioritized for funding (58%). Over time, the proportion of total nominations that were viable but not funded increased significantly, from 3.3% to 23.5%, reflecting a change in EPC Program resources.

Conclusions: A wide array of nominators use AHRQ's public nomination process to secure SRs. Professional society use increased over time, which may reflect a response to 2012 Institute of Medicine standards for producing evidence-based guidelines. Professional societies are a valuable partner for AHRQ, as they nominate topics that: meet the EPC Program's selection criteria, are amenable to SR, and can accelerate the impact of AHRQ work. While almost a third of Individual nominations are out of scope (limited detail or not appropriate for the EPC program), still about 10% are funded. Results suggest that AHRQ should continue to accept Individual nominations, while providing better submission guidance, and must investigate why some viable topics were not funded for Other groups.

Patient or healthcare consumer involvement: Individual nominations include patient and consumer perspectives.

Additional file: Table 1

STAKEHOLDER INVOLVEMENT

A model for timely synthesis and sharing of evidence to meet health system decision-maker needs

Tricco A[1], Chaudhry S[1], Zarin W[1], Abou-Setta A[2], Curran J[3], Clement F[4], Holmes B[5], LeBlanc A[6], Rodin R[7], Downey M[8], Straus S[1]

[1] Li Ka Shing Research Institute, Knowledge Translation Program, St. Michael's Hospital, Canada; [2] George & Fay Yee Centre for Healthcare Innovation, University of Manitoba, Canada; [3] IWK Health Centre, Halifax, Nova Scotia, Canada; [4] Health Technology Assessment Unit, O'Brien Institute of Public Health, University of Calgary, Canada; [5] Michael Smith Foundation for Health Research, Canada; [6] Department of Family and Emergency Medicine, Faculty of Medicine, Université Laval, Canada; [7] Public Health Agency of Canada, Canada; [8] British Columbia SUPPORT Unit, Canada

Background: The Strategy for Patient-Oriented Research (SPOR) Evidence Alliance (hereafter, the Alliance) is a pan-Canadian initiative designed to promote evidence-informed healthcare and policy that is grounded in the principles of inclusion, diversity, integrated knowledge translation, co-creation of research, transparency, and reduction of research waste. The Canadian Institutes of Health Research (CIHR) funded the Alliance upon recognizing that a concerted, well-resourced, and collective, patient-oriented approach was needed to increase capacity within the health landscape to synthesize, disseminate and integrate research results more broadly into healthcare decision-making and clinical practice.

Objectives: To provide an overview of the Alliance and demonstrate how stakeholders (e.g. patients, policy makers) are meaningfully engaged in all Alliance activities.

Methods: The Alliance is governed by six standing committees to advise on the operation, sustainability, and evaluation of its activities. A key activity of the Alliance is to promote increased use of evidence in health system decision-making by offering a research query service. This service permits stakeholders to submit their health research needs for knowledge synthesis, guidelines, and knowledge translation. An Alliance-affiliated research team then co-produces research findings with input from relevant stakeholders. Collaborative tasks often include creation of the study protocol, and participation in research conduct and dissemination activities.

Results: The Alliance currently includes 250 members who are researchers (60%), trainees (13.2%), policy-makers/healthcare providers (18.4%), and patients (8.4%). Policy-makers/healthcare providers and patients hold 16.4% and 13.7% (respectively) of 73 seats within the governance structure. To date, the Alliance has received 75 queries of which 24 successfully completed queries involved collaboration with 54 knowledge users. The completed queries all addressed a need in knowledge synthesis. Three-quarters of these came from government, 8.0% from another SPOR entity, 8.0% from a guideline developer, 4.0% from a health system manger/care provider, and 4.0% from a multinational health agency. Results from completed queries are anticipated to inform policy or practice by organizations that have a provincial reach (79.2%), national reach (16.7%), or international reach (4.2%).

Conclusions (patient/healthcare consumer involvement): The strength of the Alliance lies in the meaningful engagement of stakeholders (e.g. patients, health system decision-makers) in all activities. Stakeholders provide input into the research agenda and strategic direction of the Alliance, and engage as meaningful partners in the research process. This model of collaboration ensures that efforts to address health research priorities of populations and health systems are appropriately informed and purpose-driven, thereby reducing inefficiencies and waste.

An evidence-informed and stakeholder co-designed physical and community mobility intervention for older adults facing health inequities: the EMBOLDEN study

Neil-Sztramko S[1], Teggart K[1], Phillips S[1], Sherifali D[1], Fitzpatrick-Lewis D[1], Newbold B[1], Coletta G[1], Alvarez E[1], Kuspinar A[1], Raina P[1], Ganann R[1]

[1] McMaster University, Canada

Background: Physical mobility and social participation are requisite for independence and quality of life as one ages. Barriers to mobility lead to social isolation, poor physical and mental health, all of which are precursors to frailty. To date, most mobility-enhancing interventions for older adults have been designed by researchers with little-to-no consultation with older adults or service providers, and delivered in controlled settings. Thus, translation and scalability is often impractical.

Objectives: Using an evidence-informed public health approach, the purpose of EMBOLDEN (Enhancing physical and community MoBility in OLDEr adults with health inequities using commuNity co-design) is to combine research evidence, local context, and community needs and preferences to co-design an intervention targeting physical and community mobility, improving diet quality, social participation and system navigation.

Methods: We used an environmental scan that identified four neighborhoods with high health inequity to map existing relevant health and social programs and identify gaps. In parallel, we conducted an umbrella review to identify high-quality evidence of effective interventions targeting physical activity and/or nutrition in community-dwelling older adults. We searched five databases using a comprehensive search strategy. Two review authors independently performed title/abstract screening, full text review, critical appraisal (using AMSTAR2) and data extraction. We compiled key findings across reviews and will report these to key stakeholders to inform intervention design. We will used qualitative persona-scenario interviews with older adults and service providers to understand experiences, gaps, and priority features.

Results: The environmental scan identified key assets and gaps to be explored in the co-design process. The umbrella review identified 34 systematic reviews describing heterogeneous group-based physical activity and multi-modal interventions. No reviews focused on nutrition only. We compiled review findings to identify the best available evidence of effectiveness across five categories of outcomes: aerobic capacity, muscular strength, physical function, balance, and falls. Qualitative data collection and analysis are underway. Our next aims are to work with partners to co-design the evidence-informed intervention, implementation and evaluation plan.

Conclusions: Collaborating with community partners and using multiple sources of evidence helps to bring important contextual information to the design of a community-based intervention. Our goal is to implement a feasible intervention that meets the needs of the community with potential for long-term sustainability.

Patient or healthcare consumer involvement: Older adults and local health and social service providers are members of our research team and helped to prepare the research proposal. A larger group of stakeholders are engaged monthly as members of a Strategic Guiding Council and provide important input into each phase of the research.

Co-design, co-production and acceptability of smartphone mood monitoring app for young people with bipolar disorder

Barron Millar E[1], Kaushal P[2], McGurk K[1], Glod M[1], Sharma A[3]

[1] Newcastle University, UK; [2] Cumbria, Northumberland, Tyne and Wear NHS Foundation Trust, UK; [3] National Specialist Adolescent Mood Disorders Service, UK

Background: Bipolar Disorder (BD) has a global prevalence of 2%. BD in young people is associated with greater illness severity and poorer outcomes. Mood symptom monitoring is an important part of clinical management. Pen and paper charts are the current main approach to mood monitoring in this population, but completion rates are low and review of data is time consuming. Mood monitoring using digital technologies such as smartphone apps could reduce costs, improve accessibility and increase reliability of self-management, but there is currently no evidence to support this in a population of young people with BD.

Objectives: To co-design and co-produce a smartphone mood monitoring app to improve outcomes for young people with BD in a UK tertiary care service.

Methods: Participatory methodology to co-design and co-produce a mood monitoring app. Tertiary care patients aged 16 to 25 years with a DSM-5 diagnosis of BD, and their families, were recruited. In the Co-Design Phase we held weekly workshops with six young people and their families to identify themes of self-management. In the co-production phase the alpha-prototype of the app was developed by young people, clinicians, researchers, graphic designers and software engineers. We used a mixed methods prospective study (n = 13) to evaluate the acceptability of the alpha-prototype to young people with BD. We used the Warwick Edinburgh Mental Wellbeing Scale (WEMWBS) to measure mental well-being, the Satisfaction Usability Tool (SUT) to capture overall satisfaction with the app and we conducted qualitative semi-structured interviewing.

Results: Six themes of self-management (socialization, learning, relaxation, excitement, regulation and distraction) were incorporated in the alpha-prototype of the app. Patients requested a sliding scale, rather than Likert scale, for rating their mood on dimensions such as happiness, sadness and irritability. We developed a graphical representation of trends in mood to feedback to users about their own mood trends and a comment section was added so that users could note any triggers associated with their mood fluctuation. A 'lifeline' feature was developed, for use in crisis (e.g. suicidal ideation) so users could press one button to call through a list of emergency contacts until someone answered. Scores on the WEMWBS and SUT improved after app use. Interview data suggested the app was well received, but patients requested more personalization and interactivity. The lifeline function was used twice.

Conclusions: The co-design and co-production of a mood monitoring smartphone app was iterative in nature and maintained the patient focus by incorporating patient suggestions into the development of the app. The reported acceptability and usability of the app suggest that a smartphone mood monitoring app could be used clinically to improve engagement with mood monitoring and improve cost-effectiveness of mood monitoring strategies

Patient or healthcare consumer involvement: Patients were involved in co-design and co-production of the app.

Development of an evidence-based prediction model for medical usage at mass gatherings: interaction of evidence, expertise, and end-user demands

Scheers H[1], Van Remoortel H[1], De Buck E[2], Vandekerckhove P[3]

[1] Centre for Evidence-Based Practice, Belgian Red Cross, Belgium; [2] Centre for Evidence-Based Practice, Belgian Red Cross; Department of Public Health and Primary Care, Faculty of Medicine, KU Leuven; Cochrane First Aid, Belgium; [3] Belgian Red Cross; Department of Public Health and Primary Care, Faculty of Medicine, KU Leuven, Belgium

Background: Every year, volunteers of the Belgian Red Cross (BRC) provide preventive medical aid at more than 8000 mass gatherings and smaller events. For optimal use of resources (personnel, materials, and money), it is important to be able to predict patient load and healthcare needs at these events. The BRC's Medical Triage and Registration Informatics System (MedTRIS), containing data on more than 200,000 interventions at mass gatherings during the last 10 years, is a valuable source of information to build a predictive model on.

Objectives: To develop and validate a prediction model of patient presentation rate (PPR) and transfer to hospital rate (TTHR) at mass gatherings in Belgium, based on the three pillars of evidence-based practice (EBP): scientific evidence, in-house expertise and experience of health professionals and volunteer representatives, and preferences and resources of the target group (BRC volunteers).

Methods: We conducted a systematic review to identify potential environmental and event-specific predictors of PPR and TTHR at mass gatherings. Subsequently, we developed a nonlinear prediction model, containing these variables and using regression trees, with a subset of 28 mass gatherings in MedTRIS, and validated the model with another subset. Throughout the project, we regularly met with experienced volunteer representatives (e.g. first aid responders) and health professionals (e.g. emergency physicians), and made field visits to specific mass gatherings to better understand the practice of preventive first aid, peculiarities of the data, and preliminary findings of the model.

Results: We selected 12 potential predictor variables for our model from 16 studies identified by the systematic review. Five of these variables were predictive for PPR in the regression tree: number of days and type of the event, number and age distribution of attendants, and temperature. Internal validation of the model revealed an R^2 of 0.69. External validation indicated limited predictive value for some mass gathering types ($R^2 = 0.30$). We obtained similar results for TTHR. The meetings and field visits helped identifying strengths and weaknesses in the underlying database and resulted in recommendations to further optimize data collection and analysis, which will improve the predictive power of the model.

Conclusions: Following EBP principles, we were able to develop and validate an evidence-based prediction model, which was further finetuned by consulting different stakeholders. Implementation of the prediction model will ultimately lead to a better use of resources at preventive aid actions by the BRC.

Patient or healthcare consumer involvement: This prediction model was developed in collaboration with the Relief Service at the Belgian Red Cross, which co-ordinates the preventive aid campaigns at mass gatherings in Flanders (Belgium). Patient and healthcare needs at mass gatherings were summarized by BRC volunteer representatives, specialized health professionals, and volunteering first aid responders we met during the field visits.

Engaging patient partners in systematic reviews: the example of prevalence of persistent postsurgical pain after breast cancer surgery

Wang L[1], Cohen J[2], Devasenapathy N[3], Hong B[4], Kheyson S[5], Lu D[2], Oparin Y[2], Kennedy S[6], Romerosa B[7], Couban R[1], Busse J[1]

[1] Department of Anesthesia, McMaster University, Canada; [2] Michael G. DeGroote School of Medicine, McMaster University, Canada; [3] Indian Institute of Public Health, Delhi, India; [4] Division of Plastic Surgery, University of Toronto, Canada; [5] Department of Family Medicine, Cumming School of Medicine, Canada; [6] Department of Medical Imaging, University of Toronto, Canada; [7] Department of Anesthesia and Critical Care, University Hospital of Toledo, Spain

Background: Persistent postsurgical pain (PPSP) is common after breast cancer surgery; however, the reported prevalence varies widely. Some of the variability in the estimates of PPSP after breast cancer surgery is likely due to how pain is reported (e.g. severity, location). We conducted a systematic review to inform this uncertainty and engaged patient partners to help design our protocol.

Objectives: We aimed to engage patient partners, who were breast cancer survivors, to guide the approach and interpretation of a systematic review regarding the prevalence of PPSP after breast cancer surgery.

Methods: Patient partners were involved in the design and interpretation of the systematic review. We searched MEDLINE, Embase, CINAHL, and PsycINFO from inception to November 2018, for observational studies reporting the prevalence of PPSP (≥ 3 months) after breast cancer surgery. We used random-effects meta-analysis and multivariable meta-regression for PPSP prevalence based on patient's preference. We used the GRADE approach to rate quality of evidence, and our patient partners defined a patient-important threshold for the risk of persistent pain for assessing imprecision.

Results: There were 146 observational studies (137,675 patients) eligible for review that reported the prevalence of PPSP (ranging from 2% to 78%). The pooled prevalence of PPSP of any severity was 35% (95% confidence interval (CI) 32% to 39%). Moderate-quality evidence supported subgroup effects of PPSP prevalence for localized pain vs. any pain (29% vs. 44%), moderate or greater vs. any pain (26% vs. 44%), and clinician-assessed vs. patient-reported pain (23% vs. 36%) (Figure 1, Table 1). Two breast cancer patients were involved in this review; both endorsed that PPSP should be based on patient report, and in any location, but were divided on whether 'any pain' or 'moderate-to-severe pain' was most important. Based on patients' preference, multivariable meta-regression found the prevalence of patient-reported PPSP following breast cancer surgery of any severity or location was 46% (95% CI 36% to 56%), and the prevalence of patient-reported moderate-to-severe PPSP at any location was 27% (95% CI 10% to 43%) (Table 2). Our patient partners suggested the following rates of PPSP as important: 20% for any pain, 10% for moderate pain, and 5% for severe pain.

Conclusions: Moderate-quality evidence suggests almost half of all women undergoing breast cancer surgery develop PPSP, and approximately one in four develop moderate to severe PPSP. Involvement of patient partners in the design of systematic reviews may help ensure that results are reported in a way that is most helpful to patients.

Patient or healthcare consumer involvement: Patient partners were involved in the design and interpretation of systematic reviews.

Additional files: Tables and figures

Engaging small businesses in the practice of health promotion through workplace wellness programs

McIntosh B[1], Adepitan K[1]

[1] ACE Community Health, Canada

Background: The practice of health promotion involves engaging stakeholders from multiple sectors to help individuals in communities make better choices regarding their wellness. The practice of engaging small businesses in health promotion is of a particular importance given a major public health crisis: the COVID-19 pandemic. Collectively, small businesses employ a significant number of people, yet they have adopted fewer wellness programs than large corporations.

Objectives: To assess evidence around key issues with the engagement of small business employers in the adoption of workplace wellness programs.

Methods: We searched the Cochrane Library and PubMed databases for studies published before January 2020. One author independently screened the articles for relevance and two authors selected the articles for inclusion. One researcher extracted data for evaluation by two researchers. We assessed studies using GRADEpro software.

Results: Of the 95 studies, 15 met the inclusion criteria. Employee wellness is essential within small business communities and broader communities. Costs and a lack of support around designing, implementing, and evaluating a wellness program were identified as critical issues for small business employers. Partnering with a third-party practice that holds expertise in this field can provide support to small businesses, while helping to manage the costs associated with delivering a wellness program.

Conclusions: Barriers that limit small business employers from engaging in health promotion by providing a workplace wellness program are related to costs and a lack of expertise. A reported strategy that can be used to mitigate these key issues involves using a third-party practice.

Patient or healthcare consumer involvement: Stakeholder engagement can be used to engage the small business community in health promotion.

Enhancing uptake of systematic reviews: what is the best format for healthcare managers and policy makers?

Marquez C[1], Park J[1], Somji I[1], Blaine C[2], Chignell M[3], Ellen ME[4], Graham ID[5], Hayes A[6], Hamid JS[7], Hemmelgarn B[8], Hillmer M[9], Holmes B[10], Holroyd-Leduc J[11], Hutton B[5], Kastner M[9], Lavis JN[12], Moher D[5], Ouimet M[13], Perrier L[9], Proctor A[6], Noseworthy T[8], Schuckel V[14], Stayberg S[15], Tonelli M[11], Tricco AC[1], Straus SE[1]

[1] Li Ka Shing Knowledge Institute, St. Michael's Hospital, Unity Health Toronto, Canada; [2] BMJ Knowledge Centre, UK; [3] Department of Medicine, Faculty of Medicine, University of Toronto, Canada; [4] Ben Gurion University, Israel; [5] Ottawa Hospital Research Institute, Canada; [6] Ontario Ministry of Health and Long-Term Care, Canada; [7] Children's Hospital of Eastern Ontario, Canada; [8] Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Canada; [9] Institute of Health Policy, Management and Evaluation, University of Toronto, Canada; [10] Michael Smith Foundation for Health Research, Canada; [11] Department of Medicine, Cumming School of Medicine, University of Calgary, Canada; [12] McMaster Health Forum, Department of Health Research Methods, Evidence and Impact, and Department of Political Science, McMaster University, Canada; [13] Laval University, Canada; [14] Ministry of Health, Canada; [15] Alberta Health, Canada;

Background: Despite advances in reporting and valuing systematic reviews, they are used infrequently by healthcare managers (HCMs) and policy makers (PMs) in decision-making.

Objectives: To compare the impact of a novel systematic review format with the traditional systematic review format on the ability of HCMs and PMs to understand the evidence in the review and apply it to a relevant healthcare decision-making scenario.

Methods: We conducted two parallel randomized controlled trials (RCTs) with HCMs and PMs in Canada and the UK. HCMs and PMs were randomized separately to the intervention (novel format) or the control group (traditional format). The primary outcome was the proportion of HCMs or PMs who appropriately considered and applied the evidence from each systematic review format to the scenario presented.

Results: A total of 257 participants completed the RCTs (100 HCMs and 157 PMs). 18.6% were from Alberta, 16.0% from British Columbia, 57.6% from Ontario, 3.9% from Quebec, 0.4% from UK, and 3.5% from other provinces. Of the 157 PMs that participated in the RCT, 31.2% worked in their field for 6 to 10 years, 62.4% received training in research methodologies, and 41.4% were mostly familiar with systematic reviews. Of the 100 HCMs, 27.9% worked in their field for over 20 years, 59.0 % received training in research methodologies, and 41.0% were mostly familiar with systematic reviews. Final results of the RCTs will be available at the Cochrane Colloquium 2020.

Conclusions: This study is the first to engage HCMs and PMs from multiple settings in assessing the impact of novel formats on the use of systematic review results compared to the traditional format. The results of this study will help increase the uptake of systematic review results in healthcare management policy making decision-making, ultimately leading to informed decision-making and positively impacting the health of Canadians.

Trial Registration: ClinicalTrials.gov NCT03041454

EVOLVE: a framework for meaningful patient involvement in clinical practice guideline development and implementation

Bjorkqvist J[1], MacLennan S[1], Giles R[2], Cornford P[3], Plass K[4], Van Poppel H[5], Makaroff L[6], Van Hemelrijck M[7], Cowl J[8], MacLennan S[9]

[1] Academic Urology Unit, University of Aberdeen, UK; [2] International Kidney Cancer Coalition, The Netherlands; [3] Royal Liverpool University Hospitals Trust, UK; [4] Guidelines Office, European Association of Urology, The Netherlands; [5] European Association of Urology, Belgium; [6] Fight Bladder Cancer and World Bladder Cancer Patient Coalition, UK; [7] Translational Oncology & Urology Research, King's College London, UK; [8] Public Involvement Programme, National Institute for Health and Care Excellence, UK; [9] Academic Urology Unit, University of Aberdeen and Urological Cancer Charity, UK

Background: The importance of stakeholder involvement, including patients, in the development of clinical practice guidelines is internationally recognised. Patients' unique experience of living with the condition and its consequences gives them a lived perspective, which could contribute to improvements in the quality and relevance of guidelines. However, there is a lack of methodology and evaluation around patient involvement in guideline development. The European Association of Urology (EAU) Guidelines Office have international reach, robust guideline development processes, an established patient information section and existing links to global cancer patient organizations, making the EAU and genitourinary cancers an ideal proving ground to test an integrated framework of patient involvement in guideline development.

Objectives: This study aims to develop a framework of meaningful patient involvement to address: which stakeholders to involve, how to involve them, at what stage of the process, and to propose a stakeholder involvement evaluation tool.

Methods: First, we systematically reviewed existing models of patient involvement in guideline development. Next, we conducted semi-structured interviews with patient and clinician members of European genitourinary cancer Guideline Panels, to assess barriers and facilitators for patient involvement. Then, patients and clinicians scored a list of topic areas for considering patient involvement for importance via an international Delphi survey preceding a face-to-face consensus meeting. Finally, we designed a framework of patient involvement in guideline development and implementation based on evidence from the systematic review, interviews with key stakeholders and the Delphi and consensus process.

Results: Sixteen priority areas and technical processes for patient involvement were identified via our Delphi survey and consensus meeting and these were embedded within the EVOLVE framework. The final EVOLVE framework includes recommendations for both social and technical guideline processes. These recommendations include increasing patient understanding of the processes for guideline development through focused recruitment and provision of adequate training and support; establishing a patient advisory board; and providing access to patient networks and mechanisms for feeding in the preferences and values of wider patient groups. This framework will be tested for genitourinary cancers in Europe within the 2020-2021 guideline development cycle. Normalization Process Theory will be used to evaluate the implementation and integration of the framework and inform an evaluation tool.

Conclusions: The final EVOLVE framework provides guideline developers with clear methodology, including both social and technical processes, for patient involvement in guideline development and implementation. The next step is to test the EVOLVE framework within genitourinary cancers in Europe for the 2020 to 2021 guideline development cycle.

Patient or healthcare consumer involvement: Patient-led study.

Improving vaccination uptake among adolescents

Abdullahi L[1], Kagina B[1], Ndze V[2], Hussey G[1], Wiysonge C[3]

[1] Vaccines for Africa Initiative, Institute of Infectious Disease and Molecular Medicine, University of Cape Town, South Africa; [2] Centre for Evidence-based Health Care, Faculty of Medicine and Health Sciences, Stellenbosch University, South Africa; [3] Cochrane South Africa, South African Medical Research Council, Cape Town, South Africa

Background: Adolescent vaccination has received increased attention since the Global Vaccine Action Plan's call to extend the benefits of immunization more equitably beyond childhood.

Objectives: To evaluate the effects of interventions to improve vaccine uptake among adolescents

Methods: In October 2018, we searched the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE Ovid, Embase Ovid, and eight other databases. In addition, we searched two clinical trials platforms, electronic databases of grey literature, and reference lists of relevant articles. For related systematic reviews, we searched four databases. Furthermore, in May 2019, we performed a citation search of five other websites. We included randomized trials, non-randomized trials, controlled before-after studies, and interrupted time series studies of adolescents (girls or boys aged 10 to 19 years) eligible for World Health Organization-recommended vaccines and their parents or healthcare providers.

Data collection and analysis: Two review authors independently screened records, reviewed full-text articles to identify potentially eligible studies, extracted data, and assessed risk of bias, resolving discrepancies by consensus. For each included study, we calculated risk ratios (RR) or mean differences (MD) with 95% confidence intervals (CI) where appropriate. We pooled study results using random-effects meta-analyses and assessed the certainty of the evidence using GRADE.

Results: We included 16 studies (eight individually randomized trials, four cluster randomized trials, three non-randomized trials, and one controlled before-after study). Twelve studies were conducted in the USA, while there was one study each from: Australia, Sweden, Tanzania, and the UK. We categorized interventions as recipient-oriented, provider-oriented, or health systems-oriented. The interventions targeted adolescent boys or girls or both (seven studies), parents (four studies), and providers (two studies). Five studies had mixed participants that included adolescents and parents, adolescents and healthcare providers, and parents and healthcare providers. The outcomes included uptake of human papillomavirus (HPV) (11 studies); hepatitis B (three studies); and tetanus–diphtheria–acellular–pertussis (Tdap), meningococcal, HPV, and influenza (three studies) vaccines among adolescents.

Conclusions: Various strategies have been evaluated to improve adolescent vaccination: health education, financial incentives, mandatory vaccination, and class-based school vaccine delivery. However, most of the evidence is of low to moderate certainty. This implies that while this research provides some indication of the likely effect of these interventions, the likelihood that the effects will be substantially different is high. Therefore, additional research is needed to further enhance adolescent immunization strategies, especially in low- and middle-income countries where there are limited programmes

Patient or healthcare consumer involvement: Adolescents and their parents and community.

Innovative patient and carer partnership in creating trustworthy guidelines, from protocol to publication: a case study of BMJ Rapid Recommendations

Lytvyn L[1], Siemieniuk RA[1], Macdonald H[2], Price A[2], Lyddiatt A[3], Brandt L[4], Heen AF[4], Guyatt G[1], Agoritsas T[5], Vandvik PO[4]

[1] McMaster University, Canada; [2] The BMJ, UK; [3] Cochrane Consumers, Canada; [4] University of Oslo, Norway; [5] University of Geneva, Switzerland

Background: The BMJ Rapid Recommendations (RapidRecs) are guidelines developed in response to potentially practice-changing evidence, following the GRADE approach and the Institute of Medicine trustworthy guidelines criteria. They are produced by the MAGIC Evidence Ecosystem Foundation in partnership with The British Medical Journal (BMJ). RapidRecs are published along with infographics, with linked systematic reviews, and decisions aids, in The BMJ and MAGICapp.org. RapidRecs are developed by international panels of unconflicted clinical experts, methodologists, and patients and carers. All communication is via teleconference, phone, and email.

Objectives: We sought to determine the feasibility and effect of patient/carer partnership at each step of rapid guideline development.

Methods: For each RapidRec, we recruit patient/carer partners from general and health-specific organizations, panel member referrals (provided that patients/carers are not in their circle of care), Cochrane TaskExchange, and Twitter. Partners receive an invitation, conflict of interest form, and teleconference or phone call describing the RapidRecs project, expected commitment, and timelines. Those who participate: 1) identify and prioritise patient-important outcomes for the supporting systematic reviews; 2) identify practical issues for shared decision making; 3) engage in an education session before panel deliberations; 4) participate in deliberation teleconferences; and, 5) edit draft recommendations and manuscript as co-authors. We will conduct interviews with patient/carer partners to describe their experience, perspective on contributions, and identify strengths and weaknesses of our approach. We are also developing a systematic approach to evaluate the contributions and their impact from all guideline panel members.

Results: To date, we published 16 guidelines, including 43 partners from 14 countries. Partners were recruited from referrals (N = 12), general organizations (N = 11), health condition-specific organizations (N = 11), TaskExchange (N = 7), and Twitter (N = 2). Preliminary feedback from patients/carers and other panel members has been positive regarding the partnership process and contributions from patients/carers. Areas of improvement that we are pursuing are establishing an honorarium for patient/carer panel members, collaboration and knowledge exchange with others in the field of patient-oriented research, establishing relationships with patient/carer organizations, maximizing patient/carer involvement while managing burden, as well as documenting challenges and resources required (e.g. recruitment, education).

Conclusions: We provide an example of a patient/carer partnership initiative for guideline development that is feasible, and may produce more trustworthy, relevant, and patient-centred guidelines for shared decision-making.

Patient or healthcare consumer involvement: Patients/caregivers are involved in every RapidRecs guideline. Two co-authors (AP, AL) of this abstract are consumers in addition to their research role.

Integrating knowledge user perspectives into evidence-based clinical practice guideline development and implementation

Burnett L[1], Fahim C[1], Thombs B[2], Moore A[3], Rodin R[4], Straus SE[5]

[1] Li Ka Shing Knowledge Institute, St. Michael's Hospital, Canada; [2] Canadian Task Force on Preventive Health Care; Faculty of Medicine, McGill University, Canada; [3] Canadian Task Force on Preventive Health Care; Department of Family Medicine, McMaster University, Canada; [4] Global Health and Guidelines Division, Public Health Agency of Canada, Canada; [5] Li Ka Shing Knowledge Institute, St. Michael's Hospital; Department of Geriatric Medicine, University of Toronto, Canada

Background: The Canadian Task Force on Preventive Health Care (Task Force) develops and disseminates evidence-based clinical practice guidelines (CPGs) for use in primary care across Canada. Using an integrated knowledge translation (iKT) approach and the Knowledge to Action model, knowledge users (KUs) (e.g. clinicians and patients) are engaged to inform the development of CPGs and associated KT tools (e.g. infographics and algorithms), thereby yielding more relevant and acceptable guidelines and tools. Integrating KU perspectives to inform CPG and KT tool development may improve uptake of evidence-based guidelines, however there is a paucity of literature on how to optimally engage patients in these processes.

Objectives: (1) To identify key KU preferences across a variety of Task Force CPGs; (2) To discuss lessons learned from the Task Force's KU engagement process.

Methods: The Task Force engaged KUs at three stages in the CPG development process: (1) selecting outcomes to include in the systematic review protocol that will inform the CPGs; (2) developing CPG recommendations; and (3) developing KT tools.

In Stages 1 and 2, the Task Force used focus groups and surveys to identify patient-important outcomes related to the guideline topics. In Stage 3, the Task Force conducted usability testing interviews and focus groups to elicit feedback on the content, format, and usefulness of the KT tools. The Task Force used the Public and Patient Engagement Evaluation tool to assess KU experience with the engagement process.

Results: Since 2010, the Task Force developed 19 CPGs on preventive healthcare topics using this iKT approach. The Task Force completed Stage 1 with 242 patients across 14 guidelines and Stage 2 with 136 patients across eight guidelines. Patients were 70% female, with mean age 46 years (range: 13 to 78 years). Participants typically rated outcomes focused on intervention benefits as 'critical' for decision-making and rated harms slightly lower ('important'). Stage 3 was completed with 139 clinicians (67% female, between the ages of 20 and 79), and 74 patients (65% female, between the ages of 20 and 79) across 15 guidelines. Most found draft KT tools clear and straightforward; feedback generally focused on layout modifications and requests for additional topic-related information. Participants generally had positive experiences with the engagement process and appreciated the opportunity to contribute to Canadian healthcare.

Conclusions: The Task Force provides a model for engaging KUs using an iKT approach to elicit their values and preferences at three stages in the CPG development process. The approach helps address literature gaps around effective methods for engaging patients in CPG and KT tool development.

Patient or healthcare consumer involvement: Patients are involved with the Task Force activities and the outlined approach may help CPG development and implementation groups enhance the uptake and applicability of their CPGs.

Involvement of patients in the development of Cochrane protocols and reviews, a COPD survey, a patient blog, and a patient information leaflet

Janjua S[1], Threapleton C[1], Baker E[1], Cates C[1], Fortescue R[1], Dennett E[1]

[1] Cochrane Airways, St George's, University of London, UK

Background: We received a NIHR programme grant to produce reviews for patients with chronic obstructive pulmonary disease (COPD). We recognise that involving people living with COPD in the process of research prioritization ensures that we produce the most important reviews for this patient group.

Objectives: To develop Cochrane protocols and reviews, a COPD survey, to recruit new consumer peer referees and authors, and to involve consumers in knowledge translation.

Methods: Nine COPD patients have volunteered from a larger patient advisory group at St. George's Hospital, London. Patients are involved in five areas:

1) regular meetings every six months for 90 minutes in which 2 to 3 Airways members (and a COPD clinical specialist) guide discussion of review topics;
2) development of a one-off patient survey to capture broader opinions;
3) writing a blog;
4) providing consumer refereeing and authoring;
5) knowledge translation (patient information leaflet).

Results: During 24 months of the programme grant, we have had four patient meetings. Patients have suggested new topics we had not previously considered, including a review about air pollution. They have helped us to refine review topics and suggested possible additional approaches for reviews, including new outcomes. They have given guidance about the relative priority of possible patient-important topics (maintenance rehabilitation), and approved our decisions to progress with reviews. We obtained 200 responses from the survey. Eight themes identified have helped to develop new topics for the programme grant and the patient blog increased survey uptake. Patients have helped us to meet the requirements of the new peer review policy (every review should have a consumer referee). One patient, involved as an author, will guide us on patient perspective in interpretation of findings in that review. Patients are helping us to develop a leaflet that will consist of information about evidence from important reviews, and also about Cochrane Airways.

Conclusions: Patient involvement is an important and rewarding part of our programme grant, enabling development of review questions, protocols and complete reviews that reflect the patient perspective. Patients have also made a valuable contribution to survey development and the patient information leaflet. Patient involvement requires appropriate resources, which should be factored in during the grant application stage. We plan to involve patients in knowledge translation efforts, for example, helping to present findings in a suitable format for patients.

More than a letter of invitation: drawing on digital marketing strategy to engage stakeholders in evidence production

Steele E[1], Noel-Storr A[2], Turner T[3], Champion C[2]

[1] Cochrane, Australia; [2] Cochrane, UK; [3] Monash University, Australia

Background: Four years ago Cochrane launched two innovative solutions to pressing challenges in evidence production: Cochrane Crowd and Cochrane TaskExchange. In different ways these online platforms aim to increase efficiency in the production of systematic reviews and other health evidence projects. Their success depends in part on attracting evidence production stakeholders such as consumers, students and evidence production professionals to sign up and stay engaged over a period of time.

Objectives:

- To build awareness of platforms within target user groups (awareness)
- To build global communities of users on both platforms (acquisition)
- To facilitate the journey from sign up to engagement in activities (conversion)
- To engage users in evidence activities over time (retention)
- To encourage users to refer colleagues and friends to the platforms (referral)

Methods: Drawing predominantly on knowledge from the new discipline of Digital Marketing, we have designed strategies over a four-year period to meet objectives across the marketing funnel; that is, from awareness through to referral. Strategies have included bringing a user journey lens to all activities, upfront analysis of target groups, clever onboarding strategies, use of appropriate social media channels, running bespoke campaigns, delivering content such as webinars, blogs and articles, offering rewards to users, and establishing an incentivized referral system. This presentation will showcase examples of specific strategies used, and outcomes of each strategy.

Results: Cochrane Crowd now has 16,000 users and the more specialized platform Cochrane TaskExchange has 5000 users.

Conclusions: The emerging field of digital marketing has much to offer in terms of attracting and engaging a variety of stakeholders in the work of evidence production.

Patient partners' perspectives of meaningful engagement in synthesis reviews: a patient-oriented rapid review

Boden C[1], Edmonds A[2], Porter T[2], Bath B[3], Dunn K[4], Gerrard A[5], Goodridge D[6], Andrews Stobart C[4]

[1] Leslie and Irene Dube Health Sciences Library, University of Saskatchewan, Canada; [2] Patient Partner, Saskatchewan Centre for Patient-Oriented Research, Canada; [3] School of Rehabilitation Science, University of Saskatchewan, Canada; [4] Saskatchewan Centre for Patient-Oriented Research, Canada; [5] Student Learning Services, University Library, University of Saskatchewan, Canada; [6] College of Medicine, University of Saskatchewan, Canada

Background: Patient-oriented research is research conducted in partnership with patients. Patient partners contribute expertise from lived experience of interacting with the healthcare system, healthcare providers and day-to-day challenges associated with their healthcare concerns to these research teams. Engaging all relevant stakeholders in patient-oriented research helps focus research studies on patient-identified priorities, with the intention of generating knowledge that ensures better and more meaningful patient outcomes. A growing literature describes best practices for patient-oriented research generally, and patient involvement in systematic reviews specifically. Existing reviews of this literature have largely employed the researchers' lens. In this review, we seek to understand meaningful engagement in synthesis reviews from the patient partner perspective.

Objectives: This patient-oriented rapid review investigated patient partners' perspectives of meaningful engagement in synthesis review to 1) understand how research teams can conduct their reviews to ensure planned, supported and valued involvement of patient partners, and 2) describe the characteristics of a review team which enable meaningful engagement by patient partners. Our definition of meaningful engagement is based on Hamilton et al 2018 (1).

Methods: The review team comprised patient partners, research librarians with synthesis review experience, Saskatchewan Centre for Patient-oriented Research staff with patient-oriented research experience, and academic faculty. We searched health-related databases (OVID MEDLINE, OVID Embase, and ProQuest Nursing and Allied Health) and undertook a focused search of core patient-oriented research websites. We included documents describing patient partners' reflections on their involvement in synthesis reviews (e.g. synthesis reviews, realist reviews). Two review authors independently conducted screening and data extraction. Disagreements were resolved through consensus or adjudication. Descriptive statistics and thematic analysis will be employed to synthesize the quantitative and qualitative data respectively.

Results: The literature search found 1090 citations, of which we included 15 documents from 14 studies. Analysis of extracted data is ongoing and will be completed by May 2020.

Conclusions: Based on patient partner perceptions, findings from this review contribute to a greater understanding of best practices in engaging patient partners meaningfully in a synthesis review.

Patient or healthcare consumer involvement: The research team included two patient partners (AE, TP) who collaborated on all aspects of the review.

Reference:

1. Hamilton CB, Hoens AM, Backman CL, McKinnon AM, McQuitty S, English K, et al. An empirically based conceptual framework for fostering meaningful patient engagement in research. *Health Expectations* 2018;21(1):396-406.

Producing guidance on the ethical, governance and management of involving patients and community groups as co-applicants, team members and co-researchers in health research

Minogue V[1], Donskoy AL[2]

[1] Patient advocate; Executive member NHS RD Forum and co-facilitator service user and carer working group; independent consultant, Ireland; [2] Member service user and carer working group NHS RD Forum; PhD student Bristol, UK

Background: The evidence base for public and patient involvement (PPI) and community based participatory research (CBPR) in health service research has grown in the last 30 years. The roles and responsibilities of patients and the community have also evolved to include being co-applicants in research funding applications, co-researchers, and project team members. The barriers and enablers to these roles are becoming more clearly understood as is the role of PPI in reducing waste (Minogue and Wells, BMJ 2019). Whilst professionals understand their roles and responsibilities in a research study, PPI and community representatives may not be aware of the full implications and responsibilities of taking on these roles. Many PPI representatives are volunteers and unable to access the supportive mechanisms of employers' human resource processes. Evidence collected by a PPI group actively involved in all aspects of research activity suggested there was a lack of understanding on the part of PPI representatives, community groups, researchers, and research sponsors, of the barriers resulting from ethical, legal and governance factors relating to involvement as co-applicants, co-researchers, and team members.

Objectives: Ensure PPI representatives, community groups, research funders, sponsors, investigators and research managers are aware of the responsibilities, ethical and legal implications of PPI and CBPR and how to address the barriers to involvement. Understand the contractual and legal issues and responsibilities that are particular to PPI and CBPR co-applicancy, project team membership and co-research, from an organizational and individual perspective.

Methods: A literature review, review of current practice, data collection of narrative experiences, two workshops, consultation exercise, production of guidance, pilot phase, development of case studies relating to contractual and legal issues.

Results: The review and consultation led to collaborative working with research managers across the NHS and with the National Institute for Health Research and the Health Research Authority. Guidance for research managers, researchers and PPI was published in January 2019 with further guidance on CBPR produced for consultation in January 2020.

Conclusions: The contractual, legal and governance issues and responsibilities that are specific to PPI and CBPR co-applicants, co-researchers and team members are not fully understood by PPI and CBPR representatives, sponsors, researchers or research managers. The new guidelines have contributed to raising awareness and increasing patient empowerment in research and reducing negative experiences for PPI representatives.

Patient or healthcare consumer involvement: The project was conceived, initiated and undertaken by the service user and carer working group of the NHS RD Forum. In undertaking the project, the group liaised with NIHR INVOLVE which resulted in a joint piece of work to produce a further set of guidelines aimed at PPI representatives and researchers (published January 2019). This abstract is an updated and revised version of a Cochrane Colloquium 2019 submission.

Public-private partnerships in primary health care

Doshmangir L[1], Judian N[1], Mahdavi M[2], Gordeev VS[3]

[1] Health Services Management Research Center, School of Management and Medical Informatics, Tabriz University of Medical Sciences, Tabriz, Iran; [2] National Institute of Health Research, Tehran Universality of Medical Sciences, Tehran, Iran; [3] Pragmatic Clinical Trials Unit, Centre for Primary Care and Public Health, Queen Mary University of London, London, UK

Background: The Astana Declaration on primary health care reiterated that primary health care (PHC) is a cornerstone of a sustainable health system for universal health coverage (UHC) and health-related sustainable development goals. It called for governments to give high priority to PHC in partnership with their public and private sector organizations and other stakeholders. Each country has a unique path towards UHC and different models for public–private partnerships (PPPs) are possible.

Objectives: To examine evidence on the use of PPPs in the provision of PHC services, as well as reported challenges and recommendations.

Methods: We systematically reviewed peer-reviewed studies in six databases (Science Direct, PubMed, Web of Science, Embase, Ovid, and Scopus) and supplemented it by the grey literature search. We followed the PRISMA reporting guidelines.

Results: We included the results of 61 studies in the review. Most PPPs projects were implemented to facilitate a provision of and increase access to prevention and treatment services (i.e. tuberculosis, education and health promotions, malaria and HIV/AIDS services) for specific target groups. PPPs projects faced challenges during the starting and implementation phases. Challenges and recommendations related to PPPs in PHC were reported for education, management, human resources, financial resources and information and technology systems.

Conclusions: Despite various challenges, PPPs in PHC can facilitate access to healthcare services, especially in remote areas. Governments should consider long-term plans and sustainable policies to start PPPs in PHC and should not ignore local needs and context.

Patient or healthcare consumer involvement: Not applicable.

Stakeholder mapping to identify stakeholders for the Cochrane Cancer Network

Tomlinson E[1], Opiyo N[2], Skoetz N[3]

[1] Cochrane Cancer Network, UK; [2] Cochrane Cancer Network, Norway; [3] Cochrane Cancer Network, Germany

Background: Involving stakeholders, such as patients, health professionals and policy makers, in systematic review processes helps to ensure that the most important research questions are addressed and outcomes of interest are identified. Consulting stakeholders at an early stage in the review process can also increase acceptance of the research, thereby increasing the potential impact on decision making. The Cochrane Cancer Network is made up of six Cochrane Review Groups (CRGs) who provide reliable evidence required to make important decisions concerning cancer. CRG members indicated wanting to improve stakeholder engagement in their work and to facilitate this, we conducted a stakeholder mapping project to identify relevant organizations and individuals to engage with.

Objectives: To identify stakeholders for the six cancer CRGs and the Network overall, creating a resource that will facilitate stakeholder engagement in review production and knowledge translation activities.

Methods: We gathered stakeholder information from CRGs, guidelines and social media. First, we surveyed CRGs for information about stakeholders they have worked with and would like to work with. Then, using information from Cochrane UK, we searched guidelines in which Cochrane Reviews had been cited from 1 January 2017 to 31 September 2019 for lists of contributing stakeholders. We compiled stakeholders into a spreadsheet with a page for each CRG and searched all mapped stakeholders on Twitter to gather information about the type of organization, audience, online presence, location and to find further suggestions. Stakeholders common to all CRGs were added to a Network map. We sent each CRG their stakeholder map and edits were made following discussions about stakeholder relevance, alignment of priorities and feasibility of connection.

Results: We identified 180 stakeholders located in over 15 countries with a collective total of 7.8m Twitter followers. Of these, 63 organizations were charities. An overview is shown in Figure 1. So far the stakeholder map has been used by two CRGs to inform the planning of a priority setting exercise and a dissemination strategy for a suite of reviews.

Conclusions: There are numerous organizations and individuals working towards common goals in cancer research across the world. Identification of these through this stakeholder mapping exercise will allow us to work towards increasing stakeholder partnerships in the Cancer Network.

Patient or healthcare consumer involvement: Identification of relevant stakeholders is the first step to improving stakeholder engagement in review processes in the Cancer Network. This will lead to more relevant review questions, more thought about what matters to decision-makers and increased acceptance of Cochrane systematic reviews in evidence-based decision making. Ultimately this will increase the usage and relevance of Cochrane evidence to stakeholders, including patients and consumers.

Additional file: [Stakeholder mapping project visual](#)

Supporting rapid learning and improvement in a health-system transformation: a qualitative description of implementation supports

Waddell K[1], Lavis J[2], Bulloch H[1], Grimshaw J[3]

[1] Rapid Improvement Support and Exchange, Canada; [2] McMaster University, Canada; [3] Ottawa Hospital Research Institute, Canada

Background: Rapid-Improvement Support and Exchange (RISE) was created to support consumer- and evidence-driven rapid learning and improvement cycles among newly created Ontario Health Teams. As Ontario Health Teams mature, they will become clinically and fiscally accountable for delivering a full and co-ordinated continuum of care.

Objectives: To describe how rapid learning and improvement is being used to support health-system transformation, namely the creation of Ontario Health Teams.

Methods: We use qualitative description to describe the use of transformational design to develop supports for Ontario Health Teams. Transformational design emphasizes the use of distributed resources (knowledge, tools, and expertise) and participation of users, beyond the public including management and front-line staff, to develop solutions to system-level challenges. We use five data sources including literature reviews, documentary analysis, one-on-one key informant interviews, focus groups, evaluations of a two-day event, and the experiences of the authors participating in the work of RISE to inform this description.

Results: RISE is being iteratively designed in a manner that matches the principles of transformative design by employing participatory methods to engage stakeholders in the process of designing and providing supports. This includes the development of two communities of practice, with 269 members, the mobilization of knowledge in 13 RISE briefs and 3 rapid syntheses to date, and through the insights of two focus groups to debrief an event held for over 200 stakeholders. We will present a description of designing RISE and the supports that it provides according to ten early lessons learned about implementation supports, which among others, include: 'roll with the punches;' emphasize learning from each other rather than from experts; and have coaches and other 'on-the-ground' supports play to their comparative advantages and commit to 'warm handoffs.'

Conclusions: To date RISE has aimed to leverage expertise within the health system and fill gaps by synthesizing local and global evidence to be used in the implementation of Ontario Health Teams. However, just as in the rapid learning and improvement framework, work to collaboratively design and develop supports is constantly in need of adjustment to ensure they are tailored to support teams. This description documents the iterative design and development of RISE and provides one example of how to operationalize rapid learning and improvement to support health-system reforms.

Healthcare consumer involvement: Citizens, patients, families and caregivers are engaged as key partners in our network of experts as well as on the RISE advisory council. Further, patient and caregiver partners recently participated in a citizen panel to codesign guidance for Ontario Health Teams on how to engage citizens, patients, families and caregivers in the process of reform.

The value of stakeholder, patient and public contributions and knowledge of lived experience to the conduct and interpretation of empirical findings from synthesized evidence

Babatunde O[1], Bishop A[1], Rhodes C[1], Cottrell E[1], van der Windt D[1]

[1] School of Primary, Community and Social Care; Keele University, UK

Background: Robust evidence synthesis, including Network Meta-Analysis (NMA), is a fundamental tool for clinical decision and health policy making. However, integration of patient values with clinical experience and best available research information is not always apparent at the core of current evidence-based practice.

Objectives: This project explores novel approaches to patient and public involvement and engagement (PPIE); contribution of lived experiences to the conduct and interpretation of evidence from an NMA and a large systematic review.

Methods: The two studies used to explore the value of PPIE contributions to the conduct of evidence synthesis followed established guidance and systematic review methodology. First, an NMA (80 trials, 4533 patients, 19 treatments) was conducted to determine the comparative effectiveness of treatments for improving pain and function of patients with subacromial shoulder conditions (SSC). PPIE representatives with SSC (n = 5) participated in a discussion regarding management of SSC. This included the process of shared decision making (if any), treatment outcomes and factors which may have influenced outcomes of treatment. The group were then invited to rate the effectiveness of treatments and also rank the likelihood of benefit from each treatment (on a scale of 1 to 5) based on their own personal lived experiences. One review author performed random-effects NMA and cumulative ranking of interventions without knowledge of the details of PPIE contributions. Second, we conducted a systematic review (26 studies, 120,276 patients) and narrative synthesis of the clinical and socio-economic outcomes of musculoskeletal (MSK) triage/direct access services. We supported PPIE representatives with MSK pain (n = 8) to read and assimilate a sample paper. We encouraged them to highlight items they felt were relevant for data extraction and discussed important issues related to access to care for MSK pain based on their own lived experience.

Results: PPIE conventions highly correlated with/predicted empirical findings from the NMA which found small to moderate effects for most treatment options for SSCs, with exercise ranked highly for long-term benefit. PPIE involvement in the MSK triage/direct access review resulted in a more comprehensive data-extraction instrument. It provided a basis for within-data exploration of potential factors (e.g. socio-demographics, and health literacy) that may be associated with long-term health and socio-economic gains due to MSK triage/direct access.

Conclusions: Involvement of PPIE improved the robustness of the review process (specifically data-extraction and interpretation of results) and also validated empirical findings from the review.

Patient or healthcare consumer involvement: The project draws on PPIE involvement in bridging the research to practice gap, and shared decision-making as a core principle of evidence-based research and practice. As part of study advisory group, clinicians and relevant stakeholders also contributed to the interpretation of evidence from both studies.

ID: CRD42018085978/42014009788

What difference does patient and public involvement make to evidence synthesis?

Mann M[1], Hannigan B[1], Edwards D[1], Anstey S[1], Coffey M[2], Gill P[1], Meudell A[3]

[1] Cardiff University, UK; [2] Swansea University, UK; [3] Independent Service User Researcher, UK

Background: Research funders increasingly recommend and require patient and public involvement (PPI) in all aspects of the research cycle from identifying and prioritizing research topics to conducting and dissemination. Our project was funded by the National Institute for Health Research (NIHR) to identify and synthesise research and other evidence in the area of end of life care for people with severe mental illness. Our funded project team included PPI members, and our Stakeholder Advisory Group (SAG) consisted of representatives drawn from the mental health and end of life care (EoLC) fields. Members of the SAG included professionals from a range of practitioner backgrounds based in the National Health Service and charitable organizations, policy advisers and people with personal experience of mental health difficulties and EoLC.

Objectives: To discuss how PPI contributed to an evidence synthesis and to highlight the benefits of involvement throughout the research process.

Methods: Training was provided relating to aspects of the evidence synthesis process to members of the project team, including PPI co-investigators. Members of the SAG attended the scheduled meetings throughout the project. We held our first meeting at the commencement of the project to refine search terms and strategies for the evidence review. The second meeting was held at the completion of the literature search and screening stage. The final meeting was held to introduce the synthesis and report writing stage, where we discussed progress and plans for dissemination and maximizing impact.

Results: Members of the public were involved in identifying and prioritizing the research questions for this project, and in identifying search terms and grey literature sources. In addition, they were involved in the study selection stage, commenting on the final report and disseminating findings.

Conclusions: Engaging with patients and members of the public provided useful insights and helped contextualise findings and their relevance. PPI in evidence synthesis adds value in all steps of the review.

Patient or healthcare consumer involvement: This project has had full PPI from the beginning and this study demonstrates the benefits of this approach of meaningful involvement. This project was funded by the National Institute for Health Research (NIHR) Health Services and Delivery Research (HS&DR) programme (project number 17/100/15). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

KNOWLEDGE TRANSLATION

"It's not smooth sailing": bridging the gap between methods and content expertise in public health guideline development

Chartres N[1], Grundy Q[2], Parker L[1], Bero L[1]

[1] Charles Perkins Centre, Faculty of Medicine and Health, School of Pharmacy, The University of Sydney, Sydney, Australia; [2] Faculty of Nursing, University of Toronto, Toronto, Canada

Background: The development of reliable, high-quality health-related guidelines depends on explicit and transparent processes, methods aimed at minimizing risks of bias and the inclusion of all relevant expertise and perspectives. While the methodological aspects of guidelines have been a focus to improve their quality, less is known about the social processes involved, for example, how guideline group members interact and communicate with one another, and how the evidence is considered in informing recommendations. With this in in mind, we aimed to empirically examine the perspectives and experiences of the key participants involved in developing public health guidelines for the Australian National Health and Medical Research Council (NHMRC).

Design: This study was conducted using constructivist grounded theory as described by Charmaz, which informed our sampling, data collection, coding and analysis of interviews with key participants involved in developing public health guidelines.

Setting: Australian public health guidelines commissioned by the NHMRC.

Participants: Twenty experts that were involved in Australian NHMRC public health guideline development, including working committee members with content topic expertise (n = 16) and members of evidence review groups responsible for evaluating the evidence (n = 4).

Results: Public health guideline development in Australia is a divided process. The division is driven by three related factors:

1) the divergent disciplinary background and expertise that each group brings to the process;
2) the methodological limitations of the framework, inherited from clinical medicine, that is used to assess the evidence; and
3) barriers to communication between content experts and evidence reviewers around respective roles and methodological limitations.

Conclusion: Our findings suggest several improvements for a more functional and unified guideline development process: greater education of the working committee on the methodological process employed to evaluate evidence, improved communication on the role of the evidence review groups, and better facilitation of the process so that the evidence review groups feel their contribution is valued.

A health service approach to dissemination, knowledge translation and research impact

Minogue V[1], Morrissey M[1], Terrés A[1]

[1] Research and Evidence, Health Service Executive (HSE), Ireland

Background: The literature about research waste tells us that two of the factors in research having little or no impact are it not being disseminated and implemented. Furthermore, there are significant barriers in getting research into practice, policy, and guidance. Embedding research, engaging clinical services, and realizing impact and adding value through research, are the areas that research managers find most challenging. The Health Service Executive (HSE) in Ireland, in setting up new processes for research management and governance, also wanted to enhance a culture where research evidence is valued and has impact.

Objectives: The objectives of the project were to produce a framework for dissemination, knowledge translation, and impact, to identify a suitable Knowledge Translation (KT) framework and impact tool(s) for use in research studies and within organizations.

Methods: The presentation will describe a project designed to address the issues of embedding research evidence and realising benefit from research across services and organizations. The project was developed and delivered by two project managers and a working group. It involved scoping and reviewing two systematic reviews and circa 250 models of KT, a review of impact tools, a pilot implementation phase.

Results: Deliverables included a dissemination, a KT and impact approach with guiding principles, a KT model, and impact tools.

Conclusions: Creating evidence-based culture and embedding research has to be framed within a clear approach to dissemination and impact. Key learning points include:

- the importance of dissemination and impact planning throughout the research study;
- the KT framework must be usable and applicable in health service settings and have clear constructs;
- the need to engage knowledge users in developing a framework; and
- while lots of groups are energized to get research into practice and use evidence, this project streamlines resources to support researchers to effectively enhance the use of their research findings.

Ownership by key knowledge users and senior managers is important to successful implementation. A project to develop an approach to dissemination, knowledge translation (KT) and impact resulted in framework guidance for research managers, researchers and organizations.

Patient or healthcare consumer involvement: Access to a HSE patient advisory group is in development.

A pilot to translate Cochrane Incontinence blogshots: does it sustainably help reach a more global audience?

Johnson E[1], Ried J[2]

[1] Population Health Sciences Institute, Newcastle University, Newcastle upon Tyne, UK; [2] Cochrane Central Executive, Germany

Background: In 2017, Cochrane Incontinence adopted a new dissemination policy that included the production and circulation of visual products to disseminate the key evidence of its reviews, such as blogshots (a short summary of key points from a review with a picture that can be shared on social media). However, many of the Group's dissemination products are in English, while only 6% of the world are native English speakers and 75% do not speak any English. Producing and sharing visual dissemination products in languages other than English could improve engagement with reviews internationally.

Objectives: To assess the feasibility and impact of translating visual dissemination products for Cochrane Incontinence reviews to increase their international reach.

Methods: We developed a pilot project to produce and evaluate four translated blogshots presenting the evidence from four recent Cochrane Incontinence reviews. We sought French, Spanish and German translators from Cochrane Geographic Groups and TaskExchange. Two people worked on each translation in consecutive steps: an initial translation and editing. We sent all translators an evaluation form to collect information on their experience, while we disseminated translated blogshots via Cochrane Incontinence's Twitter account according to a defined schedule. We then collated and tabulated quantitative data on impressions and engagements from Twitter Analytics between 20 January and 28 February 2020. In addition, we gathered data regarding the countries from which the reviews were being accessed via Altmetrics before the pilot and repeated this after for comparison.

Results: We will present the results of the quantitative evaluation of the pilot scheme along with the results from translator surveys.

Conclusions: We will present conclusions based on both the quantitative data and surveys, considering these in the context of how translated blogshots may be able to contribute to the rapid dissemination of key results internationally. We will also discuss the key limitations and challenges of the pilot scheme and how these may affect the ways in which Cochrane Review Groups and Geographic Groups collaborate regarding visual dissemination materials.

Patient or healthcare consumer involvement: There was no direct consumer involvement in the pilot, but we hope that the results will contribute to making Cochrane evidence more accessible to patients and the public internationally.

A readability assessment of Plain Language Summaries and abstracts of Cochrane Reviews published during 2019

Meza N[1], Perez-Bracchiglione J[1], Franco J[2], Garnham R[3], Tapia-Benavente L[3], Olguín-Sepúlveda P[3], Urrútia G[4], Madrid E[1]

[1] Interdisciplinary Centre for Health Studies (CIESAL), Cochrane Associate Centre, School of Medicine, Universidad de Valparaíso, Chile; [2] Research Department, Cochrane Associate Centre, Instituto Universitario Hospital Italiano, Argentina; [3] Cochrane Associate Centre, School of Medicine, Universidad de Valparaíso, Chile; [4] Iberoamerican Cochrane Centre, Biomedical Research Institute Sant Pau, CIBER Epidemiología y Salud Pública, Barcelona, Spain

Background: Plain Language Summaries (PLS) of Cochrane Reviews must be written in clear and simple language to serve a relevant aim of knowledge translation, and must be accurately summarized in a succinct and readable style, as a way to deliver the results to a broad audience. Cochrane has approved standards for the elaboration of the message of PLS, and the review groups have an important role as well, since they are expected to ensure these aspects. Some groups have reported that readability of PLS might be difficult even for medical students, and those who have assessed readability have shown controversial results. Moreover, the readability of PLS over abstracts of Cochrane Reviews has not been fully elucidated.

Objectives: To analyse the readability of the PLS and abstracts of Cochrane Reviews of interventions published during 2019.

Methods: We undertook a cross-sectional study, aiming to assess the readability of all PLS and abstracts of Cochrane Reviews of interventions published during 2019. We excluded protocols, withdrawn reviews, and non-intervention reviews (prognostic, diagnostic, etc.). We retrieved all reviews from the Cochrane Database of Systematic Reviews published from 1 January to 31 December 31 2019. We extracted the following data: title, authors, country of affiliation of the corresponding author, Cochrane Review Group, and text of the PLS and abstract (including titles). We assessed the readability of the abstracts and PLS using 'Readability Formula', as suggested by the MECIR Manual, in order to obtain the Simple Measure Of Gobbledygook Index (SMOG), and other indices, which gives an estimate of the years of education needed to understand a text.

Results: We analysed 546 Cochrane Reviews. The resulting SMOG scores for abstracts were 11.32 (±1.55), while PLS scored 11.99 (±1.73), which was significantly different ($P < 0.001$, meaning that PLS were less readable). A total of 326 (59.7%) of these reviews had corresponding authors whose affiliations were from majority native English speaking countries (MNESC). There were no differences between the readability of PLS ($P = 0.05$), according to whether the corresponding authors' affiliations were based in MNESC or not; and that was the same for abstracts ($P = 0.53$). The differences in PLS and abstracts were also significant in subgroups of MNESC and non-MNESC main author (see Table 1).

Conclusions: The abstracts and PLS of Cochrane Reviews published during 2019 have a readability for lay people equivalent to 11th grade (i.e.16 to 17 years old). PLS are significantly less readable than abstracts in all the Cochrane Reviews assessed, and we found no differences when the corresponding authors' institutions were from a MNESC or not. We found no differences among the readability of PLS or abstracts, when we compared MNESC with non-MNESC groups.

Patient or healthcare consumer involvement: Patients and consumers will be able to take an active role in their health care as long as they have an adequate health literacy to understand the health information regarding their conditions.

Table 1. Readability SMOG index of abstracts and PLS in relation to the country of origin of each corresponding author's affiliation, classified according to whether these were from majority native English speaking countries or not.

	Readability score mean (SD)* of abstracts	Readability score mean (SD)* of PLSs	p-value ¥
All the CSR (n=546)	11.328 (±1.557)	11.995 (±1.739)	<0.001

	Readability score mean (SD)* of abstracts	Readability score mean (SD)* of PLSs	p-value ¥
CSR whose corresponding authors' affiliation were in majority native English speaking country † (n=326)	11.29 (±1.56)	11.87 (±1.77)	<0.001
CSR whose corresponding authors' affiliation were in **NO** majority native English speaking country † (n=220)	11.38 (±1.54)	12.16 (±1.66)	<0.001
p-value ¥	0.52	0.05	

* Statistical Measure Of Gobbledygook (SMOG) index means (SD) of the 546 Cochrane SRs assessed.
† According to The University Of Sheffield's list of majority native English speaking countries (https://www.sheffield.ac.uk/international/english-speaking-countries).
¥ Two-sample mean-comparison by Student's T test calculation.

Assessing the accuracy, reliability and quality of online health information on treatment for patients with COVID-19

Luo X[1], Ren M[2], Liu Y[2], Zhang X[3], Lv M[2], Wang L[2], Chen Y[4], Wang M[5]

[1] School of Public Health of Lanzhou University; Cochrane China Network for GRADE and Guideline Working Group, Lanzhou, China; [2] School of Public Health of Lanzhou University, Lanzhou, China; [3] The First School of Clinical Medicine, Lanzhou University, Lanzhou, China; [4] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; Cochrane China Network for GRADE and Guideline Working Group, Lanzhou, China; [5] Department of Radiology, the First Hospital of Lanzhou University, Lanzhou, China

Background: Since the outbreak of novel coronavirus disease (COVID-19) in Wuhan, Hubei Province, in December 2019, there has been widespread concern and controversy about the treatment options. Information on the internet about the treatment of COVID-19 is endless and contradictory, so that the public cannot easily identify effective information.

Objectives: The aim of this study was to assess the accuracy, reliability and quality of online health information about treatment for patients with COVID-19.

Methods: We searched the Baidu, Google and WeChat platforms on 12 March 2020, and each search selected the top 50 results for selection. In Baidu we used the advanced search function; we searched Google using the Chinese and English search term "treatment of COVID-19"; We used the mobile version of WeChat to search by "treatment of COVID-19" in the "Search" function, the scope of search is unlimited, and the search is sorted by the number of reads. We also conducted a supplemental search of the above platforms on 21 March 21 2020. We used the DISCERN instrument to evaluate the accuracy, reliability and quality of the contents regarding the treatment of COVID-19.

Results: We included 15 search results from our search on 12 March 2020, (10 from Baidu, 4 from Google, and 1 from WeChat), and then 16 search results were added by free search, and a total of 31 search results were included for accuracy, reliability and quality evaluation. Specific evaluation results will be presented at a later academic conference.

Conclusions: By evaluating and analyzing the health information of COVID-19 treatments, this study hopes to improve the recognition of health information in the mass media and to report health information scientifically, so as to provide the public and patients with effective health information and assist them in making informed health decisions.

Patient or healthcare consumer involvement: None.

Better dissemination using Cochrane's Dissemination Checklist: Cochrane UK's experience

Ryan-Vig S[1], Chapman S[1]

[1] Cochrane UK, UK

Background: In 2019, Cochrane introduced a Dissemination Checklist and Guidance for disseminating findings from Cochrane intervention reviews. At Cochrane UK we share Cochrane evidence in a variety of ways, including in blogs and blogshots (a way of sharing the key messages from a Cochrane Review in a single image).

Objectives:

- To raise awareness of the Cochrane Dissemination Checklist and Guidance.
- To share examples of how the Checklist can be applied in practice for better dissemination of Cochrane evidence.
- To raise awareness of Cochrane UK's resources based on the Dissemination Checklist.

Implications and impact on our practice: As a result of the introduction of the Checklist, at Cochrane UK we have made changes to our dissemination products. Whilst we find it useful to use the Dissemination Checklist in its entirety, we have found that there are particular items which have especially influenced the way we share evidence. We have also developed new guidance based on the Checklist, including 1) guidance for bloggers; 2) guidance for people in Cochrane creating blogshots; and 3) guidance for choosing images to share alongside Cochrane evidence. We hope that these resources, which are available for use by others in Cochrane, will improve the global efforts in sharing Cochrane evidence with a variety of stakeholders.

Patient or healthcare consumer involvement: Cochrane UK will involve consumers in the development of the Images Checklist. Prior to beginning work on the Checklist, Cochrane UK along with Cochrane Common Mental Disorders sought feedback from the public and from people with lived experience of mental health problems about appropriate image use when illustrating mental health topics. The Checklist draws on this feedback. Additionally, having new guidance for bloggers has improved our ability to work with, and support, our guest bloggers including healthcare consumers.

Citation of Cochrane systematic reviews in newspapers

Casino G[1]

[1] Iberoamerican Cochrane Centre, Spain

Background: National citation patterns of leading general medical journals (*NEJM*, *The Lancet*, *JAMA* and *BMJ*) in newspapers has been analyzed (1), as well as the contribution of authors from countries all around the world to the research literature published in these medical journals (2). Contribution of countries to the publication of Cochrane systematic reviews has also been mapped (3). But the citation of Cochrane Reviews in newspapers is not known.

Objectives: To analyze the citation of Cochrane systematic reviews in newspapers and to compare the results with those of related previous studies.

Methods: Content analysis of the full text of 22 newspapers from 14 countries since 1993 (when Cochrane was founded) to 2019. Newspapers were selected following related previous studies and categorized into four regions: the USA, the UK, Euro-American countries (European countries other than the UK, and Australia, New Zealand and Canada) and Rest of the World (other countries) (1,3). Searches were performed in Factiva database using two terms (Cochrane and review*) in six languages (English, Spanish, French, Italian, Portuguese and German). Duplicates and non-eligible retrieved news pieces were eliminated.

Results: Of the 22 newspapers analyzed, the four that most cited the Cochrane Reviews during the period 1993-2019 were *The Guardian* (172), *The Times* (131), *The New York Times* (110) and *The Australian* (85) (Figure 1). The number of annual citations was no more than 20 between 1993 and 2004, rising sharply in 2005 (62 citations) and 2006 (65), and then decreasing slightly until 2019 (42), with the lowest number in 2017 (26) (Figure 2). In 2008-15, citations of Cochrane Reviews represent 14.1%, 20.9%, 10.3% and 10.7% of citations of *NEJM* (417/2,948), *JAMA* (417/1,992), *The Lancet* (417/4,064) and *BMJ* (417/3,884), respectively (Figure 3). The number of Cochrane Reviews conducted since 1993 in the 14 countries analyzed correlates positively with the number of citations of Cochrane Reviews in the newspapers of these countries (Figure 4).

Conclusions: Citations of Cochrane Reviews in newspapers decreased slightly since 2006 and represent 10-21% of citations of *NEJM*, *JAMA*, *The Lancet* and *BMJ* in the period 2008-15. The newspapers that cite the most Cochrane Reviews are those from the UK, Australia and the US, which are the countries that produce the most of them.

Patient or healthcare consumer involvement: None

Additional file: [Figures](#)

References:

1. Casino G, Rius R, Cobo E. National citation patterns of NEJM, The Lancet, JAMA and The BMJ in the lay press: a quantitative content analysis. *BMJ Open* 2017;7(11):e018705.
2. Sumathipala A, Siribaddana S, Patel V. Under-representation of developing countries in the research literature: ethical issues arising from a survey of five leading medical journals. *BMC Medical Ethics* 2004;5:5.
3. Groneberg DA, Rolle S, Bendels MHK et al. A world map of evidence-based medicine: Density equalizing mapping of the Cochrane database of systematic reviews. *PLOS One* 2019;14(12):e0226305.

Cochrane Brazil Rio de Janeiro Affiliate's activities and the mission of promoting Cochrane in Brazil

Fontes LE[1], Martimbianco AL[1], Latorraca C[1], Pacheco R[1], Vieira R[1], Mochdece CC[1], Riera R[1]

[1] Cochrane Brazil Rio de Janeiro, Brazil

Background: In 2016, Cochrane launched the Brazilian Cochrane Network, comprising five Affiliate Centers, located in three different Brazilian regions. Since then, the Cochrane Brazil Rio de Janeiro Affiliate (CBRJ), hosted by Faculdade de Medicina de Petrópolis (FMP/FASE), has implemented a set of activities for promoting evidence-based decision making in healthcare in Rio de Janeiro, following the guidance of Cochrane's strategic plan.

Objectives: To describe the activities carried out by Cochrane Brazil Rio de Janeiro Affiliate to date.

Methods: Descriptive case study.

Results: The CBRJ currently counts with nine volunteer members who have implemented the following set of activities:

1) translation of 30 abstracts and Plain Language Summaries of Cochrane Reviews into Portuguese, through a network of 20 volunteers;
2) translation of 36 posts from the Students 4 Best Evidence blog;
3) six training workshops on Cochrane systematic reviews, that reached 120 potential review authors;
4) publication of 12 Cochrane Reviews;
5) publication of 10 Cochrane protocols;
6) 2 title registrations for Cochrane Reviews;
7) supportive role for five newcomers as authors of Cochrane Reviews;
8) consolidation of partnership with the program of scientific initiation of FMP/FASE, supporting and giving advice to eight undergraduate medical students.

Recently, the CBRJ team was in charge of the translation of abstracts and Plain Language Summaries related to the management of COVID-19. The next strategic activity is the translation of Cochrane Training materials into Portuguese, that we have already assigned to the Memsource platform.

Conclusions: We hope that Cochrane Brazil Rio de Janeiro Affiliate is achieving its purpose as an affiliate center, conducting relevant regional activities, encompassing different areas such as education, knowledge translation, and promotion of Cochrane, under the supervision of Cochrane Brazil. Notwithstanding our efforts, we believe that further strategies might be developed for improving the dissemination of Cochrane evidence in social media and the interaction with the other affiliates, building a sense of a living network in our country.

Patient or healthcare consumer involvement: We believe that either the promotion of Cochrane evidence in our region or training healthcare professionals for conducting and interpreting systematic reviews, may improve the quality of healthcare delivered and, as a natural consequence, the final patient care.

Cochrane Clinical Answers: a product retrospective

Pettersen K[1]

[1] Wiley, UK

Background: Cochrane Clinical Answers (CCAs) were developed in partnership between Wiley and Cochrane in response to feedback from clinicians that, while they highly valued Cochrane Systematic Review evidence, the length of a Cochrane Review meant it was difficult to read and apply in a busy, time-sensitive clinical setting. CCAs are aimed at health professionals at the point of care, mimicking the way they approach information-gathering and designed to help close the evidence-to-practice implementation barrier.

Objectives: To review the growth and development of CCAs since the inception of the product.

Methods: CCAs were designed by the Wiley web team in response to a brief based on market research conducted in the United States and New Zealand by Wiley in 2012, which was augmented by feedback from the Cochrane Editorial Unit and user experience testing with practising clinicians. The website was launched in November 2013 with 100 CCAs covering 10 clinical areas and written primarily by US-based clinicians. We conducted a retrospective quantitative and qualitative analysis, reviewing content growth and changes in processes since website launch.

Results: In 2020, we have over 2000 CCAs covering clinical areas from all of the Cochrane Review Groups and written by clinicians worldwide. Our selection processes evolved over time, beginning with building core content across the clinical areas with the highest disease burden (based on hospitalization rates and visits to General/Family Practitioners), and moving to creating CCAs based on all clinically-focused practice-enhancing new and updated Cochrane Reviews. CCAs were launched as an independent website, which in 2018, with the re-platforming of the Cochrane Library, became part of the suite of databases available to all subscribers of the Cochrane Library. The Editorial and Production team has grown from a single Wiley Senior Editor and Production Assistant to a team of 4 Wiley-/Cochrane-based Editors, a Wiley Copy Editor, and 2 Wiley Production Assistants.

Conclusions: CCAs have grown into an important tool within in the Cochrane Library to inform joint clinician-patient treatment choice. In April 2019, we appointed our inaugural Editor-in-Chief of CCAs to work alongside the Editor-in-Chief of the Cochrane Library to derive an ongoing strategy for development of CCAs, which has led to new developments, including preparation of CCAs to support Cochrane's Special Collections and developing a methodology to include CCAs based on Diagnostic Test Accuracy reviews.

Cochrane Library Special Collections: curating evidence in response to the COVID-19 pandemic

Mehta M[1], Abbotts K[2]

[1] Cochrane Editorial and Methods Department, UK; [2] Cochrane Knowledge Translation Department, UK

Background: With the emergence of SARS-CoV-2 and its development into a pandemic, Cochrane identified a need to disseminate evidence on focused topics of relevance to COVID-19. Special Collections, published on the Cochrane Library since 2011, serve this purpose well, curating content on a specific healthcare topic. From February 2020, a series of COVID-19 Special Collections was initiated and published on the Cochrane Library.

Objectives: To describe and evaluate the use of Cochrane Library Special Collections as a mechanism for presenting evidence around specific themes of relevance to key user groups during the COVID-19 pandemic.

Methods: Members of Cochrane's Editorial and Methods, and Knowledge Translation departments identified thematic collections that could present the most relevant evidence from Cochrane Reviews. The Collections were developed by engaging with the Cochrane community and through responding to enquiries from the community. From the start, the collections were updated on a regular basis in response to changes within WHO recommendations and changes in the progress of the COVID-19 pandemic. The Special Collections included links to Cochrane Clinical Answers, where available. Additionally, these Collections were rapidly translated for Biblioteca Cochrane (the Spanish Language version of the Cochrane Library) and also into multiple other languages. The Knowledge Translation team undertook the Collections' dissemination after publication.

Results: The first of the Cochrane Library Special Collections (on evidence relevance to critical care for COVID-19 patients) was published on 11 February 2020. As of 4 April 2020, three Special Collections have been published and updated, with at least three more in development. The current list is available at https://www.cochranelibrary.com/covid-19. In the period up to 4 April 2020, the COVID-19 Special Collections saw unprecedented levels of usage, with more than 194,00 views in total (data from Adobe analytics). The Special Collections are being updated and new Collections are being developed in response to the evolving COVID-19 pandemic, so this work is ongoing, and we will present the latest data.

Conclusions: The production and dissemination of COVID-19 Special Collections highlights that thematic collections of Cochrane Reviews can form a useful and highly accessed part of a response to an urgent need for evidence-based knowledge.

Patient or healthcare consumer involvement: Some degree of involvement in upcoming Special Collection on remote healthcare in the situation of 'social distancing'.

Communicating to support evidence use in practice: knowledge dissemination across multiple channels

Howarth C[1], Neil-Sztramko S[1], Dobbins M[1]

[1] National Collaborating Centre for Methods and Tools, Canada

Background: The National Collaborating Centre for Methods and Tools (NCCMT) champions the use of evidence in public health practice, programs, and policy decision making. The NCCMT uses several communications channels to support public health professionals and organizations use of evidence in practice, and to spread awareness of the importance of evidence-informed public health. We share resources, highlight current evidence and offer training and mentorship opportunities through our Centre or partner organizations.

Objectives: To disseminate knowledge on what works in public health directly to those who use it. We aim to support efficient and effective use of various types of evidence in practice to create positive change in public health in Canada and worldwide.

Methods: The NCCMT uses several key communications channels to reach diverse audiences globally, both online and in-person. Our online communications channels include: a monthly newsletter (the Round-Up); social media platforms including YouTube, SlideShare, Twitter, Facebook, and LinkedIn; monthly webinars; quarterly external promotion blasts where we ask external organizations to promote our new products; and a monthly evidence service to share new systematic reviews added to Health Evidence. We also have exhibitor booths at select conferences to connect in-person with our target audience.

Results: These methods of communication allow us to reach a wide audience to share resources and knowledge. For example, our monthly newsletter reaches over 15,000 subscribers each month. External promotion e-blasts connect with over 85 Canadian organizations who actively promote new NCCMT products and events within their networks. This lets us extend our reach to new audiences who may not subscribe to our newsletter or be familiar with NCCMT. Regularly scheduled webinars are centred on new or existing products and resources for evidence-informed decision making. For example, we host the Online Journal Club where attendees get practical and hands-on training in critical appraisal. From April 2019 to February 2020, the average number of attendees was 57, our highest webinar attendance was 136 attendees. These various communications channels also help us to build partnerships with other organizations and individuals.

Conclusions: The NCCMT has multiple methods for reaching our audience of public health professionals and organizations. Together, these methods of communication are a comprehensive strategy for disseminating and sharing resources and knowledge to promote and create change in using evidence in practice. The multiple communications channels not only help with dissemination but are also a useful tool in building partnerships.

Patient or healthcare consumer involvement: While patients or healthcare consumers were not involved in the development of our communications channels, they may find useful information through our communications channels especially through our partnerships with other organizations.

Content analysis of rehabilitation definitions for research purposes

Arienti C[1], Patrini M[1], Lazzarini SG[1], Kiekens C[2], Negrini S[3]

[1] IRCCS Fondazione Don Carlo Gnocchi, Milan, Italy; [2] Spinal Unit, Montecatone Rehabilitation Institute, Imola, Italy; Department of Physical and Rehabilitation Medicine, University Hospitals Leuven, KU Leuven, Leuven, Belgium; [3] Department of Biomedical, Surgical and Dental Sciences, University of Milan "La Statale", Milan, Italy; IRCCS Istituto Ortopedico Galeazzi, Milan, Italy

Background: Cochrane Rehabilitation (CR) found problems with current rehabilitation definitions in identifying all Cochrane Reviews of rehabilitation interest ("tagging"), in the possible inappropriate use of the term "rehabilitation" (according to Rehabilitation stakeholders perception) in the title of some Cochrane systematic reviews (CSRs) and in the development of the Package of Rehabilitation Interventions with the World Health Organization (WHO).

Objectives: To carry out a content analysis on the current definitions of rehabilitation from three major sources: rehabilitation stakeholders, users (represented in Google) and scientists (represented by CSRs).

Methods: The study included three parts:

1) a survey about rehabilitation definitions used by the major rehabilitation stakeholders represented by the CR Advisory Board;
2) the definitions reported in Google: 6 searches from all the continents have been performed and the first 200 results from each have been stored and studied; and
3) the search of the definition inside CSRs including those that had the term "rehabilitation" in the title and/or abstract.

We collected all the descriptions of rehabilitation inside each selected CSR. We performed:

1) a frequency analysis for the identified definitions;
2) a semantic analysis, looking for the word roots (e.g. function* to include all words like function, functioning, functional, etc) that recurred most; and
3) the number of definitions that included the most used roots.

Results: The survey received 37 answers (response rate 76%) including 31 definitions. The five most common word roots were function*, health*, person*, disabil* and process*. The Google searches returned 1240 unique websites from which 239 total and 128 unique definitions were retrieved. The frequency analysis showed that one definition was repeated 70 time whilst 108 had a single occurrence. The six most common word roots were process*, restor*, health*, person*, function*, and condition*. The content analysis showed that in 71% of the definitions "rehabilitation" is qualified with a noun, being "process" the most common. From the Cochrane Library, 93 CSRs met the inclusion criteria, 52 did not include any definition, 56 (62%) presented the term "rehabilitation" in the title and of these 11 presented a complete definition. The five most common words used in the definitions were: rehabilitation, training, exercise, patient, intervention.

Conclusions: The results showed that a unique rehabilitation definition does not exist. A sort of intuitive and common understanding is present and consistent between different stakeholders, opinion leaders and users. The results from CSRs highlighted the centrality of the intervention. Therefore, a more complete definition of rehabilitation is needed. In this direction, Cochrane Rehabilitation is working on a Consensus process to identify a more specific definition of rehabilitation suitable for research purposes that also defines inclusion and exclusion criteria.

Patient or healthcare consumer involvement: Not applicable.

Contributing to the CENTRAL database and disseminating clinical evidence on Traditional Chinese Medicine in China: managing a better way out

Qiao S[1], Liu X[1], Lu C[1], Jiang Y[1], Li X[1], Liu J[2]

[1] School of Traditional Chinese Medicine, Beijing University of Chinese Medicine, China; [2] School of Health and Social Care, London South Bank University, UK

Background: The Cochrane Central Register of Controlled Trials (CENTRAL) is a bibliography of controlled trials that contains records identified by handsearching the biomedical journals worldwide and creates an unbiased source of data for systematic reviews. The Center for Evidence-based Chinese Medicine at Beijing University of Chinese Medicine (BUCM) has done much work to contribute to this register through identifying and input abstracts on Traditional Chinese Medicine (TCM) that were published in Chinese.

Objectives: To improve the efficiency and accuracy of identifying the references on TCM and optimize the project management.

Methods:

We improved the search strategy and expanded the scope of journals in order to include more published randomized controlled trials with English abstracts, reducing the workload for translation.

We sorted all the problems encountered in the past 10 years into Q&A and developed a standard operating procedure (SOP), which lowered the difficulty threshold for volunteers who's lack of knowledge in evidence-based medicine and greatly decreased the workload of training. The SOP was developed by generations of group leaders based on the feedback given by volunteers, which contains software using, judgement of eligibility and details of the work.

Volunteers cross-checked the results to reduce the errors.

Results: Until 4 April 2020, we have submitted 24,544 records and 2641 more ongoing. Among them, there were 2327 abstracts submitted in 2019 and 2641 abstracts to be submitted in 2020. The amount of available abstract has raised 13.49%. For the 2327 abstracts in 2019, the total length of searching, screening, training and type-in was 82 days, compared with 67 days in 2020. That was a decrease of 18.29%. The 12 problems in the Q&A were mainly in software: 5, judgement of eligibility: 3 and others: 4. These measures did improve the accuracy of searching and the efficiency of input.

Conclusions: In the process of project management, the use of SOP and the possible problems sorted into the form of Q&A can significantly improve the work efficiency. Training of volunteers has also become much easier. The volunteers also significantly reduced the error rate by cross-checking the entries after they were completed. On the one hand, we suggest to continually optimize Procite software. On the other hand, Chinese researchers and clinicians should pay more attention to their English abstract writing skills.

Patient or healthcare consumer involvement: The work on contributing CENTRAL database is very meaningful for researchers and clinicians worldwide. It brings TCM studies to a wider audience and enriches the diversity of clinical research. Experiences on improving the project management might have a positive influence on similar projects all over the world.

depressionscreening100.com/phq: a practice-based perspective to using the Patient Health Questionnaire-9 to screen for depression

Levis B[1], Dehtiarova Y[1], Sun Y[1], Wu Y[1], Benedetti A[1], Thombs BD[1]

[1] McGill University, Canada

Background: The Patient Health Questionnaire-9 (PHQ-9) is the most commonly used depression screening tool in primary care. Most primary studies and meta-analyses of PHQ-9 accuracy focus on sensitivity and specificity, which estimate the probability of a screening result giving true diagnostic status. However, due to the complex nature of conditional probabilities, it can be challenging for clinicians to translate these estimates into more meaningful numbers that reflect the likelihood that a patient has depression given her or his screening score. When high sensitivity and specificity are reported, some clinicians believe it means a positive test is "virtually diagnostic," even if positive predictive value is low. One way to improve understanding of screening tool accuracy estimates is to present information in a format that is more readily understood, such as natural frequencies.

Objectives: To create a user-friendly knowledge translation tool based on sensitivity and specificity estimates from a large individual participant data meta-analysis of PHQ-9 accuracy. The tool allows clinicians to estimate, for a given PHQ-9 screening cut-off and depression prevalence, how many patients would screen positive versus negative, and how many in each group would be correctly versus incorrectly identified.

Methods: We developed a web-tool with a 100-person diagram that self-populates based on user-entered values of major depression prevalence and PHQ-9 cut-off threshold. The tool provides instructions for use, including advice for estimating underlying prevalence and selecting a cut-off, text to explain the numerical results shown in the diagram, and a FAQ section with basic information about depression screening. We consulted family physicians to improve the format and presentation and to ensure that the content is clear and addresses needs of clinicians considering using the tool.

Results: The web-tool can be found at depressionscreening100.com/phq. Illustrated example: As shown in Figure 1, by entering an underlying major depression prevalence of 10% and selecting the standard PHQ-9 cut-off score of ≥ 10, 22 of 100 patients would be expected to screen positive on the PHQ-9, 9 (39%) of whom would meet diagnostic criteria for major depression (true positives) and 13 (61%) of whom would not (false positives). Of the 78 patients expected to screen negative, 77 (99%) would be correctly ruled out (true negatives), while 1 (1%) would be a missed major depression case (false negative). Numbers in the diagram automatically update for different combinations of prevalence and cut-off.

Conclusions: The present web-tool improves clinician understanding of complex diagnostic accuracy estimates from meta-analyses by translating results into natural frequencies that are more readily understood and providing guidance on their meaning and use.

Patient or healthcare consumer involvement: We consulted with several family physicians during development.

Additional file: KT tool

Development of an evidence base to support global hepatitis elimination

Morgan RL[1], Doran-Brubaker S[2], Hiebert L[3], Ward JW[3]

[1] McMaster University, Canada; [2] Evidence Foundation, USA; [3] The Task Force for Global Health, USA

Background: Viral hepatitis is a large global health burden. With hepatitis B virus (HBV) and hepatitis C virus (HCV) accountable for an estimated 1.3 million deaths in 2016 and approximately 325 million persons living with the virus, the World Health Organization defined elimination goals of 90% reduction in incidence and a 65% reduction in hepatitis B- and C-related mortality by 2030.

Objectives: Identify, compile and disseminate evidence on global research priorities to facilitate worldwide hepatitis elimination efforts.

Methods: The Coalition for Global Hepatitis Elimination has adopted a process for identifying and compiling evidence on viral hepatitis. A Technical Advisory Board prioritize global research topics. Evidence identified from peer-reviewed and grey literature sources on each priority topic undergoes rigorous systematic review methods, including screening and data extraction. Summary "cards" of eligible studies across each topic area are tagged based on target population, key intervention, evidence type, and virus targets and disseminated on a central evidence base website.

Results: Currently, the Coalition for Global Hepatitis Elimination evidence base presents critically appraised systematic reviews and guidelines for the topic of timely HBV birth dose vaccination; regional, national, state and city plans for HBV and HCV control and elimination featuring goals, policies, interventions and targets; and care cascade outcomes associated with HCV testing strategies.

Conclusions: Systematic review methods introduce rigor and transparency to the identification of evidence to support global hepatitis elimination efforts. Coupled with an online platform, the evidence summaries and search engine allow for knowledge translation and dissemination efforts to organizations and partners.

Patient or healthcare consumer involvement: No

Disseminating evidence to healthcare practitioners on Instagram: an ecological study

Rapisarda A[1]

[1] Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Italy

Background: Social media are performing a more important status in life day by day. The habit of using them, and particularly Instagram, by healthcare workers, not only in free time but also for work reasons, is growing as an answer to an increased number of scientific influencers and institutions who spread evidence by these tools. At the same time, it seems that researching literature on standard databases has become more complicated due to many barriers (Leach and Gillham, 2008) and identified often as poorness of resources and clinicians' skills in evidence-based practice. This cannot guarantee the optimal clinical practice by the healthcare provider.

Objectives: Describing the new phenomenon of knowledge translation and dissemination of evidence to a wider public of healthcare workers through Instagram.

Methods: We conducted an ecological study. Through the "Evidence For Health" Instagram page, launched on 11 November 2019, we posted an Instagram tailored size image in English, at least once a week, in which is described some evidence from guidelines, systematic reviews or other scientific articles. We divided the posts into six topics, each characterized by color for an immediate recognition, but all similar for graphic and logos. Each post is shared with related hashtags (#EvidenceForHealth) and tags of involved institutions. We also supported the most important world health campaigns. Instagram Stories are used to interact with followers. We analyzed the data with the Instagram statistic tool and a spreadsheet. No ads or paid promotions have been done to increment the page visibility.

Results: At 4 April 2020 followers are 528, with an increment of 402% from 7 January 2020. They are mostly from Brazil (24%), Italy (16%), UK (15%), USA (10%), Australia (3%); mostly women (68%), between 25 and 34 years old (52%). Overall, 91% of the total followers are between 18 and 44 years old. Median reach per post: 473 views.

Conclusions: The optimum increment of followers reveals that sharing scientific good quality materials on Instagram could be a good way to share evidence. Simpleness in text and graphic clarity make reading and appraising evidence intuitive, quick and accessible by most. In-depth analysis of contents is also possible, thanks to the always quoted sources. Limits of this study are the observational design, which makes impossible to define how much SM disseminated evidence is implemented in clinical practice, and the practical issue of the followers as a real healthcare workers because Instagram does not permit to analyze users' profiles. Qualitative and analytic research is needed to assess the value of this method. In conclusion, we can affirm that Instagram could be a new method to share and disseminate evidence quickly worldwide.

Patient or healthcare consumer involvement: The more healthcare workers will use evidence-based practice, the more patients will be involved in healthcare. We are trustful that using social media to diffuse evidence-based practice is a valid method to achieve that.

Dissemination of knowledge translation products via WeChat in China: facilitators and barriers

Zhang X[1], Lu C[1], Fei Y[1], Liu J[1], Yu M[1], Liang C[1], Zhao L[2], Wang Y[2], Li X[1]

[1] Center for Evidence-based Chinese Medicine, Beijing University of Chinese Medicine, China; [2] School of Traditional Chinese Medicine, Beijing University of Chinese Medicine, China

Background: Evidence dissemination and knowledge translation are supplementary to each other. Our team have participated in Cochrane evidence translation including abstract and Plain Language Summary (PLS), podcast, blogshot and press release since 2014, and disseminated the translated products via WeChat, the biggest social media platform in China similar to Twitter and Facebook in Western countries, since 2017. Till March 2020, over 1000 pieces of Cochrane evidence of systematic review have been translated into simplified Chinese, and 272 items of translated products have been posted on WeChat. However, the effect of dissemination is more important than the dissemination itself.

Objectives: To explore the facilitators and barriers in knowledge translation products dissemination and seek a better method of dissemination via the WeChat platform.

Methods: We optimized dissemination formats and display by applying audience feedback, the Cochrane Dissemination Checklist and previous experience. We monitored the visitor data of all WeChat posts , analyzed them by categorizing contents and features, and compared them with previous data to evaluate the impact.

Results: Dissemination: 173 items of translated products have been posted from March 2019 to March 2020, 75% of which are Cochrane PLS. For the COVID-19 outbreak at the beginning of 2020 in China, we translated and introduced Cochrane Special Collection in February 2020 rapidly. Meanwhile, we initiated Evidence-Based Medicine (EBM) Anti-Coronavirus Action via WeChat to collect COVID-19 questions from clinical practitioners and respond with knowledge translation products and all obtainable evidence. In March 2020, the average number of readers per item was 300, four times as many as that in March 2019 (Figure 1). The top 10 reading items in the last 12 months were all about COVID-19, two items on Cochrane Special Collections, and eight on EBM Anti-Coronavirus Action (Table 1).

Figure 1. Average readers per item on WeChat: March 2019-March 2020
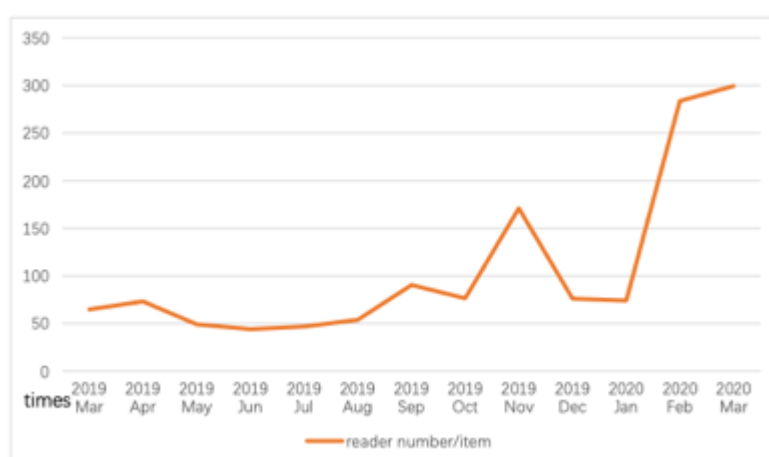
Facilitators: Rapid WeChat dissemination of public health emergency met clinicians' need and worries, which reached the highest reading and followers since established. Responding to an emergency in time is a necessary attitude and method for dissemination.

Barriers: With 3955 followers and 300 reading per item to March 2020, identification was insufficient thus we haven't attracted all the target audiences; there are over 8000 systematic reviews on Cochrane, some without updating, which resulted in a grey area to translate and disseminate.

Solutions: With the establishment of Cochrane China Network Affiliate, Cochrane's title and logo can be used on WeChat subscription account to strengthen identification; content and style of dissemination should be tailored to audience; co-operating with related subscription account or journals, and integrating knowledge translation into training and teaching process to establish the network for dissemination.

Table 1. The top 10 reading items publish on WeChat: March 2019-March 2020

Title	Reader number	Retransmission times	Publish time via WeChat
Infection rate of non-COVID-19 hospitalized population with or without mask*	1403	76	2020/2/20
Can Chinese medicine shorten the recovering time of mild COVID-19 patients?*	774	49	2020/2/26
Comparison of Chinese medicine and regular treatment for viral nucleic acid turning negative or antipyretic time*	748	60	2020/2/24
Manual airway sputum aspiration in patients with COVID-19*	681	39	2020/3/1
What is the effect on lung CT of using traditional Chinese medicine combined with standard treatment compared with standard treatment in severe COVID-19 patients for 14 days in the isolation ward?*	653	57	2020/2/28
Chinese medicine for late recovery of fibrosis after lung injury*	623	47	2020/2/29
Is eight-section brocade (Ba Duan Jin) or Tachi helpful for improving treatment efficacy of COVID-19*	601	49	2020/2/23
Special Collection: Coronavirus (COVID-19): evidence relevant to critical care**	581	45	2020/2/12
Physical interventions to interrupt or reduce the spread of respiratory viruses**	552	46	2020/2/2
Can Vitamin C shorten the recovering time of mild COVID-19 patients?*	443	18	2020/2/26

Note:
* COVID-19 question-answer action
** Cochrane COVID-19 Special Collection

Conclusions: Complying with Cochrane 2020 strategy, WeChat dissemination in China would be improved from multi-aspect and multi-level with Cochrane knowledge translation to yield a higher impact.

E-SCOPE: a strategic approach to identify and implement Cochrane and other high-quality systematic reviews in learning healthcare systems

Henry S[1], Mohan Y[1], Whittaker J[1], Koster M[1]

[1] Evidence-Based Medicine Services Unit, Department of Clinical Analysis, Kaiser Permanente Southern California, USA

Background: With more than 8000 systematic reviews published annually, it is challenging for healthcare systems to review new evidence, prioritize practices that warrant implementation, and ultimately implement them. Cochrane Reviews are among the most valuable, high-quality sources of evidence for improving quality and care delivery, but strategies to implement the findings of these reviews in learning healthcare systems are less well-understood.

Objective: To describe the knowledge translation and accelerated implementation strategies employed by the Kaiser Permanente Southern California Evidence Scanning for Clinical, Operational, and Practice Efficiencies (E-SCOPE) program.

Methods: E-SCOPE uses a strategic search algorithm to identify high-quality studies of interventions that yield improved health outcomes, quality, and/or efficiency of care delivery. To increase confidence about the validity of the effect of the interventions in question, program staff preferentially select systematic reviews and meta-analyses of interventions, further prioritizing Cochrane Reviews. Each quarterly search yields 500 to 1000 abstracts; 5% to 10% are selected for implementation consideration and reviewed for methodological rigor as well as alignment with organizational goals. E-SCOPE staff then work closely with clinical and operational stakeholders to interpret the evidence and translate evidence into a feasible implementation plan, leveraging existing processes and resources. To help ensure successful implementation and sustainability, E-SCOPE project managers oversee implementation efforts, facilitate practice owner identification, and create measurement plans using data drawn from our electronic health record system.

Results: Since 2014, the program has catalyzed the implementation of 42 practices—17 (40%) based on results from Cochrane Reviews—to improve the overall quality of care provided to our 4.6 million members. Using our knowledge translation and use model, the time from study publication to implementation averages 16 months (ranging from 4 to 36 months).

Conclusion: E-SCOPE bolsters the knowledge translation and best-practice adoption process by making optimal use of evidence-based medicine and implementation expertise and leveraging existing channels for practice implementation. Prioritizing the identification of practices with an established, high-quality, and high-confidence evidence base is generalizable and can efficiently and effectively promote rapid learning and implementation within any healthcare system. Approaches to knowledge translation prioritizing Cochrane systematic reviews have the capacity to significantly improve the quality and delivery of patient care. Examples of Cochrane Review-based E-SCOPE initiatives will be discussed.

Patient or healthcare consumer involvement: Patient values and preferences are key determining factors in every implementation plan undertaken by E-SCOPE.

Engagement, dissemination and dialogue: considerations for health researchers using social media for knowledge translation

Elliott S[1], Dyson M[2], Wilkes G[3], Zimmermann G[4], Chambers C[5], Wittmeier K[6], Russell D[7], Scott S[8], Thomson D[9], Hartling L[10]

[1] Cochrane Child Health, Department of Pediatrics, University of Alberta, Canada; [2] Alberta Research Centre for Health Evidence, Department of Pediatrics, University of Alberta, Canada; [3] Faculty of Communications, Mount Royal University, Canada; [4] Knowledge Translation Platform, Alberta SPOR SUPPORT Unit, Department of Pediatrics, University of Alberta; Adjunct Assistant Professor, Department of Community Health Sciences, University of Calgary, Canada; [5] Departments of Psychology and Neuroscience and Pediatrics, Dalhousie University; Centre for Pediatrics Pain Research, IWK Health Centre, Canada; [6] Department of Pediatrics and Child Health, Rady Faculty of Health Sciences, University of Manitoba; Children's Hospital Research Institute of Manitoba, Canada; [7] Research and Knowledge Exchange Consultant, Scientist Emeritus, CanChild Centre for Childhood Disability Research, McMaster University, Canada; [8] Faculty of Nursing, University of Alberta, Canada; [9] Knowledge Translation Platform, Alberta SPOR SUPPORT Unit, Department of Pediatrics, University of Alberta; Cochrane Child Health, University of Alberta, Canada; [10] Alberta Research Centre for Health Evidence, Department of Pediatrics, University of Alberta; Knowledge Translation Platform, Alberta SPOR SUPPORT Unit, Department of Pediatrics, University of Alberta; Cochrane Child Health, University of Alberta, Canada

Background: Despite extensive literature describing the use of social media in health research, a gap exists around best practices in establishing, implementing and evaluating effective social media knowledge translation (KT) and exchange strategies.

Objectives: Our goal was to examine successes, challenges, and lessons learned from using social media within health research, and create practical considerations to guide other researchers.

Methods: The Knowledge Translation Platform of the Alberta SPOR SUPPORT Unit formed a national working group involving platform staff, academics, and a parent representative with experience using social media for health research. We collected and analyzed four case studies which used a variety of social media platforms and evaluation methods. The case studies covered a spectrum of initiatives from participant recruitment and data collection to dissemination, engagement and evaluation. Methods and findings from each case study were summarized, as well as barriers and facilitators encountered. Through iterative discussions, we converged on recommendations and considerations for health researchers planning to use social media for KT

Results: We provide recommendations for elements to consider when developing a social media KT strategy: (1) Set a clear goal and identify a theory, framework, or model that aligns with the project goals and objectives; (2) Understand the intended audience (use social network mapping to learn what platforms and social influences are available); (3) Choose a platform or platforms that meet the needs of the intended audience and aligns well with the research team's capabilities (can you tap into an existing network? What mode of communication does it support?). (4) Tailor messages to meet user needs and platform requirements (e.g. plain language, word restrictions). (5) Consider timing, frequency, and duration of messaging, as well as nature of interactions (social filtering, negotiated awareness). (6) Ensure adequate resources and personnel are available (e.g. content creators, project co-ordinator, communications expert and audience stakeholder/patient advocate); (7) Develop an evaluation plan a priori driven by goals and type of data available (quantitative, qualitative); (8) Consider ethical approvals needed (driven by evaluation, type of data collection).

Conclusions: In the absence of a comprehensive framework to guide health researchers using social media for KT, we provide several key considerations. Future research will help validate the proposed components, and create a body of evidence around best practices for utilizing and evaluating social media as part of a KT strategy.

Patient or healthcare consumer involvement: Parents, patients, healthcare consumers and caregivers were involved throughout some of the case study projects. Additionally, a parent partner was invited to be part of the research team to offer perspective on important outcomes for successful end user engagement.

Epistemonikos is building a 'one-stop shop' of synthesized nutrition evidence and LOVES to enable rapid-learning and decision-making to reduce malnutrition

Naude C[1], Schoonees A[1], Brand A[1], Perez M[2], Verdugo F[3], Baladia E[4], Rada G[5]

[1] Centre for Evidence-based Health Care, Division of Epidemiology and Biostatistics, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa; [2] Epistemonikos Foundation, Santiago, Chile; [3] Epistemonikos Foundation, Santiago; UC Evidence Centre, Pontifical Catholic University of Chile, Santiago, Chile; [4] Evidence-Based Nutrition Network (RED-NuBE), Spanish Academy of Nutrition and Dietetics, Pamplona, Spain; [5] Epistemonikos Foundation, Santiago; UC Evidence Centre, Pontifical Catholic University of Chile, Santiago; Internal Medicine Department, Faculty of Medicine, Pontifical Catholic University of Chile, Santiago, Chile

Background: Malnutrition, in all its forms, affects every country, wealthy people, poor people, and most of the world's population at some point from infancy to old age. Given the exponential growth in the number of systematic reviews about nutrition and the complexity of skills needed to identify them, new approaches are required. Aligned with the Epistemonikos Foundation's vision, a dedicated database of all nutrition-relevant synthesized evidence would help bring evidence closer to those who need it. This initiative aims to support access to and timely production of synthesized evidence to enable rapid-learning and decision-making to reduce malnutrition.

Objectives: To share our approach, progress and lessons learned in planning, curating and building the nutrition database

Methods: The collaboration was co-ordinated and established by the Epistemonikos Foundation, with researchers, information technologists and nutrition experts from four countries contributing. For the database, we developed draft review eligibility criteria. The potentially eligible articles being screened come from two sources: traditional search in the Epistemonikos database and reviews in the Living OVerview of Evidence (LOVE) platform. LOVE is a platform that retrieves all evidence hosted in the main Epistemonikos database and classifies the information using artificial intelligence algorithms. For the traditional search, we developed a sensitive and complex Boolean search strategy to identify potentially eligible reviews. For calibration across multiple collaborators, we piloted screening in a sample of retrieved records. This further enhanced the eligibility criteria, particularly in areas of likely uncertainty. Two screeners independently assess the retrieved records, consulting a third if needed. For the second source, we are developing a LOVE for each relevant nutrition topic.

Results: The comprehensive final search yielded over 60 000 records. Screening is ongoing, co-ordinated through Red de Nutrición Basada en la Evidencia (RED-NuBE), the Centre for Evidence-based Health Care, Stellenbosch University and Epistemonikos, and database architecture and design is in progress. Screening for a pilot nutrition LOVE showed high sensitivity. Simultaneous screening via the creation of relevant nutrition LOVEs is ongoing. Records included in LOVEs are automatically removed from the database search yield.

Conclusions: We are working to establish an up-to-date, interconnected 'one-stop shop' of synthesized nutrition evidence and LOVEs to enable timely data- and evidence-informed changes at all levels of health systems to address the universal malnutrition burden. The systematic simultaneous screening approach to building the nutrition database and LOVEs strives to be efficient, while upholding high methodological standards.

Patient or healthcare consumer involvement: No direct patient or healthcare consumer involvement.

Evidence ecosystem for knowledge translation in Nigeria: the bottlenecks and gaps to evidence uptake

Oduwole O[1]

[1] Achievers University, Nigeria

Background: The main participants in this ecosystem were researchers from Cochrane Nigeria and other Cochrane contributors, healthcare professionals, guideline developers, policymakers, media organizations, consumer networks and patients. Within the last decade, researchers from Cochrane Nigeria have conducted two priority setting exercises of systematic reviews of interventions for communicable and non-communicable diseases to identify priority health topics. The process involved stakeholders like healthcare providers, policymakers, consumer representatives and patients who identified and ranked review questions. We have tried to disseminate the findings from the priority setting exercises.

Objectives: To examine the challenges and gaps to knowledge translation within the evidence ecosystem in Nigeria.

Gaps and bottlenecks to evidence uptake: It is worrisome that there is a gap between the producers of evidence and policymakers in Nigeria (Figure 1). Partly due to lack of pull for evidence on the part of policymakers and senior healthcare professionals largely due to ignorance. Because of this, most policy decisions and practice guidelines are not based on evidence.

Conclusions: Knowledge translation is a new area for us and we are making progress in this regard. Recently, there has been an effort on the part of Cochrane Nigeria to engage professional groups so that they can appreciate and use systematic reviews for practice guideline development. We also need to continue to engage other professional groups and different departments of the Ministry of Health on the need to have evidence-based practice guidelines.

Patient or healthcare consumer involvement: The priority exercises conducted involved patients and consumer advocacy groups. Additionally, a consumer advocate contributed to the outcomes of the systematic review protocols before publication.

Additional file: Evidence ecosystem

Evidence synthesis, practice guidelines and real-world prescriptions of new generation antidepressants in the treatment of depression

Luo Y[1], Chaimani A[2], Ostinelli EG[3], Kataoka Y[4], Ogawa Y[5], Cipriani A[3], Salanti G[6], Furukawa TA[1]

[1] Department of Health Promotion and Human Behavior, School of Public Health in the Graduate School of Medicine, Kyoto University, Japan;
[2] Université de Paris, Research Center of Epidemiology and Statistics (CRESS-UMR1153), INSERM, INRA, France; [3] Department of Psychiatry, University of Oxford, Oxford Health NHS Foundation Trust, Warneford Hospital, UK; [4] Hospital Care Research Unit, Hyogo Prefectural Amagasaki General Medical Center, Japan; [5] Department of Healthcare Epidemiology, School of Public Health in the Graduate School of Medicine, Kyoto University, Japan; [6] Institute of Social and Preventive Medicine, University of Bern, Switzerland

Background: Various gaps may exist in the process from evidence synthesis to its implementation. Evidence synthesis may have been biased and late; guidelines may not have reflected the most up-to-date evidence; clinicians may have made decisions without being guided by relevant evidence.

Objectives: We aimed to examine the gaps between the ideally synthesized evidence, guideline recommendations and real-world clinical practices in the prescription of new generation antidepressants for major depressive disorder (MDD) through the past three decades.

Methods: We used cumulative network meta-analyses (NMAs) to represent ideally synthesized evidence over time. We built a Shiny web app to perform and present NMAs interactively. The analyses were based on a comprehensive dataset of randomized controlled trials of 21 antidepressants in the acute treatment of MDD. The primary outcomes were efficacy (treatment response) and acceptability (all-cause discontinuation), and treatment effects were summarized via odds ratios. We evaluated the confidence in evidence using the CINeMA (Confidence in Network Meta-Analysis) framework for several consecutive NMAs. The Shiny app presents network plots, two-dimensional plots combining efficacy and acceptability for each drug, forest plots, league tables, and funnel plots. We identified and extracted recommendations from several representative practice guidelines (Table 1). We estimated the real-world prescription patterns of antidepressant monotherapy for MDD, using the Medical Expenditure Panel Survey, a nationally representative database in the US. We evaluated the gaps between the results from NMAs, recommendations and prescriptions between 1990 and 2016.

Results: The Shiny app is accessible at https://cinema.ispm.unibe.ch/shinies/GRISELDA/. It indicates dramatic changes of drugs with relative superiority, and potentially exaggerated performance of newly approved drugs. Guidelines are usually updated every 5 to 10 years (Table 1). All proposed specific antidepressant recommendations, and most recommended drugs showed relatively acceptable efficacy and acceptability in the NMA at that time. However, fluvoxamine and duloxetine, although they had already appeared barely satisfactory in efficacy, acceptability and credibility, were still being recommended more than five years later. In the US, the prescriptions for some newly launched drugs were very large, even without formal recommendations or firm evidence (Figure 1).

Conclusions: Our study revealed the gaps from evidence to real-world practice in the antidepressant treatment of MDD. Considering the initially amplified effects in the evidence, recommendations about new drugs should be made with caution. Since evidence is the cornerstone in the process of evidence-based medicine, it should be kept up to date using rigorous synthesis methods. We also provided an example of how to present and visualize the evidence interactively through a Shiny app, in order to help policy-makers and clinicians comprehend the evidence.

Patient or healthcare consumer involvement: None.

Additional files: Table 1; Figure 1

Experience of promoting rapid increasing evidence to knowledge translation

Lu CL[1], Zhang X[1], Fei YT[1], Liu JP[1], Li X[1]

[1] Centre for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, China

Background: With rapid increasing evidence, it is a huge and time-consuming process to promote the latest high-quality evidence to be translated and disseminated. Our centre is responsible for two Cochrane China working groups: Translation & Dissemination, and Traditional Chinese medicine (TCM). We have been involved in the Cochrane Central Register of Controlled Trials (CENTRAL) project since 2007 and a simplified Chinese translation project since 2014.

Objectives: To share the experience of promoting Cochrane knowledge translation (KT) and discuss the further strategy based on Cochrane China network.

What we have done to promote KT? Based on our series of normative standard operating procedures (SOP) to recruit, support, and maintain volunteers with KT since 2017, we have doubled the quantity of 248 active volunteers, published 1418 translations on Cochrane.org, and have 3955 subscribers to the public Wechat account. The maintained SOP helps involve and manage increasing volunteers to finish the latest and important projects effectively. Since the COVID-19 outbreak in January 2020, we first collected 50 related Cochrane Reviews involving public health prevention and vaccine, diagnosis, treatment and prognosis, and called on volunteers in the Wechat group to translate them immediately (Fig 1). We published these 50 reviews in one week and released 28 special posts via Wechat, which were read 4153 times until 11 March 2020. Based on the Cochrane Dissemination Checklist, we highlighted the important sentences in each posts (Fig 2) and listed the links of the previous relative posts at the end of passage (Fig 3) so that one post can link more than one evidence and people can read all relative translations. We added the English text behind the translation for experts to read better. The routine SOP and management help us to react to the emergency translation and dissemination more quickly.

What we can do to promote KT based on Cochrane China network? With the establishment of Cochrane China network, each group has their interests in health care and they are skilled at professional knowledge, so it will be easy to identify the wider audience and their need in the specific topic. We can promote KT by the collaboration of producing the high-quality evidence, translating priority evidence, disseminating more evidence, and implementing best evidence in practice. Several groups such as the Public health working group affiliated to Chongqing Medical University were involved in the COVID-19 special collections translation and editing according to their area of interest, which improved the efficiency. The new model can decrease the length of time to assign translations to individuals and improve the quality of translation. Networks will provide multiple new possibilities to promote KT in a rapid and specialized method.

Patient or healthcare consumer involvement: KT helps patient and healthcare consumers to pick out useful information from the rapidly increasing amount of evidence.

Additional files: Figure 1; Figure 2; Figure 3

External validity of definition of rehabilitation in health

Negrini S[1], Arienti C[2], Patrini M[2], Lazzarini SG[2], Kiekens C[3]

[1] Department of Biomedical, Surgical and Dental Sciences, University of Milan "La Statale", Milan; IRCCS Istituto Ortopedico Galeazzi, Milan, Italy;
[2] IRCCS Fondazione Don Carlo Gnocchi, Milan, Italy; [3] Spinal Unit, Montecatone Rehabilitation Institute, Imola, Italy; Department of Physical and Rehabilitation Medicine, University Hospitals Leuven, KU Leuven, Leuven, Belgium

Background: A definition of rehabilitation in health (medical rehabilitation) (RIH-MR) must be valid both for those in and out of the field. In scientific papers, internal validity refers to the possibility of a study to explain its findings without biases, external validity to generalizability of results to other contexts. Here we consider first the possibility to describe with the definition all we do in RIH-MR, and second the possibility to rule out anything that is not RIH-MR. In Cochrane Rehabilitation (CR) we first faced the problem when a Cochrane Systematic Review (CSR) on penile rehabilitation was published. The review dealt only with drugs; we considered inappropriate the use of the term rehabilitation but lacked a definition of RIH-MR to confirm our thesis.

Objectives: To quantify the phenomenon looking at all the CSRs claiming to study rehabilitation and comparing them to the definitions provided by CR, PubMed (MeSH Term) and the author judgment.

Methods: We performed a search of the all CSRs published in the Cochrane Database of Systematic Reviews with the presence of the term "rehabilitation" in the title. Exclusion criteria were: editorial, updated CSRs, withdrawn CSRs. We performed a content analysis of the CSRs included/excluded by each classification. For each field/intervention, the author judged the classifications of CR and PM coherent if all CSRs were included or excluded, incoherent if some CSRs were included and others excluded.

Results: Out of 14,816 records, we found 139 papers with the term rehabilitation in the title. We analyzed 89 CSRs and CR included 94.4% of CSRs, the author 91%, PubMed only 50.5%. We judged four reviews and Cancer and vestibular rehabilitation fields to be non-RIH-MR by all classifications. CR incoherently excluded one review related to exercises cardiac rehabilitation. The author excluded four reviews included by CR as the provided interventions were not considered rehabilitation. All the other CSRs were judged consistently by CR and the author. Only the neurological field was coherently included by PM, albeit excluded in all cases with application of cognitive and neuropsychological interventions. We did not find coherence for all the other fields and interventions.

Conclusion: The results highlight the possibility to exclude "single interventions" from the definition, even if they are classically considered in the field of rehabilitation. Consequently, we could provide a definition of "rehabilitation intervention", and this could be considered in terms of the professionals providing the intervention.

Patient or healthcare consumer involvement: Not applicable.

Four framework synthesis approaches to using theory in knowledge production

Brunton G[1], Oliver S[2]

[1] OntarioTech University, EPPI-Centre; University College London, Canada; [2] EPPI-Centre, Institute of Education, University College London, UK

Background: Framework synthesis is increasingly used in systematic reviews of healthcare practice and policy. However, it appears to be used in different ways.

Objectives: To demonstrate how framework synthesis methods have been used and their contribution to research synthesis methods.

Methods: We conducted a systematic review and search update of reports which applied, illustrated or discussed framework synthesis. Using framework synthesis methods, we incorporated findings into a previously developed and evolving conceptual framework. We derived higher order themes using constant comparative analysis.

Results: Searches identified 61 publications which included 37 applied reviews and 24 illustrative or situated reports. These described varied contexts, concepts, challenges and processes of framework synthesis. Framework synthesis is a realist method that uses a spectrum of approaches, the choice of which depends on the extent of existing developed theory. Four approaches emerged. Where theory was underdeveloped, early sense-making either: 1) constructed multidimensional frameworks with stakeholders' expert knowledge to fractionate or disaggregate the data into meaningful subsets; or 2) used widely recognised concepts to frame/provide a shell and then qualitatively identify sub-themes. In other cases, 3) well-established theory closely matching the topic was translated into a framework to test the fit between study data, framework and theory. Where the topic lacked an exact theoretical fit, 4) acceptable a priori theories were identified, potentially refined and adopted ('best fit'). While not currently consistently used, stakeholder engagement was widely advocated. Where stakeholders were engaged in knowledge production, findings suggested that stakeholders helped to make sense of complex issues underpinning a health condition by addressing challenges to understanding stakeholder priorities, constraints, or to sense-check theory as it developed.

Conclusions: Our review establishes a spectrum of framework synthesis applications that invoke a framework as: an established theory to test; an analogous theory to be refined; a theoretical shell to hold emergent themes; or a multidimensional framework to fractionate then integrate heterogenous data. The choice of approach depends on the fit between data and existing theory or the scale and heterogeneity of the literature. We recommend exploration of these uses of framework synthesis beyond health.

Patient or healthcare consumer involvement: Framework synthesis, when used in conjunction with stakeholder involvement, can help to develop and explore theory that underpins health issues and interventions that inform policy.

From systematic review to knowledge translation in emergency medicine: contribution of the Cochrane Pre-hospital and Emergency Care field

Lvovschi VE[1], Meyran D[2], Magee K[3], Dumouchel J[4], Auffret Y[5], Nekhili N[6], Beroud S[7], Yordanov Y[8], Miroux P[9], Trinh-duc A[10], Jabre P[11]

[1] Urgences CHU Rouen, France; [2] SMUR – Bataillon de Marins Pompiers de Marseille, France; [3] Dalhousie University, Halifax Infirmary, Canada; [4] Urgences CHU Tours, France; [5] Urgences CHU Brest, France; [6] SMUR Gonesse, France; [7] Urgences CHU Lyon, France; [8] Urgences CHU Saint-Antoine, France; [9] Urgences CHU Angers, France; [10] Urgences CHU Agen, France; [11] SAMU de Paris, CHU Necker Enfants Malades, France

Background: Knowledge Translation is a challenge in Emergency Medicine. There is a need to improve the impact of evidence-based medicine on physician practice. The Cochrane Pre-hospital and Emergency Care field (Cochrane PEC) is involved in the dissemination of relevant Cochrane systematic reviews (CSR). In January 2017, The Cochrane PEC embarked on a process to summarize relevant CSR for emergency physicians called "Practical Evidence About Real Life Situations" (PEARLS) and to ensure their dissemination.

Objectives: To assess the method, implementation processes and results of the production and dissemination of Cochrane PEC PEARLS.

Methods: Twelve international pre-hospital or in-hospital emergency physicians, using various communication channels, worked in partnership on a regular basis over a short timeframe. We report the methodological steps used and model construction from Cochrane Review selection to journal publication. We also detail a quantitative assessment of this activity.

Results: Methodology optimization and network development took twelve months. Five Cochrane PEC members identified CSR relevant to emergency medicine. Two Cochrane PEC members tagged these reviews in the Cochrane central server ("ARCHIE"). Through consensus, we identified reviews demonstrating marked benefit or harm and selected them for PEARLS development. The PEARLS, limited to 200 words, is written by two members before being presented to the working group. The title of the PEARLS is based on the conclusion provided by the CSR to underscore relevance. The background highlights the clinical and epidemiological conditions of the intervention evaluated. The clinical question is described in one sentence. Main results are presented in one or two informative sections and the limits are identified. From June 2017 to January 2020, 49 PEARLS were written (one per month) and presented during the 29 working group meetings. Subsequent discussion and approval during the meeting took an average of 30 minutes. One hundred and thirty-five hours were devoted to the entire PEARLS production. The PEARLS were published in two scientific journals of Emergency Medicine: 20 in "Mediterranean Journal of Emergency Medicine" in English and 29 in "Annales Françaises de Médecine d'Urgence" in French. Recently, two original Cochrane summaries were published in "Emergencias" in Spanish (impact factor = 3.5). No specific financial resources were allocated to this activity.

Conclusions: The dissemination of Cochrane PEARLS is the result of an international collaboration of dedicated emergency physicians. This very successful standardized knowledge translation model is easily reproducible. Our next step is to study the impact of this knowledge translation activity.

Get the picture! New guidance on choosing images for sharing Cochrane evidence

Chapman S[1], Ryan-Vig S[1]

[1] Cochrane UK, UK

Background: Choosing a good image is an important part of creating dissemination products, such as blogs and blogshots, based on Cochrane evidence. Images can also increase reach and engagement when sharing Cochrane evidence on social media. It can be challenging to choose images that are accurate, relatable and sensitive and which also conform to Cochrane Brand Guidelines. To help people choose images to share Cochrane evidence, Cochrane UK is developing an Images Checklist and Guidance, based on the Cochrane Dissemination Checklist introduced in 2019. The Images Checklist will be launched at the Colloquium in Toronto.

Objectives: (1) Introduce participants to the new Images Checklist, its development and why it is important; (2) Share general principles of good and bad practice when choosing images; (3) Show illustrated examples of images to accompany a variety of Cochrane evidence.

Conclusion: We expect that many of the items in the Dissemination Checklist and Guidance will have direct applicability to the Images Checklist. New items will need to be added to the Images Checklist, drawing on feedback from consumers and our experience of choosing images when sharing Cochrane evidence. We hope that the introduction of an Images Checklist will support those working within Cochrane to share evidence for a range of stakeholder audiences and in a variety of dissemination products.

Patient or healthcare consumer involvement: Cochrane UK will involve consumers in the development of the Images Checklist. Prior to beginning work on the Checklist, Cochrane UK along with Cochrane Common Mental Disorders sought feedback from the public and from people with lived experience of mental health problems about appropriate image use when illustrating mental health topics. The Checklist draws on this feedback.

How clear and simple for lay people are the Plain Language Summaries and abstracts in Spanish? A readability assessment of 2019 Cochrane Systematic Reviews

Meza N[1], Madrid E[1], Franco J[2], Garnham R[3], Olguin-Sepúlveda P[3], Urrútia G[4], Escobar-Liquitay C[2], Perez-Bracchiglione J[1]

[1] Interdisciplinary Centre for Health Studies (CIESAL), Cochrane Associate Centre, School of Medicine, Universidad de Valparaíso, Chile; [2] Research Department, Cochrane Associate Centre, Instituto Universitario Hospital Italiano, Argentina; [3] Cochrane Associate Centre, School of Medicine, Universidad de Valparaíso, Chile; [4] Iberoamerican Cochrane Centre, Biomedical Research Institute Sant Pau, CIBER Epidemiología y Salud Pública, Barcelona, Spain

Background: Plain Language Summaries (PLS) of Cochrane Systematic Reviews (CSR) must be written and translated in clear and simple language to serve a relevant aim of knowledge translation, and must be accurately summarized in a succinct and readable style, as a way to deliver the results to a broad variety of audiences and cultures. Cochrane has approved standards for the elaboration of the message of PLS, and the regional Cochrane groups have an important role in implementing them correctly in their own languages. Although there have been published some investigations about readability of the PLS compared to abstracts, the readability of CSR in Spanish remains unclear.

Objectives: To analyse the readability of the Spanish language PLS and abstracts of CSRs of interventions published during 2019.

Methods: We undertook a cross-sectional study assessing CSRs published during 2019, which had been translated into Spanish language. We excluded protocols, withdrawn reviews, Cochrane Clinical Answers and non-intervention CSRs (prognostic, diagnostic, etc). We retrieved the CSRs and extracted the following data: Title, authors, country of affiliation of the corresponding author, Review Group, and text of PLS and abstract in both English and Spanish (including titles). We assessed the readability of the Spanish language abstracts and PLS using the readability INFLESZ scale (or Szigriszt Pazos' perspicuity formula) – validated for measuring readability of Spanish language texts – which scores the difficulty to read a text from 0 (very hard) to 100 points (very easy) (see Table 1). We used means, standard deviations and frequencies for descriptive analyses, while for inferential analyses, we performed a two-sample mean-comparison by Student's T test.

Table 1. Interpretation for readability INFLESZ scale scores

Readability INFLESZ scale score	Interpretation
0 - 40	Very hard
40 - 55	Moderately hard
55 - 65	Normal
65 - 80	Moderately easy
80 - 100	Very easy

Results: We retrieved a total of 546 CSRs, of which 505 (92.49%) had a PLS and abstract translated into Spanish. The resulting INFLESZ scores for abstracts were 56.68 (± 6.59), while PLS scored 50.87 (± 6.62), which was significantly different (P < 0.001), i.e. PLS were less readable.

Conclusions: The Spanish language abstracts and PLS of CSR published during 2019 have a readability normal and moderately difficult, respectively, for lay people according to the INFLESZ scale. Moreover, we found that PLS are significantly less readable than abstracts in all the assessed CSR. We conducted a

parallel study on English abstracts and PLS and we understand that some of these differences in readability can be carried in the translation of the original articles.

Patient or healthcare consumer involvement: Adequate health literacy is crucial for patients and consumers in order to be able to understand the health information regarding their conditions, and to take an active role in their health care. Low readability is an important barrier for Spanish-speaking health staff and consumers to make the most in using such information.

How much do Cochrane editors and authors, and healthcare consumers accept the new reporting style?

Glujovsky D[1], Ciapponi A[1], Bardach A[1]

[1] Instituto de Efectividad Clínica y Sanitaria, Argentina

Background: For decades the P value-based interpretation and reporting of trials' results dominated the publications. Nowadays, the scientific community agrees that this binary approach is not enough. Institutional reform is necessary for moving beyond statistical significance in any context – journals, education, academic incentive systems, or others. Several papers in this special issue focus on reform. Goodman (2019) notes considerable social change is needed in academic institutions, in journals, and among funding and regulatory agencies. Trafimow (2019), who added energy to the discussion of P values a few years ago by banning them from the journal he edits, suggested five "non-obvious changes" to editorial practice. These suggestions, which demand re-evaluating traditional practices in editorial policy, will not be trivial to implement but would result in massive change in some journals. Version 6 of the Cochrane Handbook, in the chapter for "Interpreting results and drawing conclusions" makes some comments about the point estimate, the confidence interval and the P value, and suggests some narrative statements. However, we are not sure how much Cochrane editors and authors and healthcare consumers agree with these suggestions and, even more importantly, if they all reach the same conclusions when they look at specific results.

Objectives: To evaluate how Cochrane editors and authors interpret different results and to analyse which form of reporting they agree with mostly.

Methods: We are conducting an online survey among Cochrane editors and authors, and we will also conduct a survey among a sample of healthcare consumers. They are receiving some different results and have to state which is the reporting style that fits better with their interpretation of the results.

Results: will be shown at the Colloquium.

Conclusions: It will be interesting to see the level of agreement or disagreement among the Cochrane community, and what healthcare consumers think about it.

Patient or healthcare consumer involvement: Healthcare consumers will participate in the survey.

How to integrate knowledge syntheses into learning health systems: reflections from academic-health system partnerships in Alberta, Canada

Thomson D[1], Brooks S[2], Hartling L[1]

[1] Cochrane Child Health; University of Alberta, Canada; [2] University of Alberta, Canada

Background: The Alberta SPOR SUPPORT Unit Knowledge Translation (KT) Platform is a Canadian research-funded initiative established to support research on patient-identified priorities, connect relevant stakeholders, and apply evidence-based solutions to improve patient outcomes through various KT services including knowledge synthesis (KS). In the first phase of AbSPORU (2014 to 2021), we were mandated to establish support services by building relationships with health system partners. Aligned with the growing interest in learning health systems (LHSs) our Phase 2 mandate (2021 to 2025) has expanded from working with health systems to contributing to LHSs. In response, we have worked to align existing academic, health system, and government partnerships to build an initiative called the 'Implementation Science Collaborative' (ISC), which aims to accelerate the implementation and improvement of evidence-based practices in Albertan settings.

Methods: To build the ISC model, we engaged with key stakeholders to learn about existing barriers and facilitators to moving evidence into practice. This engagement helped us identify opportunities to 1) align stakeholder priorities, and 2) leverage existing processes and infrastructures to facilitate LHS procedures that accelerate implementation of evidence-based practice. KS service providers perform a dual function as LHS stakeholders and as facilitative infrastructure. Including KS services providers as stakeholders helps them develop highly relevant and usable reviews for multi-stakeholder groups, in turn meeting their own goals of promoting evidence-based practice. By including KS providers in LHS operational processes, the ISC aims to institutionalize their services as a mechanism to facilitate access to priority-aligned evidence.

Results: Misaligned stakeholder priorities is a primary barrier for establishing academic-health system collaborations that are essential for emerging LHSs. Thus, a central activity of the ISC is designed to compile priorities of different stakeholders, identify synergies and collaboration opportunities across sectors, and provide support for well aligned implementation projects. KS service providers are essential to LHSs as they are the gateway to designing evidence-based health care. By incorporating KS providers into both engagement and operational components of the ISC model, the collaboration is improving the province's capacity to provide appropriate and timely evidence-based interventions to Albertans. As cross-sectoral priorities emerge, providers such as the KT Platform can develop KS products to incorporate into the LHS in ways that support high priority health innovation implementation and improvement strategies.

Conclusion: Integrated KS provision in LHSs can provide directly relevant and highly usable KS products to inform implementation of health innovations and quality improvement. The ISC model and functions will be presented in detail.

Patient or healthcare consumer involvement: None

How will we know whether we're making a difference? Evaluating Cochrane's knowledge translation impacts

Schmidt B[1], Head K[2], Anthony J[2], Kredo T[1], de Silva D[3]

[1] Cochrane South Africa; South African Medical Research Council, South Africa; [2] Cochrane Central Executive, UK; [3] The Evidence Centre, UK

Background: Cochrane is known for providing high-quality research syntheses. Much work has been done to try to encourage the use of this evidence in health-decision making. But how well are we achieving our goals? In 2020, the Cochrane Knowledge Translation (KT) Department considered how to evaluate their KT programme and how to support others in Cochrane in integrating evaluation into their work. Evaluation is fundamental to rapid-learning health systems. In order to learn and adapt, we need to evaluate not only what impacts we are having, but what is helping and hindering us from getting evidence into practice.

Objectives: We aimed to understand:

- the facilitators and barriers to using Cochrane evidence amongst our four KT target audiences: consumers and the general public, healthcare practitioners, policymakers and healthcare managers, researchers and research funders;
- what Cochrane is doing to address those facilitators and barriers;
- how we will know whether KT activities are making a difference; and
- what support others in the Cochrane community need to evaluate their KT work.

Methods: We collected feedback from more than 200 people within and external to Cochrane using interviews, discussion groups and online surveys. A working group made up of Cochrane community members helped to develop and test evaluation tools. Using this information, we created a visualization for our KT activities, which was used to develop an evaluation plan.

Results: Interviews with external audiences indicated that three-quarters of the healthcare consumers, practitioners, researchers and policy makers said that the work Cochrane does could be useful to them, but fewer than one third had actually used evidence syntheses in the past three years – and most had not used ours! Barriers included accessibility, lack of confidence and perceived issues with our content. Cochrane's KT team and fields, centres and groups all have work underway to address these issues. We have developed a framework and data collection tools for evaluating Cochrane's KT work at strategic level. We have also developed a suite of resources and training to help Cochrane groups evaluate their own KT activities globally. These tools are available for use by the Cochrane community.

Conclusions: Cochrane's vision is a world of improved health, with decisions informed by high-quality, relevant and up-to-date research evidence. Essential to achieving Cochrane's strategic objectives is an understanding of whether Cochrane is making a difference and how the organization, outputs and processes could be developed to further achieve the aims. We now have the tools to start to help us do this. The next step is to start embedding those tools so that evaluation and the rapid-learning healthcare approach becomes fundamental to what we all do.

Patient or healthcare consumer involvement: Representatives from healthcare consumer organizations around the world provided feedback about the things that facilitated and acted as barriers to them accessing and using evidence.

Human post-editing to evaluate and compare the quality of three machine translation engines for Russian translations of Cochrane Plain Language Summaries

Yudina E[1], Gabdrakhmanov A[1], Ried J[2], Ziganshina LE[1]

[1] Cochrane Russia, Russian Federation; [2] Cochrane Central Executive, Germany

Background: Translation and multi-language activities are a priority for Cochrane and critical to its Knowledge Translation (KT) activities to enable uptake of Cochrane evidence globally. High-quality machine translation (MT) can help facilitate efficient delivery of translated content in different languages, including Russian. The potential of MT and available options have been rapidly growing recently. It is important to understand which MT engine performs best for specific languages.

Objectives: To compare and evaluate the quality of three off-the-shelf MT engines for Russian translations of Cochrane Plain Language Summaries (PLS) using human post-editing.

Methods: We compared three MT engines, DeepL MT, Google Translate MT and Microsoft Translator MT, as part of our standard translation workflow within Memsource translation management system. We selected 90 PLSs published in the Cochrane Library from May 2018 to April 2019 and not yet translated into Russian. We translated 30 PLSs each with the three MT engines. We invited 10 experienced volunteer translators and editors to post-edit the machine translations, and randomly assigned them three pre-translated PLS per MT engine, so nine PLS in total. Two editors performed a second and final review. Memsource Machine Translation Quality Estimation (MTQE) provided an initial artificial intelligence-powered estimate of how much editing would be required for each machine translated text. The Memsource analysis feature allowed precise recording and numerical presentation of the amount of human editing required for each MT engine at both editing steps. We analysed and interpreted those data after machine translation and each consecutive human post-editing step to assess the quality of the three MT engines.

Results: Google Translate MT had on average the highest ratings for translation quality: the overall quality estimate after machine translation was the highest, whilst the amount of required human revisions was the lowest at both editing steps. DeepL MT followed closely after the Google Translate MT showing overall slightly lower quality estimates after machine translation and requiring overall slightly more editing. Microsoft Translator MT had the lowest quality estimate ratings and required the most revisions at both human editing steps.

Conclusions: Among the three MT engines that we tested, Google Translate MT appeared to perform best for Russian translations of Cochrane PLSs, while DeepL MT also showed good results. We would recommend Google Translate, and DeepL MT as the second-best option, for machine translation of Cochrane PLSs into Russian. Future developments in MT research and the MT market may mean that a different MT engine will become preferable. While Google Translate MT performed slightly better than DeepL, we have opted for DeepL as default MT engine in our translation workflow, as DeepL offers preferable IP and copyright terms.

Patient or healthcare consumer involvement: About one fifth of the volunteer editors in our study were consumers.

Implementation and evaluation of the RAISE initiative: supporting low- and middle-income countries to adapt and implement health systems guidelines

Fahim C[1], Quinn de Launay K[1], Baddeliyanage R[1], Davenport-Huyer L[1], Marten R[2], Hadush Gebregiorgis A[3], Wiysonge CS[4], Puchalski-Ritchie L[1], Tricco A[1], Pham B[1], Peer N[1], Tuncalp O[5], Akala O[2], Langlois E[6], Straus SE[1]

[1] Knowledge Translation Program, St. Michael's Hospital, Unity Health Toronto, Canada; [2] Alliance for Health Policy and Systems Research, World Health Organization, Switzerland; [3] World Health Organization Ethiopia, Ethiopia; [4] South African Cochrane Centre, South Africa; [5] Human Reproduction Programme World Health Organization, Switzerland; [6] WHO Partnership for Maternal, Newborn & Child Health PMNCH, Switzerland

Background: There is limited evidence on how to build capacity for health policy and systems research (HPSR) implementation in low- and middle-income countries (LMIC). The World Health Organization's Research to Enhance the Adaptation and Implementation of Health Systems Guidelines (RAISE) initiative aims to support six LMIC teams to adapt and implement health systems guidelines. We describe our approach to developing a Technical Support Centre (TSC) in response to RAISE teams' needs and our plan to evaluate the impact of the TSC.

Objectives: To develop and evaluate a tailored TSC support program based on RAISE teams' needs.

Methods: Informed by the Knowledge to Action model, the Theoretical Domains Framework, and the Consolidated Framework for Implementation Research, we conducted a needs assessment to develop the TSC activities. The needs assessment was composed of key informant interviews and surveys with RAISE participants from six LMICs (Nigeria, Ghana, Mozambique, Zambia, India, Colombia). Two researchers analyzed data using a rapid analysis approach. We developed the TSC activities in response to the identified needs.

Results: A total of 28 RAISE participants responded to the needs assessment survey and/or participated in the interviews. Participants included 6 principal investigators, 13 co-investigators, 5 research co-ordinators/assistants, and 4 decision-makers/knowledge users. The teams requested support on: engaging and training policy-makers to support evidence-based implementation in LMIC contexts and KT theories, models and frameworks to support health system guidelines adaptation and implementation. Anticipated challenges to health system guidelines adaptation and implementation included: limited institutional resources and supports, maintaining policymaker and stakeholder engagement, delays in regulatory approvals, changes in government policies/priorities, and lack of understanding of KT and HPSR methods. In response to identified needs, the TSC activities include: an in-person, kickoff workshop on KT and HPSR methods, a webinar series iteratively tailored to team needs, online discussion boards to support collaboration, implementation support (e.g. review of protocols, study materials) and coaching calls, and in-country workshops to support local capacity. We developed and registered on Open Science Framework a comprehensive four-phase, mixed methods study design protocol to determine the impact of the TSC program.

Expected Impact: To advance the knowledge on capacity building to support high-quality health system guidelines adaptation and implementation in LMICs.

Stakeholder Involvement: The TSC activities were designed using an integrated KT approach and are tailored to the needs of the RAISE teams.

Improving medical content accessed by millions: Wikipedia and the WikiJournal of Medicine

Dawson J[1], May M[2], Hutton MO[2], Dellavalle R[2]

[1] Children's Hospital of Eastern Ontario Research Institute and Cochrane, Canada; [2] University of Colorado School of Medicine, USA

Background: Wikipedia's health content receives over 10 million visits per day from around the world, across 286 languages. Efforts to improve the quality of the content and ensure that articles are not biased are ongoing and include student and trainee-based initiatives. The Cochrane Skin Wikipedia Project, which includes recruiting and training medical students and dermatology residents to improve Wikipedia with Cochrane Evidence, has improved over 73 skin-related Wikipedia articles as of 2020. Collectively, these articles have received over 27.8 million page views since 2018. Wikipedia activity, editing (via character counts and references added), and article page views can be meticulously tracked on Wikipedia. However, providing academics and students with formal recognition and credit for their contributions on Wikipedia is not common.

Objectives: Our goal was to work with the student group to improve a Wikipedia article and submit the article to The Wikijournal of Medicine, an open-access peer-reviewed journal that accepts Wikipedia article submissions.

Methods: In 2019, Cochrane Skin collaborated with Cochrane's Wikipedian in Residence to improve and polish the Leprosy Wikipedia article.

Results: Over a 12-month period, the students, Wikipedia volunteers, and international Leprosy experts reviewed and added over 79 new references and 18733 characters to the article. This article was submitted to the WikiJournal of Medicine in March 2020 and students are presently working through the peer-review process.

Conclusions: This new model of improving and polishing an existing Wikipedia article, submitting to an open-access journal, and working through the peer-review process is a valuable experience for students and helps ensure that medical information accessed via Wikipedia is accurate and up to date.

Patient or healthcare consumer involvement: Medical articles on Wikipedia are viewed millions of times a day by patients and healthcare consumers around the world. Ensuring that the evidence shared on Wikipedia is unbiased, up to date, and accurate is an important dissemination strategy to ensure people are accessing high-quality information pertaining to their health.

Improving Wikipedia skin disease content

Hutton MO[1], Dawson JE[2], Lee KC[3], Shumaker PR[4], Doney E[5], Scott HM[5], Dellavalle RP[1]

[1] Department of Dermatology, University of Colorado School of Medicine, Aurora, Colorado, USA; [2] Ottawa Hospital Research Institute, Ottawa, Ontario, Canada; [3] Department of Dermatology, Brown University, Providence, Rhode Island, USA; [4] Department of Dermatology, Naval Medical Center, San Diego, California, USA; [5] Cochrane Skin, Centre of Evidence Based Dermatology, University of Nottingham, Nottingham, UK

Background: Wikipedia is one of the most popular sources of health information for the public and appears among the top search engine results for dermatologic diseases. Because of its influence on health information, it is important to ensure that Wikipedia's content is evidence-based, unbiased, and up-to-date. The Cochrane-Wikipedia Partnership was founded in 2014 with the goal of ensuring that evidence-based health-related content is shared on Wikipedia. Cochrane Review Groups, Centres, and Fields engage with Wikipedia to recruit and train editors and share high-quality Cochrane Review evidence in Wikipedia articles.

Objectives: This project seeks to evaluate the impact of editing skin-related Wikipedia articles to include evidence-based information from Cochrane Skin reviews.

Methods: From May to August 2018, five medical students from U.S. medical schools were recruited to become Wikipedia editors. They were provided with dermatologist and Cochrane mentors and trained in editing technique. Over a six-month period, the trainees improved 40 skin-specific English language articles on Wikipedia. Articles were improved by adding paraphrased conclusions, background information, and references from 60 Cochrane Reviews.

Results: The dermatology-related Wikipedia articles that have been edited by the medical student team have amassed 28 million new views by March 2020. The top five viewed articles were on the topics of psoriasis, leprosy, cellulitis, melanoma, and alopecia areata. These five articles accounted for 38% of the total article view count.

Conclusions: The Cochrane Skin Wikipedia initiative aims to incorporate evidence-based information into Wikipedia health articles. Wikipedia's vast influence and accessibility makes it an effective dermatology education tool. We have shown that a small Wikipedia editing initiative has the potential to share evidence-based information with many people (i.e. 28 million Wikipedia article views in 19 months). The Cochrane Wikipedia initiative also provides an opportunity for medical professionals to actively contribute to evidence-based, informative articles. The Wikipedia editing team edit relevant pages when new Cochrane Skin reviews are published. Future directions of the initiative include recruiting more trainees, improving skin-related Wikipedia content in other languages, and making further improvements to increase article quality ratings. For more information about how to participate, please visit: https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Medicine/Cochrane/How. This work has been published in the Journal of the American Academy for Dermatology.

Patient or healthcare consumer involvement: Cochrane Skin reviews all have consumer input through authorship or peer review to make sure they are relevant to patients. The result of the project is Wikipedia articles which are more accurate, up-to-date and evidence-based, so that healthcare consumers and potential patients viewing them have the best health information possible.

Innovative ways to facilitate the understanding and communication of traditional Chinese medical terms and theories in a patient decision aid

Mu W[1], Fu S[2], He Y[3], Li Y[1], Wang B[4], Huang Y[1]

[1] Department of Clinical Pharmacology, Second Affiliated Hospital of Tianjin University of Traditional Chinese Medicine, China; [2] College of Chinese Medicine, Tianjin University of Traditional Chinese Medicine, China; [3] Pharmacy Department, Second Affiliated Hospital of Tianjin Unversity of Traditional Chinese Medicine, China; [4] Second Affiliated Hospital of Tianjin Unversity of Traditional Chinese Medicine, China

Background: In English-speaking countries, there are guidance documents on writing in plain language. For example, Cochrane Methods released the Standards for the Reporting of Plain Language Summaries in 2013, setting out rules and elaborating on the mandatory attributes of the language used in a summary of a Cochrane systematic review. As far as we know, there is no such guidance for writing medical texts in plain Mandarin, especially concerning traditional Chinese medicine (TCM).

Objectives: In developing a patient decision aid (PtDA), our research team looked at methods to best present information about TCM theories and therapies.

Methods: As there is no rule to follow, we organized a panel meeting for brainstorming after the PtDA was drafted. The PtDA developers, a university lecturer in TCM classics, and a pharmacist-in-chief of Chinese materia medica joined the discussion. Innovative ways were proposed and discussed for feasibility until consensus was reached.

Results: In our PtDA designed for stable angina patients, we provided information about four medication therapies, including two Chinese patent medicines. To better communicate these herbal components and their health effects to the patients, we broke down the whole formula into clusters of couplet medicinals or single herbs and presented them in order of their functionality and quantity. The health effects of a group of couplet medicinals or a single herb were first explained in the classical TCM language; for example, "Huangqi, a herbal medicine, boosts qi and frees the vessels". Their pharmacological effects corresponding to or supporting the health effects were then provided in the modern biomedical language in a different colour; for example, "Huangqi protects heart muscle and relaxes blood vessels". The explanations in classical TCM terms were based on a TCM textbook. The corresponding or supporting pharmacological effects were based on findings of pharmacological research.

Conclusions: This method creates a visually structured text to make this part of the PtDA more comprehensible. This helps healthcare consumers get instant access to the information they need in order to distinguish between different options. The PtDA can be found at https://decisionaid.ohri.ca/AZsumm.php?ID = 1930.

Patient or healthcare consumer involvement: Two patient representatives were invited to review the finalized PtDA and give comments on readability.

Is the minimal important difference (MID) considered in systematic reviews?

Ciapponi A[1], Glujovsky D[1], Comandé D[1], Ciapponi M[1]

[1] Instituto de Efectividad Clínica y Sanitaria (IECS-CONICET), Argentina

Background: The minimal important difference (MID) is the smallest change in a treatment outcome that is considered important for patients but also for the health system. As Cochrane requires GRADE to classify the quality of the evidence, MID is needed to define the imprecision and the magnitude of the effect. Review authors should use their judgment in deciding what constitutes appreciable benefit and harm and provide a rationale for their choice. If review authors fail to find a compelling rationale for a threshold, the GRADE guideline suggests a default threshold (25 ± 5% in relative terms) for appreciable benefit and harm to rate imprecision.

Objectives: To analyse how Cochrane Systematic Reviews (SRs) and Cochrane protocols compare to non-Cochrane SRs regarding the reporting and interpretation of results, if appropriate, according to MID.

Methods: We analysed all Cochrane SRs and protocols published in November to December 2018. We excluded updates. We also analysed the top five impact factors journal list (1 New England Journal of Medicine, 2 Lancet, 3 BMJ, 4 JAMA, 5 Annals of Internal Medicine), based on the 2017 Thomson Reuters Journal Citation in medicine general and internal category, where Cochrane ranked 13th. The PubMed search strategy to identify non-Cochrane SRs was: "The New England journal of medicine"[Jour] OR "Lancet (London, England)"[Jour] OR "British medical journal"[Jour] OR "JAMA"[Jour] OR "Annals of internal medicine"[Jour]. Filters: Meta-Analysis; Systematic Reviews; Publication date from 1 January 2018 to 31 December 2018. We explored how Cochrane SRs and protocols, and non-Cochrane SRs dealt with the MID concept as defined above, in every relevant review/protocol section.

Results: We have analysed eight Cochrane SRs, eight Cochrane protocols and nine non-Cochrane SRs so far. We will give the rest of the references at the Colloquium. No Cochrane SRs, 38% of Cochrane protocols and 56% of non-Cochrane SRs reported the MID. All the analysed protocols were from the same author, who assessed a dichotomous outcome for different drugs in the same condition. Sixty-six per cent of the non-Cochrane SRs evaluated continuous primary outcomes and 33% the risk difference, and in all cases the statements were supported by references (see the statements used in Box 1). Only the two SRs that evaluated continuous outcomes formally incorporated patients' point of view. Only three of 14 (21%) of the SRs not reporting a clear MID statement mentioned some effect size considerations (see Box 2) but it was not possible to assess reporting consistency in these cases.

Box 1. Examples of correct statement of MID in the methods section

- For example, the minimally important difference is about 1.0 cm for the 10-cm VAS for pain (Dworkin 2008). For the SF-36 items, the minimally important difference of 10 points was used for the individual domains (Ward 2014)
- Minimal clinically important differences (MCID), patient derived scores that represent changes in a score that have meaning for patients, have been suggested for the ADAS-cog 3 points in severe AD (Howard 2011) and MMSE (1.4 points in mild AD (Schrag 2012).
- A minimum absolute risk difference of 4% in the primary composite outcome of death or major disability by a corrected age of 18 to 24 months, corresponding to a minimally important number needed to treat of 25 infants to prevent 1 major adverse outcome (Askie 2011).
- Clinically important change (as defined by individual studies) (e.g. 'much improved' or more than 50% reduction/increase on a relevant rating scale) (Same statement for the 3 CPs)

Box 2. Examples of use of unexplained MID for result interpretation

- ...none reached clinical or statistical significance in VAS scores.
- Compared with failure or relapse, treatment success was positively associated with the use of.... and provided modest benefits.
- Trial sequential analysis confirmed that the required information size was met

Conclusions: When MID was considered in the methods section, the reporting of the results was consistent with the definition used. In most of the cases, MID was not reported and authors did not state how they defined it to analyse imprecision and magnitude of effect.

Patient or healthcare consumer involvement: None, since it was a methodological study about published systematic reviews.

Knowing what works: the Cochrane blog "Wissen Was Wirkt" provides easy and understandable access to Cochrane Evidence in German

Borchard A[1], Braun C[2], Meerpohl JJ[2], Nußbaumer-Streit B[3], Puhl A[2], Rüschemeyer G[2], von Elm E[1]

[1] Cochrane Switzerland, Unisanté, University of Lausanne, Lausanne, Switzerland; [2] Cochrane Germany Foundation, Freiburg, Germany; [3] Cochrane Austria, Danube University Krems, Krems, Austria

Background: Doctors and patients alike are overwhelmed with the sheer amount of complex scientific health information. They need free and easy access to understandable and unbiased information. In line with Cochrane's Strategy to 2020 aim to make evidence more widely accessible, the Cochrane Centres in Germany, Switzerland and Austria co-operate in providing and disseminating an accessible format of Cochrane evidence via "Wissen Was Wirkt" (Knowing What Works), the German-language Cochrane blog that started in May 2015.

Objectives: To disseminate Cochrane content via a joint blog in three German-speaking countries and to ensure its continuous development.

Methods: We regularly provide blog posts that summarise the results of Cochrane Reviews in an easily understandable and approachable manner. We also contextualise methods of evidence-based medicine using interesting and personal stories. Articles are written by the staff of the three German-speaking Cochrane Centers and by selected external authors. In 2019, a detailed author's guide was established to support the drafting of articles and to facilitate quality assurance. We disseminate our blog posts via the website wissenwaswirkt.org and our social media channels. We use Google Analytics to evaluate user access statistics.

Results: We have so far published 227 articles on "Wissen Was Wirkt", on average almost one article per week. In 2019, more than 160 000 users visited the blog (about 13 300 per month); web pages were accessed more than 578 000 times in total. In 2018, we carried out a user survey. Its results were used for the further development of the blog. We will provide updated user statistics at the time of the presentation.

Conclusions: The blog is an excellent medium for presenting evidence-based information about healthcare, mostly from Cochrane Reviews, in a way that is accessible to lay persons and professionals alike. Although it is time consuming to produce high quality content for the blog and co-ordinate activities between the three Cochrane Centers, "Wissen Was Wirkt" is a good example of how such a co-operation can succeed and how it can be continuously evaluated and improved. So far, we have published mostly text formats. We are now also exploring video and graphic formats to respond to different user preferences.

Patient or healthcare consumer involvement: The target audience for the blog are consumers. Most articles have been written by Cochrane staff, some by Cochrane Consumers. Other articles feature consumer stories. In 2020, we have – for the first time – asked users to comment on an article before its publication to ensure its readability and relevance.

Knowledge synthesis projects and knowledge translation considerations: patterns and trends in a specialized research program

Sebastianski M[1], Bialy L[1], Thomson D[1], Hartling L[1]

[1] Alberta Strategy for Patient-Oriented Research (SPOR) SUPPORT Unit Knowledge Translation Platform, Canada

Background: Knowledge synthesis (KS) is the cornerstone for evidence-based healthcare decision-making and optimizing patient outcomes, and a key component of the Knowledge to Action Framework. Yet, KS activities will only have impact if accompanied by a strong knowledge translation (KT) plan including knowledge users and decision-makers. Supporting KS is an integral part of the Alberta Strategy for Patient-Oriented Research (SPOR) SUPPORT Unit KT Platform. Since April 2016 the KT Platform has provided consultation and support services on a range of KS projects.

Objectives: 1) To understand the supports and services required by our clients, maximize resource utilization, and build on lessons learned; and 2) to identify opportunities and mechanisms to integrate KT into the KS process to optimize research uptake and impact.

Methods: From April 2016 to February 2020 the KT Platform supported 142 projects. For each project, we collected data from applications for services and consultation notes with a follow-up survey sent to the principal investigator of each completed project (n = 83). We designed the survey to gather information on KT considerations, knowledge user involvement, and resulting dissemination products. We analysed quantitative data using descriptive statistics and performed thematic analysis of qualitative data.

Results: Among all 142 projects, the majority focused on clinical topics (68%) with systematic reviews the most common methodology (55%), followed by rapid reviews (24%) and scoping reviews (13%). Survey response rate was 76% (63/83 completed projects). One third of respondents (33%) indicated a KT plan. Various knowledge users were involved, most commonly healthcare practitioners (41%), individual patients and/or patient organizations (34%) and policy makers (27%). Knowledge users were involved in: interpreting findings (82%), messaging and disseminating results (45%), shaping research questions (33%), deciding on methodology (31%), and helping with data collection and tool development (11%). Review results have been used for a variety of purposes, e.g. decision aids, clinical practice guidelines, grant applications, and subsequent projects. To date, 51 reviews (61%) have yielded dissemination products including 43 presentations and 70 publications.

Conclusions: The KT Platform has supported a large number of KS projects with the potential to impact patient outcomes; however, only a third had a KT plan. Among those with a KT plan a variety of knowledge users were involved and a range of products have emerged intended to bridge the research-practice gap. Nevertheless, our data demonstrate that more work is required to enhance linkages between KS research and KT activities to ensure optimal impact on patient outcomes.

Patient or healthcare consumer involvement: Patients and patient organizations were involved in some of the supported projects, however no patients or consumers were involved in this quality improvement project.

Knowledge translation among students: the ExME initiative

Viteri-García A[1], Angeletti E[2], Ardiles S[3], Fu Carrasco R[4], Casino G[5], Gorgues M[6], Meza N[7], Montesinos-Guevara C[1], Perales L[8], Pizarro AB[9], Rondon Rueda L[4], Rosas JC[10], Verdugo-Paiva MF[11], Bonfill Cosp X[6]

[1] Universidad UTE, Centro Asociado Cochrane de Ecuador, Centro de Investigación en Salud Pública y Epidemiología Clínica (CISPEC), Ecuador; [2] Universidad Nacional de la Matanza, Argentina; [3] Universidad Autónoma de Barcelona, Facultad de Medicina, Spain; [4] Universidad de Ciencias Médicas (UCIMED), Costa Rica; [5] Cochrane Iberoamérica, Spain; [6] Cochrane Iberoamérica, Servicio de Epidemiología del Hospital de la Santa Creu i Sant Pau, Spain; [7] Interdisciplinary Centre for Health Studies (CIESAL), Cochrane Associate Centre, School of Medicine, Universidad de Valparaíso, Chile; [8] Universidad de Guadalajara, Mexico; [9] Pontificia Universidad Javeriana, Facultad de Enfermería, Colombia; [10] Pontificia Universidad Javeriana, Facultad de Medicina, Colombia; [11] Fundación Epistemonikos; Centro Evidencia UC; Pontificia Universidad Católica de Chile, Chile

Background: Knowledge translation (KT) in health sciences has been defined as "a dynamic and iterative process that includes synthesis, dissemination, exchange and ethically sound application of knowledge to improve the health, provide more effective health services and products and strengthen the health care system". It does not only consist of the dissemination of information but rather in the real use of the knowledge produced by clinical research, clinical experience and patients' preferences. Many initiatives have been undertaken to address the need of improving KT; one of them included the participation of health sciences students, Students for Best Evidence (S4BE). S4BE is a platform made from students for students about evidence in health sciences; it reached great success among English speaking countries. We have identified that this type of resource is necessary to be available on a multilingual platform due to language barriers that are common in many non-English speaking countries. Therefore, Estudiantes por la Mejor Evidencia (ExME) is a Spanish initiative within the S4BE Cochrane Project.

Objectives: To produce, translate, and disseminate content related to the best evidence in health sciences. The platform has been hosted on a web page, in blog format, to help students make better clinical and health decisions.

Methods: ExME is an initiative that gathers students and stakeholders from seven different Spanish-speaking countries. We have developed a system of contributions for students of all levels. Undergrad and grad students write the blog entries from their own queries and submit them, then postgraduate, masters and doctoral students check the manuscript and give comments and advise to the authors. Subsequently, an editorial process in charge of the co-ordinating committee takes place. Finally, the blog entry is published on the website of the initiative. Translations follow a similar path. After being published, every blog entry is shared on the social media accounts to be spread throughout the internet.

Results: We identified several students that are interested in producing, translating and disseminate entries for the ExME initiative. They prepared the first blogs that are ready to be published. Also, the co-ordinating committee developed prioritized topics according to the interests of the students currently involved and new blogs are being prepared.

Conclusions: Knowledge translation initiatives like ExME are desperately needed in health sciences. The involvement of students in the early stages of their carriers is crucial to improve it. We expect that more students continue to join our initiative.

Patient or healthcare consumer involvement: Not applicable at the moment, but it is in our best interest to open a space in the platform for the perspectives of patients.

Knowledge translation and dissemination of Cochrane evidence on pregnancy, childbirth, maternity and breastfeeding into Portuguese language via YouTube

Pedrazzi A[1], Torloni MR[2], Pacheco R[3], Latorraca C[4], Fontes LE[5], Martimbianco AL[6], Riera R[7]

[1] Hospital Sírio-Libanês, Brazil; [2] Universidade Federal de São Paulo (Unifesp); Cochrane Brazil, Brazil; [3] Cochrane Brazil Rio de Janeiro; Centro Universitário São Camilo; Universidade Federal de São Paulo (Unifesp), Brazil; [4] Cochrane Brazil Rio de Janeiro; Universidade Federal de São Paulo (Unifesp), Brazil; [5] Cochrane Brazil Rio de Janeiro; Faculdade de Medicina de Petrópolis (FASE), Brazil; [6] Cochrane Brazil Rio de Janeiro; Universidade Metropolitana de Santos (Unimes); Centro Universitário São Camilo, Brazil; [7] Universidade Federal de São Paulo (Unifesp); Cochrane Brazil Rio de Janeiro; Hospital Sírio-Libanês, Brazil

Background: Knowledge translation is characterized by a dynamic and iterative process that encompasses the synthesis, dissemination, exchange and ethical application of knowledge to improve health, provide more effective and safe healthcare technologies and qualify the health systems [1]. Knowledge translation encompasses the second goal (Making our evidence accessible: accessible language / multi-lingual) of the Cochrane Strategy to 2020 [2].

Objectives: To describe an initiative to promote the knowledge translation and dissemination of Cochrane evidence on pregnancy, childbirth, maternity and breastfeeding into Portuguese language.

Methods: Descriptive case study.

Results: This initiative was planned and developed in the Post-graduation Program of Evidence-based Health, Universidade Federal de São Paulo with the support of the Cochrane Brazil Rio de Janeiro. The initiative comprises a set of activities as follows. (1) To produce 20 videos in Portuguese addressing 20 themes on pregnancy, childbirth, maternity and breastfeeding. The videos will be based on the findings of Cochrane Reviews, and on the blogs 'Cochrane Maternity Matters' from Evidently Cochrane (www.evidentlycochrane.net/tag/maternity-matters) and 'New baby series: fads, fashions and evidence for new parents' from Cochrane UK website (uk.cochrane.org/news/new-baby-fads-fashions-and-evidence-new-parents). The videos will be available in the 'Mãe-Estar', a video blog on YouTube, with additional links for Facebook and Instagram. (2) To quantitatively evaluate the videos access, through statistical analyses of the number of visualizations, likes, shares and comments. These analyses will measure the extent to which the initiative has reached the population, making it possible to identify the most viewed videos and the origin / location of the users. (3) To estimate the knowledge retained by users who watched the videos, through a survey made available before the start and after the end of the video.

A pilot video, addressing the vitamin D supplementation before pregnancy, has already been produced and is available from: www.youtube.com/watch?v=ED89INgFj3M

Conclusion: We hope this initiative of knowledge translation could contribute for disseminating Cochrane evidences involving pregnancy, childbirth, maternity and breastfeeding around the Portuguese-speaking countries.

Patient or healthcare consumer involvement: Disseminating evidence of Cochrane Reviews in the Portuguese language is a way of contributing so that Portuguese-speaking world population can make health choices informed by evidence.

References:

1. Straus SE, Tetroe J, Graham I. Defining knowledge translation. *Canadian Medical Association Journal* 2009;181(3-4):165-8. doi: 10.1503/cmaj.081229
2. Cochrane. Strategy to 2020. www.cochrane.org/sites/default/files/public/uploads/pdf/strategy_to_2020_22june2016.pdf

Knowledge translation to patients: the project of patient-oriented Cochrane Rehabilitation Cochrane Corners

Oral A[1], Lazzarini SG[2], Grubisic F[3], Kiekens C[4], Negrini S[5], Arienti C[2]

[1] Department of Physical Medicine and Rehabilitation, Istanbul Faculty of Medicine, Istanbul University, Istanbul, Turkey; [2] IRCCS Fondazione Don Carlo Gnocchi, Milan, Italy; [3] Department of Rheumatology, Physical Medicine and Rehabilitation, University Hospital Center "Sestre Milosrdnice", Zagreb, Croatia; [4] Spinal Unit, Montecatone Rehabilitation Institute, Imola, Italy; Department of Physical and Rehabilitation Medicine, University Hospitals Leuven, KU Leuven, Leuven, Italy; Belgium; [5] Department of Biomedical, Surgical and Dental Sciences, University of Milan "La Statale", Milan; IRCCS Istituto Ortopedico Galeazzi, Milan, Italy

Background: Shared decision-making, respecting patients' preferences and values are important concepts in the field of evidence-based medicine. 'Knowledge translation' should not only address health professionals but also all other stakeholders including the end-users, funders, and policy makers. While the evidence is particularly used by health professionals to deliver the best care to the patients, the end-users/beneficiaries are the patients. Cochrane Corners could serve as a good knowledge translation tool to convey Cochrane evidence in a simpler way to stakeholders. However, there are gaps in the application of knowledge from the perspective of the beneficiaries.

Objectives: To involve patients, their family members, and carers in the dissemination of Cochrane evidence and to provide them with health information to be used in shared decision-making together with their healthcare providers.

Methods: The dissemination of Cochrane evidence to different stakeholders may differ. The structure of patient-oriented Cochrane Corners needs to involve patient-identified priorities, their needs, and their perspectives on patient-centred outcomes. They also need to better describe the benefits and harms of certain treatments.

Results: Cochrane Rehabilitation produced 34 Cochrane Corners (as of the end of 2019) summarizing published Cochrane Systematic Reviews from a rehabilitation perspective to inform rehabilitation professionals of the Cochrane evidence to be used in rehabilitation practice for the best rehabilitative care of patients. While continuing with efforts to disseminate Cochrane evidence to health professionals to ensure the use of the best available evidence in rehabilitation practice, Cochrane Rehabilitation has planned the initiative of 'knowledge translation to patients' by producing patient-oriented Cochrane Corners to improve patients' knowledge of treatment options, to make them aware of the benefits and harms of certain treatments in order to enable their proper participation in decision making as well as in goal-setting.

Conclusion: Knowledge translation to patients using Cochrane Corners will be helpful for use in shared decision-making to ensure better health decisions with the involvement of patients, their family members, and/or carers.

Patient or healthcare consumer involvement: Patients and healthcare consumers are involved during the writing of Cochrane Corners and overall, in the interpretation of results from patients' perspectives.

Learning from the audience: a qualitative study on perception of systematic review summary formats

Prokop-Dorner A[1], Zając J[2], Bała M[3]

[1] Department of Medical Sociology Jagiellonian University Medical College, Poland; [2] Department of Hygiene and Dietetics Jagiellonian University Medical College, Jagiellonian University Medical College; [3] Chair of Epidemiology and Preventive Medicine Jagiellonian University Medical College, Poland

Background: In the context of increasing complexity of the medical discourses and the enormous amount of information available, providing medical professionals and lay people with 'accessible, credible information to support informed decision-making' (www.cochrane.org/about-us) is an adequate mission for research institutions. In order to support medical professionals in their daily work and to ensure development of medical awareness among a lay audience, the reception of the information disseminated by researchers should be evaluated and readers preferences should be taken into account.

Objectives: To gain an understanding of recipients' perception of the available formats of presenting findings of systematic reviews.

Methods: We conducted five Focus Group Interviews (FGI) with university employees (n = 6), public health students (n = 8), pharmacists (n = 7), patients and caregivers (n = 6) as well as physicians and nurses (n = 6). We presented nine information formats to the study participants: Plain Language Summary (PLS), audio-recorded PLS, 'Summary of findings' table (SoF), vlogshot, blogshot, infographics, press-release (PR), comic drawing, and abstract. During a moderated discussion, participants were encouraged to share their individual opinions about perceived usefulness of the formats, their strong sides and flaws. We then transcribed the video-recorded interviews and inductively coded them. In order to identify patterns of preferences, we used the technique of constant comparison and data visualization.

Results: The gathered material provided an insight into people's variety of preferences. For the study participants important characteristics of the presented information were: its trustworthiness; an applicatory character; comprehensibility; information structure, graphical means used and clarity as well as general effect. We categorized opinions about the presented forms of popularizing results from SR into three groups: positive, negative and ambiguous. Positive comments concerned the structured formats, presenting practical benefits for the reader, using graphic representations, comprehensibility of the information, precision of the information and clearness of conclusions. Negative comments related to the lack of information about possibility to apply the conclusions as well as the incomprehensibility of the applied terminology on one hand, and the low precision or lack of intervention effect on the other. When talking about some formats, ambiguous opinions weighed pluses and minuses.

Conclusions: The data gathered suggest that each format should be carefully revised and accompanied with clear guidelines. Tailoring the presentation of SR findings to the needs of the various categories of targeted recipients may result in a greater efficiency of the efforts to disseminate them.

Patient or healthcare consumer involvement: Our research efforts aimed at understanding the needs of various categories of recipients of Cochrane SR summaries, including the lay audience, patients and their families.

Learning health systems and evidence ecosystems: a great fit?

Vandvik P[1], Brandt L[1], Agoritsas T[2], Bernstein S[3]

[1] MAGIC Evidence Ecosystem Foundation, Norway; [2] University Hospitals Geneva, Switzerland; [3] AHRQ, USA

Background: The continued challenges of making Evidence-Based Medicine (EBM) work in healthcare policy and practice have triggered a series of evidence ecosystem concepts, one of them being the digital and trustworthy evidence ecosystem (Evidence Ecosystem) spearheaded by MAGIC in partnership with Cochrane and other international partners. This is happening while other communities aiming to implement EBM in healthcare advocate for Learning Health Systems (LHS). To what extent these concepts align and how they optimally could interact warrants further exploration, for example to inform Cochrane efforts in evidence synthesis.

Objectives: To describe the Evidence Ecosystem and LHS concepts in terms of what problems these are trying to solve and proposed solutions.

Methods: The Evidence Ecosystem represents a conceptual framework for a cyclical and continuous process, with a focus on more efficient evidence synthesis and trustworthy decision support through common standards, methods, processes and platforms (Figure 1). Recently a major LHS initiative has been launched in the US; The AHRQ evidence-base Care Transformation Support (ACTS, figure 2). The ACTS community includes over 140 organizations working together to reach the quadruple aims of US health care. The main focus is on digital decision support, delivered care and continuous quality improvement in practice. MAGIC is conducting pilots with ACTS on digital production and access to evidence and decision support through interoperable platforms (e.g. MAGICapp). This provides an opportunity to study concerted evidence synthesis and guidance at the international level (e.g. BMJ Rapid Recommendations) can inform LHS at national and local levels. Here LHS is set up to cover the final Evidence Ecosystem steps of downstream implementation, evaluation of impact and production of more relevant and reliable evidence.

Results: We find the two concepts completely aligned and complementary in describing problems and proposed solutions concerning evidence from its inception to documented improvements in delivered health care. Both concepts underscore the need for an overarching infrastructure to provide orchestration, governance and support to organizations currently working in silos. The Evidence Ecosystem can provide trustworthy and digital decision support informed by systematic reviews, ideally at the international level to increase efficiency and reduce duplication of work. LHS can assist healthcare organizations working at national and local levels in optimally re-using and adapting the decision support to improve delivered care.

Conclusions: The synergy between Evidence Ecosystem and concepts are striking but it remains to be seen how these could optimally interact. In the absence of an international orchestrator in the Evidence Ecosystem and explicit links to LHS we will likely continue to work in silos.

Patient or healthcare consumer involvement: Patient partners are contributing in the development of the Evidence Ecosystem and LHS concepts.

Linguistic analysis of Cochrane systematic reviews of oncology interventions Plain Language Summaries: cross-sectional study

Šuto J[1], Buljan I[2], Marušić A[2]

[1] Medical student, University of Split School of Medicine, Split, Croatia; [2] Department of Research in Biomedicine in Health and Cochrane Croatia, University of Split School of Medicine, Split, Croatia

Background: People have difficulties in understanding health information. Thus, Plain Language Summaries of Cochrane systematic reviews are meant to be written in the way everyone could understand and make responsible decisions, presenting a bridge to overcome the gap between the healthcare users and professionals.

Objectives: To assess the language characteristics of Plain Language Summaries (PLSs) of systematic reviews of oncology interventions in comparison to the language of corresponding scientific abstracts.

Methods: In this cross-sectional study we included all Cochrane scientific abstracts and corresponding PLSs of systematic reviews of oncology interventions available in the Cochrane Library up to February 2019, including Breast Cancer Group, Childhood Cancer Group, Gynecological, Neuro-oncology and Orphan Cancers Group, Hematology Group, and Lung Cancer Group. Language characteristics of PLS included text readability, measured using the Simple Measure of Gobbledygook (SMOG) index, and prevalence of words related different language tones (clout, authenticity, emotions and analytical) measured using LIWC software and conclusiveness of the PLS.

Results: We collected 275 PLSs and corresponding scientific abstracts of systematic reviews of oncology interventions. In general, SMOG index of PLSs was slightly above the recommended 12 years of education for health information materials, and the readability did not differ across Cochrane Review Groups. In general, the PLSs from the Colorectal Cancer Group were the shortest, whereas the PLSs from the Hematological Malignancies Group had the lowest proportion of words reflecting emotional tone. The Lung Cancer Group had the highest proportion of PLSs with negative conclusions compared to the other groups. The Childhood Cancer Group did not have any summaries with positive conclusions. Gynecological, Neuro-oncology and Orphan Cancer Group and Breast Cancer Group had no PLSs where the authors provided an opinion. Overall, PLSs with "no evidence" conclusion had the lowest SMOG index, as well as the fewest number of words compared to other conclusiveness categories. PLSs with positive and equal conclusiveness had the lowest proportion of words with analytical tone, while the PLSs with "no evidence" for any definite conclusion had a greater proportion of words related to clout tone. The comparison of language in PLSs and scientific abstracts is underway and will be presented at the Colloquium.

Conclusions: PLSs of Cochrane systematic reviews of oncological interventions have low readability and low emotional tone, as well as conclusiveness of the review findings. We intend to further compare PLSs and scientific abstracts, to understand the differences in the language of scientific and popular health information texts in order to suggest possible ways of improving the usability of information from Cochrane systematic reviews.

Patient or healthcare consumer involvement: Based on the results we will have information which can be used as the recommendations for improving the writing in PLSs.

Making child health evidence usable to the public: what do parents want?

Elliott S[1], Anzinger H[2], Hartling L[3]

[1] Cochrane Child Health, University of Alberta, Canada; [2] Department of Pediatrics, University of Alberta, Canada; [3] Alberta Research Centre for Health Evidence, Faculty of Medicine and Density, University of Alberta; Cochrane Child Health, University of Alberta, Canada

Background: Connecting parents to research evidence is known to improve health decision making. However, guidance on how to develop effective knowledge translation (KT) tools, which synthesize child health evidence into a form understandable by parents is lacking.

Objectives: To conduct a comparative usability analysis of three online KT tools to identify differences in tool effectiveness, identify which format parents prefer, and to better understand what factors affect usability for parents.

Methods: We evaluated a Cochrane Plain Language Summary (PLS), blogshot, and a Wikipedia page, on a specific child health topic (acute otitis media). We used a mixed-methods approach, involving a knowledge test, written usability questionnaire, and a semi-structured interview. We analyzed differences in knowledge and usability questionnaire scores for each of the KT tools using Kruskal-Wallis tests, considering a critical significance value of P = 0.05. We used thematic analysis to synthesize and identify common parent preferences among the semi-structured interviews. We derived key elements that parents wanted in a KT tool through author consensus using questionnaire data and parent interviews.

Results: Sixteen parents (9 female) aged 39.6 ± 11.9 years, completed the study. Parents preferred the blogshot over the PLS and Wikipedia page (P = 0.002) and found the blogshot to be the most aesthetic (P = 0.001), and easiest to use (P = 0.001). Knowledge questions and usability survey data also indicated the blogshot was the most preferred and effective KT tool at relaying information about the topic. Four key themes derived from thematic analysis, describing elements parents valued in KT tools. Parents wanted tools that were 1) simple, 2) quick to access and use, 3) trustworthy, and 4) informed how to manage the condition. Out of the three KT tools assessed, blogshots were the most preferred by parents, and encompassed these four key elements.

Conclusions: It is important that child health evidence be available in formats accessible and understandable by parents to improve decision making, use of healthcare resources, and health outcomes. Further usability testing of different KT tools should be conducted involving broader populations and other conditions (e.g. acute versus chronic) in order to generate guidelines to improve KT tools for parents.

Patient or healthcare consumer involvement: Here we disseminated three online KT tools for evaluation by parents (our key knowledge user/consumer). We then engaged with these parents on how best to design and develop KT tools which meet their unique needs. We initially engaged directly with our parent advisory group for this work, and then from their suggestions and networks invited other parents to participate.

Making the evidence accessible: improvement of Wikipedia entries by incorporating Cochrane Review results in haematology

Jakob T[1], Piechotta V[1], Tomlinson E[2], Kleinstäuber S[1], Estcourt L[3], Skoetz N[4]

[1] Cochrane Haematology, Department I of Internal Medicine, Center for Integrated Oncology Aachen, Bonn, Cologne, Dusseldorf, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany; [2] Cochrane Cancer Network, Royal United Hospital, Bath, UK; [3] Cochrane Haematology, Haematology and Transfusion Medicine, NHS Blood and Transplant, and Radcliffe Department of Medicine, University of Oxford, Oxford, UK; [4] Cochrane Cancer, Department I of Internal Medicine, Center for Integrated Oncology Aachen, Bonn, Cologne, Dusseldorf, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany

Background: The methodology of Cochrane Reviews is considered the gold standard for the production of high-quality evidence. Together with the strict conflict of interest policy and patient focus, the evidence produced is reliable and relevant to healthcare professionals, policy makers and consumers. Especially for the latter, finding trustworthy and easily accessible information on the internet can be challenging. Wikipedia is the largest and most popular multilingual online encyclopedia created, maintained as an open collaboration project by a community of volunteer editors. Until now, only 16 of 80 Cochrane Reviews in the field of haematology are cited on Wikipedia.

Objectives: To disseminate Cochrane evidence to a broader audience by improving relevant Wikipedia entries with up-to-date and patient-relevant results of Cochrane Haematology reviews in plain language.

Methods: On the basis of the PICO-scheme we systematically identified all populations and interventions that are covered within up-to-date reviews of Cochrane Haematology. In a second step, we mapped outcomes reported within the 'Summary of findings' tables of these reviews to the populations and interventions. By focusing on the results of the 'Summary of findings' tables of the reviews we made sure we included results for the prioritized main outcomes. We then searched for existing Wikipedia entries in German and English language covering topics around the identified target populations and interventions, checked their contents and assessed whether the provided information included complete and accurate evidence. Where we identified evidence gaps, we adapted Wikipedia articles in a systematic and meaningful way by using a standardized wording considering the effect estimates and the GRADE assessment of the 'Summary of findings' table. Generally, the wording was based on the Plain Language Summaries of the reviews. A Masters student is realizing the project as part of her Masters thesis.

Results: We identified 26 target populations and 11 interventions of our reviews for which Wikipedia entries exist. The Masters student was trained in interpreting 'Summary of findings' tables and plain language wording. We created accounts for the Wikipedia internal Cochrane Dashboard to monitor the progress so that by mid-2020 the results from up-to-date reviews get incorporated in to Wikipedia in a meaningful way.

Conclusion: Editing Wikipedia entries with high-quality Cochrane Review results is an easy way to disseminate evidence to a broad and lay audience. In times of 'fake news' this might help to incorporate trustworthy information to support informed decision making and making sure that high-quality health research is purposefully used.

Patient or healthcare consumer involvement: By making sure we cite the review results in plain language we make evidence more accessible to patients and healthcare consumers. Plain Language Summaries of the reviews, which are commented on and checked by consumers before publication, will form the basis of wording.

Meta-analyses for regulatory purposes: the case of esketamine

Gastaldon C[1], Papola D[1], Ostuzzi G[1], Barbui C[1]

[1] WHO Collaborating Centre for Research and Training in Mental Health and Service Evaluation Department of Neuroscience, Biomedicine and Movement Sciences Section of Psychiatry, University of Verona, Italy

Background: US FDA and the European Medicines Agency (EMA) are responsible for the evaluation and the approval of medicines developed by drug companies. These evaluations are based on the results of marketer studies and submitted to the medicine regulatory agencies. Currently, efficacy proof relies on clinical studies considered individually, whereas meta-analytical evidence is not required. A recent example of the approval process concerns a nasal spray of esketamine (S-ket) for the treatment of treatment-resistant depression in 2019 by the FDA and EMA.

Objectives: Using the example of S-ket, to underline the importance of meta-analyses in regulatory processes, aiming to draw implications for clinical practice, research and regulatory science.

Methods: We performed a systematic review and meta-analysis of trials on S-ket submitted to the FDA and EMA.

Results: We found four Phase III trials: three short-term placebo-controlled (PLB) efficacy trials and one withdrawal trial (Table 1). Only one short-term efficacy trial found a difference between S-ket+antidepressant (AD) and PBL+AD. Re-analysis of the efficacy data of the three short-term studies revealed an overall mean difference of -4.08 in MADRS-score (95% confidence interval (CI) -6.20 to 1.97, Figure 1), suggesting that S-ket may improve depressive symptoms compared to PLB. Pooled data on acceptability showed that S-ket was significantly less acceptable than PLB (Figure 2). Moreover, S-ket increased by seven times the risk of dissociation over PLB, with approximately 25% of S-ket treated patients experiencing severe dissociation during treatment (Figure 3). The withdrawal trial showed that participants who discontinued S-ket after improvement with S-ket+AD were more likely to relapse compared to patients that did not discontinue. However, this design carries several limitations; generalizing results from this type of design to patients with a current depressive episode may be challenging.

Conclusions: Re-analysis of the clinical data on S-ket submitted to the FDA and EMA on acceptability and safety outcomes showed a significantly worst profile of S-ket compared to PLB. The statistically significant superiority of S-ket to PLB regarding efficacy can hardly be considered clinically significant. Further, it does not overweight the risks related to the acceptability and tolerability profile of S-ket. This information can only be generated using secondary, i.e. meta-analytical data, the exclusion of which is a real disservice for the approval process. The case of S-ket offered a unique opportunity to reflect on the evidence supporting the licensing of new agents and, generally, to critically appraise the approval process for new drugs. We propose that the evaluation process of medicines should be complemented by regulatory meta-analyses of all relevant clinical studies. Based on meta-analytical evidence, agencies could develop a more systematic and transparent approach to summarise the submitted evidence and its quality, using, for example, the GRADE tool.

Additional files: Tables and figures

Plain Language Summaries: dissemination in the Polish context

Zajac J[1], Koperny M[2], Kolcz M[2], Mocko K[2], Bala MM[1]

[1] Cochrane Poland, Systematic Reviews Unit, Chair of Epidemiology and Preventive Medicine, Jagiellonian University Medical College, Poland; [2] Cochrane Poland, Systematic Review Unit, Jagiellonian University Medical College, Poland

Background: One of the main goals of Cochrane is making high quality medical information accessible for the lay audience. This aim in Poland is mainly achieved by dissemination of Plain Language Summaries (PLS) using fan pages on Twitter and Facebook.

Objective: To describe our experiences with dissemination of Cochrane evidence using social media and to explore the changes in the profile of followers, fans and topic preferences between the 2016 and 2020.

Methods: Cochrane Poland was established in 2015, as a Branch of the Nordic Cochrane Centre. We started the Polish Facebook fan page in March 2016, and a Twitter profile nine months later. PLS are translated by volunteers (including students taking part in the educational project "Humanities for Health" at the University of Warsaw) and their quality is ensured by medical editors also from the Jagiellonian University.

Results: Up to March 2020 Polish translations were available for 1033 PLS and 57% were promoted by fan pages. Initially the translations were done only by volunteers, also students and graduates of medicine from Jagiellonian University, Medical College. Subsequently within the project "Humanities for Health" we started co-operation with students at the Institute for English Studies, University of Warsaw. At the moment we also accept volunteers who sign up via join.cochrane.org website. Since 2015 we co-operated with 161 volunteers and editors, some of them co-operate with us till now as permanent translating volunteers and editors. Initially the highest percentage of disseminated PLSs was related to preventive medicine with a special focus on alternative medicine practices (e.g. acupuncture), nutrition, and child health. We received negative feedback from our followers regarding posts about alternative medicine which were perceived as promotion of those practices even though those reviews did not report positive effect. Therefore, we decided to change the profile of presented topics and started disseminating posts about medical interventions and specific drugs. The profile of people following our posts changed over time, especially in the proportion of men and women and age. Our initial audience was younger, right now we have older audience (35 to 44 years old). On average the organic range of published information is doubled and reach about 350. Total page likes has increased by 15% since 2019 and a similar increase was recorded on our Twitter profile.

Conclusion: In the Polish context comparable efforts on Facebook and Twitter provide more successful dissemination results in the Facebook environment.

Patient or healthcare consumer involvement: We try to approach patient organizations to co-operate with us in the dissemination of PLS. So far we have a successful collaboration with a cystic fibrosis patient organization, for whom we prepare short texts about systematic reviews.

Presenting evidence and practice: a visual interactive map of review-level evidence and digital interventions for drug/alcohol misuse available in England

Burchett H[1], Stansfield C[2], Sutcliffe K[2], Macdowall W[1], Richardson M[2], Dick S[3], Dickson K[2], D'Souza P[2], Khouja C[4], Kwan I[2], Raine G[4], Thomas J[2]

[1] London School of Hygiene & Tropical Medicine, UK; [2] EPPI-Centre, UCL Institute of Education, University College London, UK; [3] University College Cork, Ireland; [4] Centre for Reviews & Dissemination, University of York, UK

Background: There is high unmet need for interventions to prevent, treat and aide recovery from drug/alcohol misuse. Digital interventions have the potential to help overcome barriers to the supply of, and demand for, such interventions. They can use a variety of strategies such as tracking consumption or counselling, delivered through a range of modes, such as apps, videoconferencing and websites. However, it is not known what interventions are available, what evidence exists, or what evidence/practice gaps exist.

Objectives: To develop a visual interactive map that presents the findings to the following research questions: (1) Which types of digital drug/alcohol interventions have been evaluated in systematic reviews? (2) Which types of digital drug/alcohol interventions are currently available in England?

Methods: Public Health England, with the study team, developed a 'pathway' of prevention, treatment and recovery as a framework for categorizing interventions. We populated the map using screening results of database searches for systematic reviews (RQ1) and available interventions identified through an online survey with drug/alcohol commissioners, providers and intervention developers/evaluators (RQ2). We summarized each systematic review by extracting synthetic statements and descriptive characteristics and we appraised quality using AMSTAR2. We wrote a brief description of each available intervention, based on the survey results and information on the intervention's website. We checked these for accuracy with intervention developers. EPPI-Mapper was used to develop the interactive map.

Results: A visual, interactive online map was developed showing 23 systematic reviews and 33 interventions available in England. The number of available interventions and high, medium and low quality systematic reviews, at each pathway point, is visually depicted. Users click on parts of the map to read more about each review and available intervention. The map highlights the predominance of reviews and interventions targeting drug/alcohol misuse prevention, particularly interventions offering feedback or tracking consumption. There was a relative lack of research and practice targeting treatment or recovery and, whilst there were peer support, relapse and overdose prevention interventions, no systematic reviews had focused on these.

Conclusions: This map provides a visual, interactive overview of review evidence and available interventions. It highlights gaps in research as well as in practice. Visual maps showing both evidence and practice could help commissioners to identify what interventions exist, as well as what evidence there is for interventions. It could also help developers identify gaps where new interventions are needed, and researchers to identify evidence gaps.

Patient or healthcare consumer involvement: Although healthcare consumers were not involved in this project, an advisory group of potential users of the map (commissioners and service providers) met regularly to shape the work.

Quality of evidence matters: is it reported and interpreted in infertility journals?

Glujovsky D[1], Ciapponi A[1], Sueldo C[2], Valanzasca MDP[3]

[1] Argentine Cochrane Centre, Institute for Clinical Effectiveness and Health Policy (IECS), CONICET, Argentina; [2] Obstetrics and Gynaecology Department, University of California San Francisco-Fresno, Argentina; [3] Independent researcher, Argentina

Background: Despite the increasing use and popularity of systematic reviews (SRs), it is crucial to assess the certainty or quality of evidence (QoE) that they provide, since the design itself does not guarantee high QoE. The GRADE approach is the soundest system for rating the quality of a body of evidence in SRs and other evidence syntheses.

Objectives: To evaluate if authors of published SRs reported the level of QoE in infertility journals and to analyse if they used an appropriate wording to describe that QoE.

Methods: We selected the five journals focused on human reproduction with the highest impact factors, according to the 2017 impact factor. Firstly, we performed a search in September 2018 on PubMed, identifying the potential SRs with meta-analysis (limits, type of article: meta-analysis) published in 2017. Secondly, we screened the studies by title and abstract to include only those that were SRs of interventions, and where the main subject of study was infertility. We analysed if they classified the published evidence in the full text and also if they did it in the abstract. When study authors did not evaluate the QoE, we used GRADE to analyse it. We described how often the authors used a tool for QoE and the level of the QoE published in the selected SRs. We also analysed if using a tool for QoE was associated with the P value or not, and if the P value was associated with the level of the evidence or not. Finally, we analysed if the study authors made any effort to adapt the wording used in the abstract to the QoE and the magnitude of the described estimated effect.

Results: Study authors reported QoE in only 21.4% of the included SRs and in less than 10% of the abstracts. Although we did not find important differences in the reports of QoE between those that showed statistically significant differences and those that did not, P value was associated with the wording chosen by study authors. In general, magnitude of the effect was not expressed with consistent wording in 54.8% (23/42) of the SRs, while level of QoE was not expressed with consistent wording in 92.9% (39/42) of them. Whereas magnitude of the effect was more consistently expressed in studies with statistically significant findings, QoE was better expressed in those cases in which the P value was over 0.05.

Conclusions: We found that in 2017, in major infertility journals, less than 25% of study authors reported the overall QoE when publishing a systematic review. Study authors still focus more on the discussion as to whether the found difference was due to chance, and less on limitations in the study design, imprecision, indirectness, inconsistency among the primary studies, and publication bias. Study authors should make efforts to interpret results in the context of those evaluations.

Patient or healthcare consumer involvement: One of the authors is a patient.

Readability, linguistic characteristics and conclusiveness of Cochrane Plain Language Summaries: a cross-sectional study

Banić A[1], Fidahić M[2], Šuto J[3], Ščepanović R[1], Vuka I[1], Puljak L[4], Buljan I[5]

[1] PhD candidate at Translational Research in Biomedicine, School of Medicine, University of Split, Split, Croatia; [2] Faculty of Medicine, University of Tuzla, Tuzla, Bosnia and Herzegovina; [3] Graduate Student at the School of Medicine, University of Split, Split, Croatia; [4] Center for Evidence-Based Medicine and Healthcare, Catholic University of Croatia, Zagreb, Croatia; [5] Department of Research in Biomedicine in Health and Cochrane Croatia, School of Medicine, University of Split, Split, Croatia

Background: Cochrane is constantly trying to improve the dissemination of evidence in health to different populations. One of the most important formats, aimed at the lay population are Cochrane Plain Language Summaries (PLSs), which should be written in a simplified language, easily understandable and providing clear message for the consumer.

Objectives: The aim of this study is to assess Cochrane PLSs; to which extent they are customized for lay persons, providing readable, comprehensible and conclusive message to consumers.

Methods: The study analysed a large number of PLSs (N = 4405) of Cochrane intervention reviews in English language collected up to February 2019 to assess the level of readability (SMOG – Simple Measure of Gobbledygook readability formula) and emotional tone. In addition to available Review Group, authors collected descriptive data: year of publishing, number of authors and available languages in which the PLS is provided. Finally, two independent authors assessed the conclusions of Cochrane PLSs and categorized the conclusions into one of the nine categories: "positive", "positive inconclusive", "no evidence", "no opinion", "negative", "negative inconclusive", "unclear", "equal", "equal inconclusive".

Results: Median number of words per summary was 330 (interquartile range (IQR) 213 to 437), with high levels of analytical tone and low levels of emotional tone, which indicates indicate that PLSs are written in an objective style, but with low emotional engagement for the reader. Median number of years of education needed to read the PLSs was 14.9 (IQR 13.8 to 16.1), indicating that the person needs almost 15 years of education in order to read the content with ease. The most prevalent conclusiveness category was the "no opinion" category, indicating that the PLS did not provide the clarified answers about the effectiveness of the therapy, and the categories were similarly dispersed across Cochrane groups (Figure 1).

Conclusions: PLSs are predominantly written in an objective style, with low levels of emotional tone and relatively high readability score, which makes them difficult to read for lay population without medical education. Future analysis will focus on the comparison of linguistic characteristics between PLSs with different conclusiveness categories, with aim to determine whether the conclusion type is related to writing style. Most of the PLSs did not provide a clear opinion regarding the effects of the intervention. Our results indicate that PLSs are not so plain, and that further effort is needed to write PLSs that will be better suited for lay audience.

Patient or healthcare consumer involvement: Patients were not involved in this study.

Additional file: Figure 1

Reliability and feasibility of the Health Evidence quality assessment tool for systematic reviews on the effectiveness of public health interventions

Read K[1], Belitae E[2], Sully E[1], Dang N[1], Dobbins M[1]

[1] National Collaborating Centre for Methods and Tools (NCCMT), Canada; [2] McMaster University, Canada

Background: Health Evidence aims to make it easier for decision-makers to use evidence in their programs and policies. We provide access to over 6000 quality-appraised systematic reviews on the effectiveness of public health interventions. Each review is rated independently by two reviewers using the Health Evidence quality assessment tool and dictionary and consensus is achieved through discussion. The tool includes 10 questions to help assess the methodological quality of public health relevant reviews. The development of this tool has been previously published.

Objectives: To assess the inter-rater reliability and feasibility of the Health Evidence quality assessment tool for systematic reviews on the effectiveness of public health interventions.

Methods: Three reviewers independently assessed a sample of 60 systematic reviews of public health interventions from the Health Evidence registry. All systematic reviews were

1) relevant to public health or health promotion practice;
2) examined the effectiveness of an intervention;
3) include raw data on outcomes; and
4) described a search strategy.

Reviewers had different levels of experience with critical appraisal generally, and previous use of the Health Evidence tool specifically (novice, intermediate, expert). Reliability between the three raters was assessed with the intraclass correlation coefficient (ICC) using a two-way random-effects model, absolute agreement. The average measure was used to report the results, with an ICC of > 0.75 classified as good reliability. The time to complete the assessment form was also tracked to look at the feasibility of using this tool in practice.

Results: Reviewers conducted quality appraisals on 20 articles each month over June, July, and August 2019. After each set, the team met to resolve conflicts prior to completing the next month's set. Overall agreement between all three raters showed good to excellent reliability (ICC = 0.898; confidence interval 0.843 to 0.936) reporting on average measures. In general, time to complete a single quality assessment was under 15 minutes, indicating that the tool is also feasible to apply.

Conclusions: The results of this study suggest that the Health Evidence quality assessment tool is reliable for assessing the methodological quality of systematic reviews on the effectiveness of public health interventions. The next steps will be to compare a selection of reviews using this tool to other comparable critical appraisal tools in the field to identify similarities and differences

Short-term use of therapeutic opioids for children and future misuse: a systematic review and qualitative study of decision-maker information needs

Ahrari M[1], Hartling L[1], Dyson M[1], Ali S[1]

[1] Department of Pediatrics, University of Alberta, Canada

Background: Despite an overall decline in opioid prescriptions in Canada, healthcare visits, hospitalizations, and deaths due to opioid-related harms continue to rise for children. Decision-makers (including families, clinicians, and policy-makers) require high quality syntheses to inform decisions regarding opioid use for children. Previous research has found that how systematic review (SR) results are presented may influence uptake by decision-makers. Evidence summaries are appealing to decision-makers as they provide key messages in a succinct manner.

Objectives: A) To conduct a SR examining the association between short-term therapeutic exposure to opioids and development of opioid use disorder, and B) To gain perspectives from policy decision-makers on the usability and presentation of results through the form of an evidence summary.

Methods: A) We conducted a SR following methods recommended by Cochrane. A medical librarian conducted a comprehensive search and two authors were involved in study selection, data extraction and quality assessment. Studies were eligible if they reported primary research in English or French, and participants had therapeutic exposure to opioids before age 18 years. Results were described narratively. B) Decision-makers were recruited through purposive and snowball sampling methods. They participated in interviews to discuss an evidence summary about the SR. Interviews were transcribed and data was analyzed using latent content analysis.

Results: A) Of 4072 unique citations, 16 studies (634,556 participants) were included. Five studies were comparative and explored the association between therapeutic exposure to opioids and opioid misuse; 11 studies were non-comparative and only reported on prevalence of misuse following therapeutic exposure. One comparative study showed an association between short-term therapeutic use and opioid misuse. The other four studying association lacked information on the duration of exposure; still, all suggested an association between therapeutic exposure and misuse. B) Decision-makers had mixed preferences for the presentation of evidence, depending on their role. A majority shared preferences for having statistics, methods and key characteristics of studies included in the evidence summary. They generally liked key messages highlighted on the first page, but noted the summary should not be too text-heavy.

Conclusions: Preliminary evidence suggests a link between lifetime therapeutic opioid use and opioid misuse; however, there is insufficient evidence available to determine whether short-term therapeutic exposure to opioids in childhood is definitively associated with these disorders. While this SR contributes evidence to guide clinical practice and future research, the qualitative findings help in understanding the type and format of information needed by policy decision-makers. PROSPERO Registration: 122681.

Patient or healthcare consumer involvement: Healthcare decision-makers were engaged to inform the development of effective knowledge translation tools.

SPIRIT-Outcomes and CONSORT-Outcomes: Enhanced trial outcome transparency, less bias, improved systematic reviews, better health

Butcher NJ[1], Monsour A[1], Mew EJ[1], Askie L[2], Grimshaw J[3], Moher D[3], Chan A[4], Gavin F[5], Smith M[6], Offringa M[7]

[1] Hospital for Sick Children Research Institute, Canada; [2] National Health and Medical Research Council Clinical Trials Centre (NHMRC), University of Sydney, Australia; [3] Ottawa Hospital Research Institute, Ottawa, Canada; [4] Women's College Research Institute, University of Toronto, Canada; [5] Canadian Drug Expert Committee, Ottawa, Canada; [6] Cochrane Consumer Executive, Ottawa, Canada; [7] Hospital for Sick Children Research Institute, University of Toronto, Canada

Background: Clinicians, patients, and policy makers rely on published results from clinical trials to help inform evidence-based decision-making. Readers require complete and transparent information with respect to what was planned, what was done, and what was found. Inadequate reporting of trials is well-documented in the medical literature, including for study outcomes. Key information about the selection process, definition, measurement, and analysis of outcomes is often missing or poorly reported in trial protocols and subsequent published reports, impairing reproducibility of results, knowledge synthesis efforts, and prevention of outcome switching and other reporting biases.

Objectives: This international project developed the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT-Outcomes) and Consolidated Standards of Reporting Trials (CONSORT-Outcomes) 2020 reporting guidelines to provide harmonized guidance for describing outcomes in trial protocols and reports, respectively.

Methods: We developed the SPIRIT- and CONSORT-Outcomes reporting guidelines using the EQUATOR (Enhancing the QUAlity and Transparency Of health Research) framework for reporting guidelines. This included the generation and evaluation of candidate outcome reporting items via expert consultations, a scoping review, a three-round international Delphi survey, and a two-day in-person expert consensus meeting. We involved a diverse group of stakeholders throughout the process including those with experience in the design, conduct, oversight, publication, and interpretation of clinical trials (trialists, biostatisticians, health economists, trial registries, research ethics board members, epidemiologists) and stakeholders who use the results of clinical trial reports (journal editors, clinicians, systematic review authors, health technology assessors). Patient and public representatives and research funders contributed during the consensus meeting.

Results: We identified 133 outcome reporting items from the scoping review and expert consultations, the majority of which are not currently included in CONSORT or SPIRIT reporting guidelines. We consolidated items into 67 candidates for Delphi voting, which was completed by 124 participants from 22 countries. After the Delphi survey, 19 items met criteria for further evaluation at the consensus meeting for inclusion in CONSORT-Outcomes and 30 for inclusion in SPIRIT-Outcomes. The consensus meeting and post-consensus meeting finalization process ultimately yielded 8 SPIRIT-Outcomes and 16 CONSORT-Outcomes extension items, focused on outcome definition, rationale of outcome selection, composite outcomes components, minimal important difference and change, measurement properties of study instruments, outcome assessors, and planned adjustments for multiplicity.

Conclusions: SPIRIT-Outcomes and CONSORT-Outcomes provide new frameworks aimed to enhance trial transparency in areas known to be associated with outcome reporting bias.

Patient or healthcare consumer involvement: The patient member, public member, research funder, and industry representatives attended the consensus meeting, helped finalize item wording, and advised on implementation strategies.

SYSVAC: Global registry of systematic reviews and complementary online course to facilitate vaccine recommendation development

Jo C[1], Burchett H[2], Henaff L[3], Younger Z[1], Wichmann O[1], Harder T[1]

[1] Robert Koch Institute, Germany; [2] London School of Hygiene and Tropical Medicine, UK; [3] World Health Organization, Switzerland

Background: National Immunization Technical Advisory Groups (NITAGs) are independent advisory committees that develop evidence-based recommendations to guide national immunization programs and policies. Systematic reviews are recommended to be used in this process, since they synthesize findings from numerous studies and, when done well, provide reliable estimates about intervention effects. However, conducting systematic reviews requires significant resources (time, staff) that many NITAGs do not have. A large number of systematic reviews on vaccination-related topics already exist, and every year more are published. Increasing NITAGs' access to and use of existing systematic reviews could facilitate their process for developing vaccine recommendations.

Objectives: The Robert Koch Institute – in collaboration with the World Health Organization (WHO) and the London School of Hygiene and Tropical Medicine – aims to increase NITAGs' access to and use of existing systematic reviews by developing a user-friendly registry of systematic reviews on vaccination-related topics and an online course on how to use existing reviews in vaccine decision-making. Both the registry and course will be hosted by WHO's NITAG resource center.

Methods: To inform the development of both products, we designed a survey for potential end users (i.e. NITAG members and secretariats, WHO Strategic Advisory Group of Experts (SAGE) members and secretariat) to assess their baseline needs and behaviors with regard to the use of existing systematic reviews. We also assessed the feasibility of appraising the quality of reviews in the registry, by applying the Assessing the Methodological Quality of Systematic Reviews 2 (AMSTAR 2) instrument to a sample of 20 reviews on vaccination-related topics. Lastly, we organized an international experts workshop on methods for using systematic reviews.

Results: Results from the end user survey will be reported at the Cochrane Colloquium. Our feasibility assessment suggested that AMSTAR 2 is easy to apply to reviews in the registry and that the time required for review decreases as users of the instrument gain more experience with it. We used insights from the experts workshop to develop a list of possible concepts for the registry and course (i.e. "Basic," "Expanded," and "Deluxe" versions of both products), a ranked list of ideas for how both products could be set up, and a script for the e-learning course.

Conclusions: Development of both the online registry and course are underway. We will conduct the survey and prepare an article summarizing results from the experts workshop in the second quarter of 2020 and pilot-test draft versions of the registry and course in the last quarter of 2020. Multiple types of formative research are being conducted to maximize the potential utility of the registry and course for NITAG members and secretariats.

Teachers' assessments of the Informed Health Choices key concepts intended for teaching primary school children about health claims: preliminary results

Aranza D[1], Milavić B[2], Marušić A[3], Milić M[4], Poklepovic Pericic T[5]

[1] University Department for Health Studies, University of Split, Croatia; [2] Department of Anthropological Kinesiology and Health, University of Split Faculty of Kinesiology, Split, Croatia; [3] Department of Research in Biomedicine and Health and Cochrane Croatia, University of Split School of Medicine, Croatia; [4] Department for Kinesiological Education and Methodology, University of Split Faculty of Kinesiology, Croatia; [5] Cochrane Croatia, University of Split School of Medicine, Croatia

Background: The Informed Health Choices (IHC) Project developed a list of key concepts to help people understand and judge trustworthiness of treatment claims, with 12 of the key concepts being included in the IHC primary school resources.

Objectives: To evaluate attitudes of primary school teachers about 12 IHC key concepts intended for primary school children.

Methods: The IHC key concepts were presented in primary school settings in urban agglomeration of the City of Split during county expert meetings and teachers' school councils. We asked participating teachers to complete a questionnaire consisting of three parts. Demographic characteristics included teachers' gender, age and work experience. For assessing their overall understanding of the presented concepts, teachers were asked to choose one of the following statements: Not understood, Somewhat understood, Understood, and Understood very well. Furthermore, based on their competencies and their interests, teachers were asked to use a Likert scale of 1 to 6 (1 = lowest, 6 = highest) and provide a numeric score regarding the four criteria: 1) relevance and importance, 2) usefulness, 3) understandability and 4) impact on children of each of the 12 key concepts for both third grade (age 9) and sixth grade (age 12) primary school children. The study was approved by the Ethics Committee of the University of Split School of Medicine and was funded by the Croatian Science Foundation project called "Professionalism in Health – Decision making in practice and research, ProDeM" under Grant agreement No. IP-2019-04-4882.

Results: In total, 304 teachers completed the questionnaire. A total of 83.5% (N = 254) reported having understood the IHC key concepts well or very well (Figure 1). Participants' median age was 42.5 (interquartile range 18), with 22.7% of teachers (N = 69) having five years of work experience or less, 15.79% (N = 48) from five to 10 years, and 61.51% (N = 187) more than 10 years of work experience. In relation to the four criteria, all concepts scored well with mean values ranging from 4.30 to 4.97 on a 1 to 6 scale for sixth grade children, and from 3.92 to 4.77 for third grade children (Table 1). Overall assessments of all four categories for each of the concepts were scored relatively high with mean values ranging from 17.65 ± 4.48 on a 4 to 24 scale to 19.44 ± 3.72 for sixth grade children, and from 16.58 ± 5.20 to 18.61 ± 4.26 for third grade children (Figure 2). Metric characteristics for the overall assessments of all concepts showed they were homogenous, confident (Cronbach alpha from 0.836 to 0.941) and sensitive.

Conclusions: Most teachers understand IHC key concepts well. Teachers considered that teaching primary school children about the IHC concepts was relevant. Slightly lower assessments for the third grade children indicate that, in order to use the IHC key concepts for teaching at this level of primary education, it necessary to explore means of adapting and simplifying the teaching materials.

Patient or healthcare consumer involvement: Primary school teachers involved.

Additional files: Figures

The Cochrane Rehabilitation Ebook Project: a knowledge translation initiative

Gimigliano F[1], Moretti A[1], Lazzarini SG[2], Arienti C[2], Ceravolo MG[3], Kiekens C[4], Negrini S[5]

[1] Department of Mental and Physical Health and Preventive Medicine, University of Campania "Luigi Vanvitelli", Naples, Italy; [2] IRCCS Fondazione Don Carlo Gnocchi, Milan, Italy; [3] Department of Experimental and Clinical Medicine, School of Medicine, Marche Polytechnic University, Ancona, Italy; Centre for Health Care Management, School of Medicine, Marche Polytechnic University, Ancona, Italy; [4] Spinal Unit, Montecatone Rehabilitation Institute, Imola, Italy; Department of Physical and Rehabilitation Medicine, University Hospitals Leuven, KU Leuven, Leuven, Belgium; [5] Department of Biomedical, Surgical and Dental Sciences, University of Milan "La Statale", Milan, Italy; IRCCS Istituto Ortopedico Galeazzi, Milan, Italy

Background: Cochrane Rehabilitation (CR) is working on the production of an Ebook to systematically present all the relevant evidence of rehabilitation interest included in the Cochrane Systematic Reviews (CSRs). In line with the Cochrane Knowledge Translation (KT) strategy, the Ebook aims to fill the knowledge practice gap addressing four different audiences producing one different summary for each of them, structured to meet the needs of the different end-users according to their different knowledge skills, and outcomes of interest: clinicians, medical and health professional students, policy decision-makers and rehabilitation health care managers, patients and caregivers (consumers).

Objectives: To present the CR Ebook project.

Methods: After identifying all the CSRs relevant to rehabilitation, residents from two Italian universities have been involved and instructed on how to write the different summaries, using a structured template, based on the Cochrane's Dissemination checklist, and following the Cochrane Norway language guidelines. Two Physical and Rehabilitation Medicine (PRM) physicians (AM, MGC) revised them, a second revision was provided by a member of CR and a third by at least two among a group of international editors coming from the European Physical and Rehabilitation Medicine (PRM) Bodies supporting and co-authoring the ebook. A last check was then performed by the another PRM physician (FG), before the final decision by the European PRM Bodies took place.

Results: To date, we have identified 375 CSRs published between 2014 and August 2019. Of these, 145 (3 for 2018, 68 for 2017, 60 for 2016 and 14 for 2015) have been assigned to and summarized by the residents. Ninety-six of them (45 for 2017, 48 for 2016 and 3 for 2015) have been finally approved and uploaded on the Ebook website. The remaining ones are undergoing one of the two processes of revision.

Conclusions: The CR Ebook will be officially launched in the European Bodies Meeting, postponed from March to September 2020, and will be progressively filled with a new set of summaries completed and approved. The process started with the CSRs published in 2016 and 2017, and is now proceeding simultaneously onwards and backwards. The project is continuous: the number of CSRs to be summarized is meant to increase every time a CSR is tagged as relevant to rehabilitation, making the Ebook a "live" and updated source of evidence. CR is also planning to translate the summaries into different languages, as is already done for other KT products (i.e. blogshots), in order to reach the widest possible audience.

Patient or healthcare consumer involvement: Patients or healthcare consumers have not been directly involved in the project, but represents one of the different audiences the Ebook addresses.

The development of online knowledge database for Chinese medicine using a whole-evidence approach

Liu S[1], Chen J[1], Yang L[1], Guo X[1]

[1] The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, China

Background: The evidence of many Chinese medicine interventions has not been systematically explored and validated with the evidence-based medicine (EBM) approach. There is a limited number of high-quality randomized controlled trials in Chinese medicine. Currently, ancient classical books and expert consensus are still playing an important role in the Chinese medicine clinical practice. The best available evidence of Chinese medicine may include expert consensus, modern studies and ancient literature evidence. Under the circumstances, it is necessary to integrate and transform scattered, different types of evidence to efficiently facilitate rapid guidance for clinical practice.

Objectives: To develop an online knowledge database of Chinese medicine using a whole-evidence approach.

Methods: We developed a "whole-evidence approach", which aims to collate, synthesize and evaluate evidence arising from clinical trials, experimental studies, classical literature and expert consensus. We collected modern clinical and experimental studies and evaluated them according to the Cochrane Review methods. We used the GRADE approach to assess the certainty of evidence. The classical literature was searched and screened in Zhong Hua Yi Dian, one of the most comprehensive Chinese ancient literature collections. The analyses revealed similarities and differences between traditional and modern evidence. The collected expert consensuses were compared with the above forms of evidence. The intelligently retrievable, shareable evidence-based knowledge database was developed by computer technology to show the available evidence map for Chinese medicine.

Results: We systematically assessed the whole-evidence of Chinese medicine for 29 conditions, involving Chinese herbal medicine, acupuncture and other Chinese medicine treatments. The evidence of randomized controlled trials, non-randomized controlled studies, non-controlled studies, classical literature and expert consensus were synthesized and compared to provide summaries into clinical practice. Larger weight was assigned to randomized controlled trial during whole evidence in clinical decision making. Eighteen monographs have been published in English and Chinese. The evidence-based online Chinese medicine knowledge database was established and could be easily retrieved, indexed and updated.

Conclusions: The whole-evidence approach has summarized traditional and modern evidence, expert consensus, clinical and experimental evidence together and is a major milestone in Chinese medicine. The online Chinese medicine knowledge database with whole-evidence approach can improve the efficiency of evidence practicing, promotes rapid clinical decision-making and benefits patients.

Patient or healthcare consumer involvement: Clinician, educator, patients.

The rehabilitation definition for scientific research purposes by Cochrane Rehabilitation: first results

Negrini S[1], Meyer T[2], Stucki G[3], Arienti C[4], Pollock A[5], Kiekens C[6]

[1] Department of Biomedical, Surgical and Dental Sciences, University «La Statale», Milan; IRCCS Istituto Ortopedico Galeazzi, Milan, Italy; [2] School of Public Health, Bielefeld University, Germany; [3] Department of Health Sciences and Medicine, University of Luzern; Swiss Paraplegic Research, Nottwil, Switzerland; [4] IRCCS Fondazione Don Carlo Gnocchi, Milan, Italy; [5] Nursing, Midwifery and Allied Health Professions Research Unit, Glasgow Caledonian University, Glasgow, UK; [6] Spinal Unit, Montecatone Rehabilitation Institute, Imola, Italy; Department of Physical and Rehabilitation Medicine, University Hospitals Leuven, KU Leuven, Leuven, Belgium

Background: Since its launch in 2016, Cochrane Rehabilitation (CR) has increasingly found the need to better define what rehabilitation is and what it is not. We found that currently available definitions of rehabilitation fall short with regard to defining exactly what is needed for research purposes, specifically inclusion and exclusion criteria on an operational level.

Objectives: To specify a definition of rehabilitation suitable for research purposes that also defines inclusion and exclusion criteria.

Methods: The methodology was developed based on discussions during the CR Executive and Methodology Committees Meetings and with the Advisory Board of CR, which includes all relevant rehabilitation stakeholders, during its meeting in Kobe in June 2019. The methodology was then refined to include: (1) a survey involving a) the Advisory and b) the Executive Committee members and c) the participants invited to the consensus meeting here reported to collect the current definitions promoted by CR stakeholders; (2) a consensus meeting (CM) held in Milan (Italy) in February 2020; and (3) a Delphi procedure to reach a final definition.

The CM was organized in three parts. Part 1: presentation and discussion of current definitions and related problems. Part 2: work in smaller groups to prepare a proposal of the new operational definition of rehabilitation for research purposes. Part 3: discussion and voting procedures on the proposal to be submitted to the Delphi procedure. The first round of the Delphi procedure will be conducted among the meeting participants and the second round among the CR Advisory Board members. The results of the two rounds will be reviewed and synthesized, and a third Delphi round will be held again among the participants to the Meeting.

Results: During the meeting it was decided that the proposed definition should contain all the essential key words that constitute the inclusion and exclusion criteria for research purposes, following the PICO (Population, Intervention, Control, Outcome) model as much as possible. The current definition to be proposed for the Delphi Rounds is: In a health care context rehabilitation is: "a multimodal person-centered process, including functioning interventions targeting (1) body functions, and/or (2) activities and participation, and/or (3) the interaction with the environment" (Intervention), with the goal of "optimizing functioning" (Outcome) for "(1) persons with health conditions (a) experiencing disability or (b) likely to experience disability and/or (2) persons with disability" (Population). This definition is currently provisional, and it should not be used until the final results of the Delphi are published.

Conclusions: These results will inform all the future work of CR and could also serve the scientific rehabilitation community.

Patient or healthcare consumer involvement: Representatives of healthcare consumers are present in the Advisory Board of CR and contributed at all stages of the project when the Advisory Board was involved.

The Systematic Review Data Repository 2.0 (SRDR 2.0): summary data for end-users

Saldanha I[1], Jap J[1], Jap R[2], Senturk B[1], Smith B[1], Balk E[1]

[1] Brown University School of Public Health, USA; [2] Independent Consultant, USA

Background: Reports of systematic reviews (SRs) are often too voluminous to be useful for decision-makers. This problem can potentially be solved through interactive displays of key summary data from SRs, with the option to dig deeper into the evidence as needed. To make summaries of SR data more readily accessible to decision-makers digitally and interactively, we have developed and user-tested a new free platform, the Systematic Review Data Repository 2.0 (SRDR 2.0). SRDR 2.0 is funded by the Agency for Healthcare Research and Quality (AHRQ) in the United States.

Objectives: To describe the development of SRDR 2.0 and to demonstrate its use.

Methods: We followed three steps:

1) identifying potential structured elements from common items reported in 10 published SRs;
2) finalizing the list of structured elements through iterative discussions with 21 international academic and non-academic stakeholders with diverse perspectives; and
3) building, refining, and user testing a prototype for SRDR 2.0.

We have developed SRDR 2.0 as a platform that provides the information of greatest interest to stakeholders in a structured and user-friendly manner.

Results: The elements in SRDR 2.0 are organized into three levels: Level 1 (meta-data and other information related to the entire SR), Level 2 (related to specific populations, interventions/exposures, comparators, outcomes, study designs, and settings of interest), and Level 3 (related to results for specific outcomes). As illustrated in Figure 1, SRDR 2.0 gathers information from three different types of sources – the SR report (i.e. publication), SRDR (i.e. online data system), and various external online locations. SRDR 2.0 displays the requested information through interactive, user-friendly formats, such as accordion-style headings, where the user clicks on headings to reveal underlying information, and mouse-overs, where underlying information is revealed only when the user hovers the mouse over a heading (see Figure 2).

Conclusions: SRDR 2.0 aims to be a new platform for sharing summary SR data digitally and interactively with diverse end-users, such as guideline developers, clinical decision support tool developers, clinicians, patients, and other consumers. The vision is that these diverse end-users will be able to query SRDR 2.0 and usable information about a given SR in a structured format. SRDR 2.0 is part of a suite of new advancements being planned by AHRQ to make its SR reports and other products interactive. At the Colloquium, we will demonstrate live the use of SRDR 2.0. By helping present the relevant information from various types of SRs to decision-makers, SRDR 2.0 has the potential to greatly facilitate evidence-based decision-making.

Patient or healthcare consumer involvement: We involved a consumer representative among the stakeholders who helped define the structured elements. The SRDR 2.0 platform is designed to be used by patients, consumers, and various other end-users.

Additional file: [Figures](#)

The use of evidence and its contribution to action in sexual and reproductive health: lessons from contribution mapping in Nigeria

Esu E[1], Arikpo D[1], Chibuzor M[1], Oringanje C[1], Borst R[2], Kok M[2], Meremikwu M[1]

[1] Cochrane Nigeria, Calabar Institute of Tropical Diseases Research and Prevention, University of Calabar Teaching Hospital, Calabar, Nigeria;
[2] Erasmus School of Health Policy & Management, Erasmus University Rotterdam, The Netherlands

Background: Despite the development of methods, strategies and structures, the translation of evidence about sexual and reproductive health (SRH) into policies and practices and ultimately better health, remains a challenge.

Objectives: We used a novel method, contribution mapping, to assess the use of evidence and its contribution to action in SRH research in Nigeria.

Methods: This project was undertaken from September 2019 to January 2020. A qualitative approach was used, combining a search of the literature and policy documents and in-depth interviews (face-to-face and teleconference calls) with Nigerian stakeholders working in SRH. Eleven SRH research projects conducted in Nigeria between 2015 and 2019 were identified. These cases cut across non-governmental organizations, government agencies and academia. Two researchers conducted in-depth interviews with the use of digital audio recorders and notetaking to capture the non-verbal expressions of study participants. Each interview on average lasted about 35 minutes. We obtained verbal and written consent before each interview. We transcribed all interviews verbatim and conducted thematic analysis.

Results: The origins of research were generally based on problems identified in SRH programming and in routine medical practice. Research team compositions among academics was generally the same. Additional research team members were recruited based on the competencies and skills which were required to complete the research projects. For the NGOs and government agencies, the research teams tended to be larger and there was a lot more interactions with stakeholder external to the research team. The underlying motivation for many of the research projects was to solve problems while also advancing career progression. The stakeholder's ability to influence changes at the policy level appeared to be largely through their membership of National Technical Working Groups in SRH. It is through this medium that they share their results using mainly PUSH knowledge exchange mechanisms. Cultural sensitivity still exists around sexuality, especially for young people. Funding streams for research are not properly institutionalized. Several participants felt interaction between researchers and policy makers was insufficient. There were different mechanisms and pathways through which change happened as a result of the projects. Some achieved change by advocacy and for other stakeholders, their dissemination efforts led to interactions that became the precursors for change.

Conclusions: In Nigeria, knowledge platforms for SRH exist and some efforts are made to base policy and recommendations of research evidence. These knowledge platforms however seem to be clustered at federal level with only minimal activity at state level.

Patient or healthcare consumer involvement: This project involved healthcare consumer groups in priority setting at the beginning of the project.

The use of GRADE on dose–response meta-analysis in 2019: a cross-sectional study

Hui X[1], Li J[1], Lu ZX[2], Dai H[3], Ma XY[4], Yan PJ[5], Yao L[6], Yang KH[7]

[1] School of Public Health, Lanzhou University, Lanzhou, China; [2] First clinical medical college, Lanzhou University, Lanzhou, China; [3] First Clinical Medical College, Lanzhou University, Lanzhou, China; [4] Second Clinical Medical College, Lanzhou University, Lanzhou, China; [5] Institution of Clinical Research and Evidence Based Medicine, Gansu Province People's Hospital, Lanzhou, China; [6] Health Research Methodology/Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, Ontario, Canada; [7] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China

Background: An increasing number of dose–response meta-analyses (DRMAs) have been published in the past several years. The GRADE system is considered to be a common, sensible and transparent approach to grading quality of evidence and strength of recommendations. The quality of evidence has four levels: high, moderate, low and very low. However, the use of GRADE system on DRMAs has not been investigated.

Objectives: To investigate GRADE evidence assessment of DRMAs published in 2019 through a cross-sectional survey.

Methods: We will search PubMed to identify DRMAs published from 1 January 2019, to 31 December 2019, using the following search strategy: (meta-analysis [Title/Abstract]) AND (dose–response [Title/Abstract]) AND ("2019/1/1" [Date – Publication]: "2019/12/31" [Date -Publication]). Two review authors will independently screen the literature, extract the baseline characteristics and collect all data in Excel 2019 (Microsoft, Washington). Any disagreement will be resolved by the third review author. We will use Stata 15.0 (STATA, College Station, TX) for the analysis, with a P value of < 0.05 denoting significance.

Results: This study is ongoing and will be submitted to a peer-reviewed journal for publication.

Patient or healthcare consumer involvement: There was no patient or healthcare consumer in this project.

To blog or not to blog? Using blogs as a tool to provide consumer friendly evidence-based information on optimal aging

Farran R[1], Neil-Sztramko S[2], Dobbins M[1]

[1] School of Nursing, McMaster University, Canada; [2] Department of Health Research Methods, Evidence and Impact, McMaster University, Canada

Background: The integration of technology into everyday life has resulted in unprecedented access to information. However, in the age of "fake news" it is difficult to know which messages relevant to healthy aging are trustworthy. Difficulty understanding scientific research and identifying practical takeaways can also be barriers to making evidence-informed health decisions. The McMaster Optimal Aging Portal (the Portal), is a website that bridges the gap between research and decision making by curating, evaluating, and synthesizing research and resources to provide older adults, caregivers, public health professionals, clinicians, and policymakers with evidence-based information on optimal aging. One knowledge translation tool used by the Portal is Blog Posts that translate the best available evidence on healthy aging into easy-to-understand messages accessible to citizens.

Objectives: To share the process of creating and disseminating high-quality Blog Posts that translate information on healthy aging, and key lessons learned.

Methods: Potential blog topics are identified quarterly through scientific abstracts available on the Portal, news headlines, national and global Google search trends, citizen and expert suggestions, and Portal website analytics. The core of each Blog Post is based on findings from a high-quality systematic review fed to the Portal through one of four internationally recognized databases for public health, clinical and policymaking evidence. Using an established template, an outline with background information and key evidence is produced and sent to a professional writer who ties everything together in engaging, and easy to understand language.

Results: Around 26 health-related Blogs Posts are produced and published annually. Generally, each blog includes: an engaging title; summary, what the blog is about; relevant image; a "bottom line", highlight of main takeaways; background information on the disease and/or intervention; what the research tells us; featured online resources rated for evidence use, transparency, and accessibility; and references. Since 2014, 196 health-related Blog Posts have been published. Overall, Blog Posts are the most popular form of content on the Portal and contributed 58% of the Portal's content-related page views in 2019.

Conclusions: The Portal gives citizens, health professionals, policymakers and researchers direct access to trusted, evidence-based information about optimal aging to help users remain healthy, active and engaged as long as possible. Blog Posts are one useful tool that assist in the dissemination of this information in consumer friendly language.

Patient or healthcare consumer involvement: Consumers – namely older adults, their caregivers, and health experts – are involved in various stages of the blog production process. This includes blog topic selection (e.g. provision of topic suggestions), and content contribution (e.g. sharing personal experiences or professional expertise) or review, where appropriate/needed.

Towards rapid learning health systems: supporting the Central Asian and European countries in using research evidence for policy-making

Reinap M[1], Kuchenmüller T[1]

[1] WHO Regional Office for Europe, Denmark

Background: Despite available research evidence that can offer solutions to many public health and health systems challenges and enable rapid learning, deficits in knowledge translation remain a major barrier for using evidence in making health decisions. This is especially the case for countries that have limited capacities, are in transition and where access to research evidence is limited, e.g. if there are language barriers. Since 2013 the Evidence-Informed Policy Network (EVIPNet) Europe has been working extensively on strengthening national capacities in 23 countries in the WHO European Region, including countries such as Kazakhstan and Kyrgyzstan. These efforts focus on effectively and systematically translating and using the best available research evidence in decision-making.

Objectives: To present the activities and progress as well as discuss enablers and barriers in implementing and institutionalizing knowledge translation (KT) and evidence-informed policy-making (EIP) for rapid learning health systems in European and Central-Asian countries.

Methods: We use network activities to offer an enabling platform for capacity strengthening, fostering regional collaboration and exchange of knowledge and experience. We have helped building capacity, including in conducting rapid syntheses for rapid learning health systems. This was also achieved by making tools and resources available that focused on knowledge translation (KT) tools, such as the evidence brief for policy (EBP) and policy dialogues (PD). We also worked with countries to identify major factors that facilitate or hinder establishment of a KT platform and supporting the countries to work towards institutionalization of it.

Results: EVIPNet Europe has continued to make progress in the promotion of, increasing capacities and involving stakeholders in KT in its member countries in Europe and Central Asia. They have acknowledged the importance of evidence-informed decision-making and recognize existing knowledge translation deficit and gaps in their national capacities and legal framework for KT platforms. Many countries have excellent examples of KT tools and proposed concrete steps to institutionalize evidence-informed decision-making and to continue foster and strengthen national capacities based on a comprehensive analysis. The next steps of developing formal KT platforms, including establishing rapid response services in Hungary are expected in 2020 and strengthening the Network's focus on Eastern European and Central Asian countries.

Conclusions: Our efforts in supporting countries in building capacities in using research evidence for policy-making have led to an increase in awareness and concrete steps in institutionalization of KT. We will continue to work with countries to support their efforts in this regard.

Patient or healthcare consumer involvement: The network establishes links in countries between researchers, policy-makers and other stakeholders, including civil society actors representing public, patients or consumers involved in policy making.

Using a distribution-based approach and systematic review methods to derive minimum clinically important differences

Watt J[1], Veroniki AA[2], Tricco A[1], Straus S[1]

[1] University of Toronto, Canada; [2] University of Ioannina, Greece

Background: Clinical interpretation of changes measured on a scale is dependent on knowing the minimum clinically important difference (MCID) for that scale: the threshold above which we (e.g. clinicians, patients, and researchers) perceive a difference in an outcome. Until now, approaches to determining the MCID were based upon individual studies or surveys of experts. However, the comparison of treatment effects derived from a pairwise meta-analysis or network meta-analysis (NMA) to all trial-specific results in a meta-analysis could improve our clinical understanding of treatment effects derived from meta-analysis models. Furthermore, the calculation of a MCID based on a systematic review could enhance clinical decision-making when the MCID for a scale is unknown.

Objectives: To demonstrate how a distribution-based approach of pooled standard deviations (SDs) can be used to estimate MCIDs.

Methods: We approximated MCIDs using a distribution-based approach that pooled SDs associated with baseline mean or mean change values for two scales (i.e. Mini-Mental State Exam (MMSE) and Alzheimer Disease Assessment Scale – Cognitive Subscale (ADAS-Cog)), as reported in parallel randomized controlled trials (RCTs) that were included in a systematic review of cognitive enhancing medications for dementia (i.e. cholinesterase inhibitors and memantine). We excluded RCTs that did not report baseline or mean change SD values. We derived MCIDs at 0.4 and 0.5 standard deviations (SDs) of the pooled SD.

Results: We showed that MCIDs derived with a distribution-based approach approximated published MCIDs for the MMSE and ADAS-Cog. For the MMSE (51 RCTs, 12,449 patients), we estimated a MCID of 1.6 at 0.4 SDs and 2 at 0.5 SDs based on baseline SDs and we estimated a MCID of 1.4 at 0.4 SDs and 1.8 at 0.5 SDs based on mean change SDs . For the ADAS-Cog (37 RCTs, 10,006 patients), we estimated a MCID of 4 at 0.4 SDs and 5 at 0.5 SDs based on baseline SDs and we estimated a MCID of 2.6 at 0.4 SDs and 3.2 at 0.5 SDs based on mean change SDs. MCIDs were unchanged when we excluded studies in which SDs were estimated from other measures of uncertainty (e.g. standard error, 95% confidence interval).

Conclusions: A distribution-based approach using data included in a systematic review can approximate MCIDs. Our approach performed better when we derived MCIDs from baseline as opposed to mean change SDs. This approach could facilitate clinical interpretation of outcome measures reported in RCTs and systematic reviews of interventions. Future research should focus on the generalizability of this method to other clinical scenarios.

Patient or healthcare consumer involvement: Two clinicians (Straus and Watt) were involved in the design and interpretation of study results.

Using theory of change to assess impact of knowledge translation initiatives

Dobbins M[1], Howarth C[1], Husson H[1], Steinberg M[2]

[1] National Collaborating Centre for Methods and Tools, Canada; [2] Marla Steinberg Consulting, Canada

Background: The National Collaborating Centre for Methods and Tools (NCCMT) champions the use of evidence in decision-making. We share knowledge about what works in public health and provide high-quality resources as well as essential training and mentoring to support capacity development for evidence-informed decision making. The result is that practitioners make informed decisions so that all people living in Canada can achieve optimal health. As an essential component to this work, the NCCMT recognizes the importance of high-quality evaluations that are underpinned by a rigorous framework.

Objectives: The NCCMT developed an evaluation framework to effectively illustrate the impact of the NCCMT in relation to the reach of the centre's activities, the quality of available products and services, and the impact on public health professionals and organizations. The goal of this framework is to guide evaluations of our programs and services to ensure that the NCCMT's resources and training continue to be high quality and meet the needs of public health professionals globally.

Methods: The NCCMT contracted an external program evaluation expert to assist in the development of a revised evaluation framework that would better capture: reach, quality and impact. The contractor reviewed a number of internal documents: vision, mission, goals; annual workplans; annual reports (submitted to funder), and previous evaluation reports. A review of program evaluation literature was also conducted. An inductive approach was used to identify evaluation themes.

Results: A revised evaluation framework, influenced by theory of change, was developed and is now being implemented at the NCCMT. The evaluation framework encompasses four stages, each with specific outcome indicators identified, that progress from creation of products and services, to knowledge and awareness of these resources, to engagement and capacity development, and finally to policy and practices changes that become embedded within organizational structures. Specific outcome indicators include: confidence, knowledge, skills, organizational supports for evidence-informed decision making, using evidence in decision making, conducting rapid reviews; critical appraisal of evidence, improved programs, efficiency, engagement of workforce, and workforce competence for evidence-informed decision making.

Conclusions: The innovative evaluation framework informed by theory of change resulted in significant streamlining of the NCCMT's evaluation activities, and has provided clearer direction with respect to the purpose of evaluation, as well as enhanced ability to capture the centre's reach, quality and impact.

Patient or healthcare consumer involvement: Not applicable.

Writing evidence summaries for the general public: lessons learned from collaboration with patient partners

Linkins L[1], Iorio A[1], West C[2], Geerts W[3], Duffett L[4], Jessie M[5], Dubois S[2]

[1] McMaster University, Canada; [2] CanVECTOR Patient Partner, Canada; [3] Sunnybrook Health Sciences Centre, Canada; [4] The Ottawa Hospital, Canada; [5] Jewish General Hospital, Canada

Background: There is a growing need to provide information to patients that goes beyond simple commentaries. Patients want to understand how the results of high-quality research apply to them to better understand what they should be discussing with their own doctor about their care. One of the proposed ways to strengthen the connection between expert physicians/researchers and the general public is to involve people who have the disease and are willing to contribute to both research and education as patient partners.

Objectives: To describe the collaboration between patient partners and a Canadian group of thrombosis experts (CanVECTOR) on translating evidence from research studies/reviews into summaries that are easily understood by the general public.

Methods: Research articles to be translated for the public are selected from two different pools: 1) recent critically appraised articles highly rated by jury members including patient partners, or 2) "best evidence to date" articles that answer questions raised by patient partners. Summary writers are primarily thrombosis residents or clinical fellows who are mentored by an expert thrombosis consultant. Summaries are written online using an electronic platform that allows for collaboration between mentee, mentor, editor and copy editor. The format of the summary, designed with input from patient partners, is structured to be easy to read with minimal extraneous detail. Results are presented in two different ways for the lay audience – numerically within a Table and in simple language text (Figure 1). Patient partners review the final draft of the summary and may make additional suggestions to improve clarity.

Results: From May 2018 to present, 18 Patient Evidence Summaries have been published on the CLOT+ website. The top summary has been viewed over 21,000 times. Posters and a workshop based on the process of writing the summaries have been presented at a national conference. Several key thrombosis education groups show feeds linking to the evidence summaries on their websites. This process for writing patient evidence summaries has resulted in lessons learned around three key principles:

1) Context: Patient partners appreciate the "What's the Issue?" section because it helps to show where the research fits within what is known (or not known) about the topic. This requires input from content experts.
2) Clarity: connections linking concepts together that are well understood by experts can be confusing for lay people. This requires careful attention to definitions, and a detailed review by patient partners.
3) Applicability: it is important for a patient partner to know if the data presented is directly relevant to them as an individual. This is the key to knowledge translation.

Conclusions: Translating high quality research evidence into lay summaries only works if there is respectful collaboration between content experts and patient partners.

Patient or healthcare consumer involvement: Patient partners were involved in all aspects

Additional file: [Figure](#)

GLOBAL HEALTH AND EQUITY

A survey of characteristics and potential contribution of registered studies for 2019 novel coronavirus disease (COVID-19)

Yang M[1], Shang Y[1], Tian Z[2], Xiong M[2], Lu C[3], Jiang Y[1], Zhang Y[4], Zhang Y[1], Jin X[5], Jin Q[6], Zhang Y[1], Willcox M[7], Liu J[3]

1 Centre for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China; 2 Centre for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, Beijing; Dongzhimen Hospital, Beijing University of Chinese Medicine, Beijing, China; 3 Centre for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, Beijing; Institute of Integrated Traditional Chinese Medicine and Western Medicine, Guangzhou Medical University, Guangzhou, China; 4 School of Traditional Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China; 5 Third Affiliated Hospital, Beijing University of Chinese Medicine, Beijing, China; 6 Dongzhimen Hospital, Beijing University of Chinese Medicine, Beijing, China; 7 School of Primary Care, Population Sciences and Medical Education, University of Southampton, Aldermoor Health Centre, Southampton, UK

Background: The World Health Organization characterized the 2019 novel coronavirus disease (COVID-19) as a pandemic on 11 March 2020. As of 31 March, 803,541 people were confirmed infected with COVID-19 in 201 countries, areas or territories with cases. As a novel coronavirus disease, there is no specific treatment to guide the clinical practice, and the majority of care is under the guidance of clinical experience and symptomatic treatment. The first clinical trial was registered on the China clinical trials registry on 23 January 2020. Facing the increasing ongoing trials, it would be important to understand the research questions and characteristics of these registered studies. We were also interested in how the emerging evidence can inform clinical practice by providing reliable evidence in the prevention and treatment of COVID-19.

Objectives: To review the characteristics of registered trials on COVID-19 and to provide guidance for future trials to avoid duplicated effort.

Methods: We searched all the studies on COVID-19 registered before 3 March 2020 on eight registry platforms worldwide. We extracted and analyzed the data on the design, participants, interventions, and outcomes. We screened promising trials based on study design, rationale, and resource availability.

Results: We found 393 studies registered between 23 January 2020 and 3 March 2020 . Of these, 380 (96.7%) studies were from mainland China, 3 from Japan, 3 from France, 2 from the US, and 3 were international collaborative studies. A total of 363 studies (92.4%) recruited participants from hospitals and 266 studies (67.7%) aimed at therapeutic effect, others were for prevention, diagnosis, prognosis. Two hundred and two studies (51.4%) were randomized controlled trials (RCTs). The average sample size was 1061 and ranged from 8 to 150,000 per study. One hundred and seventy-nine of 266 therapeutic studies (67.3%) tested Western medicines including antiviral drugs (17.3%), stem cell and cord blood therapy (10.2%), chloroquine and derivatives (8.3%), 15 (5.6%) tested Chinese medicines, and 72 (27.1%) tested integrated therapy of Western and Chinese medicines. Only 31 studies among 266 therapeutic studies (11.7%) used mortality as primary outcome, while the most reported secondary outcomes were symptoms and signs (47.0%). One hundred and six studies (27.0%) were funded by the government, and 268 (68.2%) demonstrated ethical approval. One hundred and seventy-nine of 266 (67.3%) studies had not started recruiting until 3 March 2020. Only nine RCTs were evaluated as promising trials.

Conclusions: The majority of studies focused on assessing therapeutics for COVID-19 but inappropriate outcome setting, delayed recruitment and insufficient numbers of new cases in China implied many studies may fail to complete. Strategies and protocols of the studies with robust and rapid data sharing from international collaboration are warranted for emergency public health events, helping to accelerate priority setting for timely evidence-based decision-making. **Patient or healthcare consumer involvement:** Trials on COVID-19 concern patients and healthy people worldwide.

Addressing inequities in healthcare research

Kumar N[1], Hill M[2], Kariyawasam D[2], Jahanfar S[2], Haas D[1]

[1] Indiana University School of Medicine, USA; [2] Central Michigan University School of Public Health, USA

Background: It has been well established that interventions may have variations in safety and efficacy across various populations. Response can differ based on multiple factors, including age, sex, gender, and race. Genetic differences in the expression of metabolizing enzymes or therapy targets drive some of these altered effects. However, research does not adequately explore these potential differences and resulting guidelines are unclear about how to proceed in these cases. Trial authors should be encouraged to make individual patient data (IPD) available by key demographics, which would allow systematic review teams to run a more flexible and sophisticated meta-analysis to analyze potential differences.

An example: A retrospective cohort study reporting labor outcomes after induction noted that non-white race was independently associated with increased odds of delivering by cesarean, hemorrhage, transfusion, and peripartum infection (Singh 2018). Another study analyzing labor outcomes by race after specifically using vaginal prostaglandins for induction noted that black mothers were more likely than any other group to undergo cesarean sections and have these performed due to non-reassuring fetal heart rate tracings. Hispanic mothers in this study were more likely to have postpartum hemorrhage than other groups (Stephenson 2015). Considering these populations have much higher rates of maternal mortality, it may be ideal to use particular induction methods over others to minimize these risks. Unfortunately, limited data exist to establish efficacy on a racial basis.

Driving factors: Some of the inequities arise from a fear of perceived exploitation of minorities, given a history of unethical research practices. Minority researchers are more likely to focus on disparities but are less likely to get federal funding. One review of inequalities in the research noted that black investigators were half as likely to receive the National Institute of Health (NIH) grants as white investigators even after controlling for education, training, and experience. Furthermore, only 10.9% of NIH grant reviewers, who are chosen from the already diluted pool of successful grant winners, are underrepresented minorities (Konkel 2015).

Conclusions: While healthcare equity is driven by many other factors such as implicit bias, mistrust and systemic barriers, we must start analyzing interventions in a multifactorial manner to explore the intersection of demographical differences in both efficacy as well as the magnitude of response. When planning a systematic review, authors should consider if the intervention has potentially different responses amongst different populations and, if so, run an IPD meta-analysis, which is widely considered to be the "gold standard" approach. The results should then be disseminated in a culturally competent evidence package, so consumers and providers can weigh the risks and benefits of each intervention on a more individualized basis.

Assessment of health equity considerations in WHO guidelines from 2014 to 2019: a cross-sectional survey

Dewidar O[1], Tsang P[2], León M[3], Mathew C[1], Antequera A[3], Baldeh T[4], Akl EA[5], Schünemann H[6], Petkovic J[1], Pottie K[7], Alonso-Coello P[8], Tugwell P[9], Piggott T[4], Welch V[1]

[1] Bruyere Research Institute, University of Ottawa, Canada; [2] Queensway-Carleton Hospital, University of Ottawa, Canada; [3] Iberoamerican Cochrane Centre, Sant Pau Biomedical Research Institute, Spain; [4] Department of Health Research Methods, Evidence, and Impact, McMaster University, Canada; [5] Department of Internal Medicine, American University of Beirut, Lebanon; [6] Department of Medicine, McMaster University Health Sciences Centre, Canada; [7] Departments of Family Medicine and Epidemiology and Public Health, University of Ottawa, Canada; [8] Iberoamerican Cochrane Centre, Sant Pau Biomedical Research Institute (IIB-Sant Pau), CIBER de Epidemiología y Salúd Pública (CIBERESP), Spain; [9] Ottawa Hospital Research Institute, Clinical Epidemiology Program, Canada

Background: The World Health Organization (WHO) supports the UN to deliver the 2030 Agenda for Sustainable Development Goals (SDG), which primarily calls for ending poverty, protecting the planet and reducing health inequities, through the development of health guidelines. These guidelines are developed using the GRADE approach to ensure that high quality of evidence is used. The GRADE group developed a series of Evidence to Decision (EtD) frameworks in 2016 which build on the GRADE approach while providing a set of expanded criteria, including impact on health equity, to explicitly collect details on factors that influence the strength and direction of recommendations. The WHO released the second edition of its technical handbook on guideline development in 2014, which included a chapter on integrating gender, equity and human rights into guidelines.

Objectives: To evaluate how and to what extent health equity considerations are assessed in WHO guidelines.

Methods: We conducted a cross-sectional survey on guidelines approved by the WHO guideline review committee (GRC) published between January 2014 and May 2019. We assessed guidelines for health equity considerations using the PROGRESS (Place, Race, Occupation, Gender, Religion, Education, Socioeconomic Status, Social Capital) framework. We also assessed how likely impact on health equity was assessed in the research evidence section of each recommendation of these guidelines using criteria based on differences in baseline risk, value of outcomes for socially disadvantaged populations, health inequity as an outcome, equity related subgroup analysis, and applicability.

Results: We identified 111 WHO guidelines published in this time period and 91% (101/111) of them were focused on socially disadvantaged populations. The use of the EtD frameworks progressively increased from 10% (2/20) in 2014, to 100% (8/8) in 2019. Mention of health equity anywhere in the guideline increased from 55% (11/20) in 2014 to 100% (8/8) in 2019. For the 40 guidelines that reported using and published their EtD frameworks, likely impact on health equity was supported by evidence statements in 28% (94/332) of the recommendations. The most common evidence provided was differences in baseline risk of outcomes (23%; 78/332 recommendations), and the least frequently reported as differences in the magnitude of effect (0.9%, 3/332).

Conclusion: Consideration of the importance of health equity in WHO guidelines has increased since the adoption of the EtD framework. However, there is a lack of evidence provided to support judgements for likely impact on health equity. For the United Nations' global agenda to succeed, high quality medical research is needed in conjunction with informed decision. Better harnessing these characteristics of rapid learning health systems would eventually lead to an improvement in individual experience and health outcomes.

Patient or healthcare consumer involvement: no

Association between the burden of behavioural risk factors and the number of primary studies included in Cochrane Systematic reviews

Madera Anaya MV[1], Frassetto C[2], Siebert M[3], Megranahan K[4], Roberts D[5], Plugge E[6]

[1] University of Cartagena, Colombia; [2] Catholic University of Sacred Heart, Italy; [3] Université de Rennes, France; [4] University of London, UK; [5] University of Oxford, UK; [6] Cochrane UK, UK

Background: Alcohol and drug use (A&D) and dietary risk are two increasingly important risk factors. It has been reported that there is a weak correlation between Cochrane Systematic Reviews (CSRs) of risk factors and their global risk burden. Likewise, only 19% of the corresponding authors in the CSRs are from low- and middle-income countries (LMICs) even though more than 83% of the global population live in these areas of the world.

Objectives: To examine whether there is an association between the burden of a risk factor in countries of specific income bands as defined by the World Bank, and the number of primary studies included in CSRs conducted in those countries.

Methods: We extracted data from primary studies included in CSRs assessing two risk factors (A&D and dietary risk) as outcomes. For each, we obtained data on its overall burden in disability-adjusted life years (DALYs) by World Bank income levels and calculated the correlation between DALYs and the number of primary studies and also their participants.

Results: We included 1601 studies from 96 CSRs. Only 18.3% of the global burden for A&D is in high income-countries (HICs) but they produced 90.5 % of primary studies and 99.5% participants. Only 14.2% of the dietary risk burden is in HICs but they produced 80.5% of primary studies and 98.1% participants.

Conclusions: This study demonstrates the significant imbalance of research heavily weighted towards HICs. More initiatives and collaborations are required to address this inequality and promote studies in LMICs for future global health research to address the disparity.

Patient or healthcare consumer involvement: This study did not involve healthcare consumers.

Does 'context' influence the conclusions of equity-focused systematic reviews? A methodology study

Magwood O[1], Pottie K[2]

[1] Bruyere Research Institute, Canada; [2] University of Ottawa, Canada

Background: Equity-focused systematic reviews of effectiveness provide decision-makers with the best available evidence on interventions for disadvantaged populations. Current guidance for conducting equity-focused systematic reviews acknowledges that the context in which one lives contributes to inequities. Unfortunately, there is a lack of guidance on when and how to report contextual details in these circumstances. It remains unknown how current review authors of equity-focused reviews identify and use contextual data in their reviews and what impact this has on their findings and conclusions.

Objectives: To describe the extent to which contextual factors are considered in equity-focused systematic reviews, describe the methods used, and assess the implications of their context-related findings on conclusions (implications, practice and research).

Methods: A methodology study examining the evidence on methodological aspects of equity-focused systematic reviews. We systematically identify equity-focused systematic reviews through database searching and reference searching. A mixed methods synthesis approach that incorporates quantitative and qualitative elements as appropriate guides the synthesis of results.

Results: Context includes geographical, epidemiological, socio-cultural, socio-economic, ethical, political and legal domains. The inclusion of contextual information may impact the conclusions of equity-focused reviews and is dependent on the author's decision to apply a multiple-context versus single-context approach. Stakeholder engagement represents one strategy to determine whether a narrow or broad approach should be used to ensure that systematic review findings are useful for their intended audience.

Conclusions: This study has implications for systematic review authors, clinical guideline developers and policy-makers. It may catalyze a discussion on improved reporting of contextual domains both in equity-focused systematic reviews and primary studies to improve their credibility, transferability, and implications for decision-makers. Such improved understanding of methodological challenges may lead to improved reporting, analyses and conclusions for topics and populations with concerns of equity, and downstream may lead to better guidelines and improved professional practice to improve health equity of disadvantaged populations.

Patient or healthcare consumer involvement: Not applicable.

Effective methods used in community sanitation and hygiene promotion in the developing world: a scoping review

Nzioki JM[1]

[1] School of Public Health, Jomo Kenyatta University of Agriculture and Technology, Nairobi

Background: World Health Organization data on the burden of disease shows that approximately 3.1% of deaths (1.7 million) and 3.7% (54.2 million) of disability-adjusted-life-years (DALYs) worldwide are attributable to unsafe water, sanitation and hygiene. In Africa and low- and middle-income countries in South East Asia 4% to 8% of all disease burdens are attributable to poor hygiene and sanitation. Over 99.8% of all deaths in the developing world are attributable to the same factors and 90% are deaths of children under five years. Studies have shown that more than 2 billion people lack access to adequate sanitation and hygiene and increased access to sanitation and better hygienic practices have significant positive health and other impacts. Development agencies and governments have developed and adopted several interventions to help address the problem of poor hygiene and sanitation.

Objectives: To search, document and provide a synthesis of effective methods used in promoting hygiene and sanitation in communities across the developing world.

Methods: This was a scoping review. We reviewed a total of 42 evaluation surveys. We collected data from 8 out of 42 papers which met the inclusion criteria to be considered into the data set.

Results: Results indicated that social mobilization and community participation methods which include community-led total sanitation (CLTS), participatory hygiene and sanitation transformation (PHAST) are effective methods in promoting community hygiene and sanitation. Although children's hygiene and sanitation training (CHAST) and sanitation marketing have been implemented to improve hygiene and sanitation, currently no evaluation studies have been conducted to evaluate the effectiveness of these approaches.

Conclusions: To achieve access to adequate sanitation and hygiene for all, and to end open defecation and reduce the burden of diarrheal diseases by the year 2030 (Target 6.2 of the Sustainable Development Goals), governments and development agencies working in the developing world need to fast track adoption and scaling up of CLTS and PHAST. CHAST and sanitation marketing strategies should also be embraced because they are grounded on sound scientific principles and have potential to improve hygiene and sanitation. However, there is need to carry out evaluation surveys to measure their effectiveness in improving sanitation and hygiene in the developing world.

Patient or healthcare consumer involvement: None.

Effects of interventions to enhance the self-reliance of refugees and internally displaced persons: a systematic review

Saad Z[1], Ismail H[2], Akl E[1], Kashash R[3], Lotfi T[1]

[1] The Global Evidence Synthesis Initiative (GESI) Secretariat, American University of Beirut (AUB), Lebanon; [2] International Committee of Red Cross, Lebanon; [3] Palestine Red Crescent Society, Lebanon

Background: The frequency of man-made disasters such as wars has been increasing for the past 30 years. By end of 2018, almost 70.8 million individuals were forcibly displaced worldwide as a result of persecution, conflict, violence, or human rights violations. Aid professionals are trying to identify interventions to encourage self-reliance among refugees and internally displaced people (IDPs) living in various environments. Self-reliance is commonly defined as the ability of refugees to live independently from humanitarian assistance.

Objectives: This systematic review aims to 1) identify evidence related to designing interventions that enhance the self-reliance of refugees and internally displaced people; 2) assess the effects of these interventions on self-reliance; 3) explore the barriers and facilitators of implementing these interventions; and 4) engage stakeholders in the review process.

Methods: Our inclusion criteria are the following: 1) quantitative or qualitative design, reports and case studies; 2) population includes refugees and internally displaced people living in or outside camps worldwide; 3) interventions designed to enhance self-reliance of population with no restriction on the type or duration; and 4) self-reliance as one of the outcomes. We will search seven databases and websites related to humanitarian emergencies. We will conduct study selection and data abstraction in duplicate and independently. Depending on the type of study, we will critically appraise the included studies using the Modified Critical Appraisal Tool (MCAT), the Mixed Methods Appraisal Tool (MMAT), or CASP Critical Appraisal Checklists. We will conduct a narrative synthesis, and a quantitative synthesis when appropriate.

Results: This work is in progress and the results will be ready by the time of the Colloquium. We will describe the identified interventions and report the barriers and facilitators contributing to their success in achieving self-reliance.

Conclusions: Recommendations and challenges will be concluded when the study is completed by the time of the Colloquium to assist aid workers in developing and implementing interventions that intend to enhance self-reliance for refugees.

Patient or healthcare consumer involvement: We will hold a stakeholder engagement meeting and participants in this meeting will include: 1) people who have supported refugees and internally displaced persons in conflict settings in achieving self-reliance; 2) refugees and internally displaced people with experience in self-reliance interventions such as receiving funding to start a business or being employed to achieve financial independence; and 3) funders of such interventions.

Efficacy and security of internet-based cognitive behavioural therapy versus face-to-face cognitive behavioural therapy in patients with mild or moderate depression: a systematic review

Solis Chimoy BDL[1], Valenzuela Castillo E[1], Rivera Torrejón O[2], Huaman Menacho J[3], Villanueva Lopez D[4], Huaman Sánchez K[5]

[1] Association for the Development of Student Research in Health Sciences, Lima, Peru; [2] San Fernando Scientific Society, National University of San Marcos, Lima, Peru; [3] Federico Villarreal University, Lima, Peru; [4] University of Piura, Peru; [5] National Institute of Health, Peru

Background: Depression is a mental health problem with emotional, cognitive, physical and behavioral symptoms that can interfere with a person's daily life. Its prevalence is on the rise, reaching 29.2% in countries with a medium human development index and 6% in low-income countries according to the World Health Survey conducted in 2013. There is a varied group of treatment strategies such as psychotherapy, whose variant, cognitive behavioral therapy unites concepts about emotions, behaviors and thoughts. It can be done in an autonomous way or with the participation of a therapist through the Internet, called Internet-based cognitive behavioral therapy (ICBT), showing effects comparable to the traditional format in primary healthcare settings. It could be an alternative to increase the coverage of psychotherapeutic treatment for depression and reduce the saturation of the public health system, characteristic of low and middle-income countries such as Peru, where more than 10% of the population does not have access to these services.

Objectives: To synthesize the available evidence about the efficacy and safety of ICBT compared to the face-to-face modality of this type of psychotherapy.

Methods: We searched MEDLINE (PubMed), LILACS (BIREME), the Cochrane Library and the Cochrane Central Register of Controlled Trials (CENTRAL) and SCOPUS without date or language filtering. In addition, we reviewed the website clinicaltrials.gov and the references of included clinical trials and relevant systematic reviews on the topic. We will only include randomized controlled trials. The inclusion and exclusion criteria were validated with a psychiatrist expert in the management of depression. In addition, the protocol of this review will be uploaded to an online platform.

Results: This systematic search identified 2192 studies, after eliminating duplicates there were 1485. Two review authors will independently select studies for inclusion, there will be a third review author in case there is no consensus in the process. Rayyan QCRI will be used for the selection of titles and abstracts, The Risk of Bias 2.0 (Cochrane) tool will be used to assess the risk of bias. We are currently in the study selection phase; the final results will be presented at the Colloquium.

Conclusions: This systematic review will provide reliable evidence about effective therapies for psychotherapeutic management of depression based on telemedicine with a focus on primary care, mainly in low- and medium-income countries, the results will form the basis for future cost-effectiveness studies of these interventions.

Patient or healthcare consumer involvement: This study has a direct implication for patients but mainly will be discussed with the thematic experts.

Evidence-based evaluation of safety of genetically modified food: a report on adverse events of carcinogenesis

Shen C[1], Yin X[2], Jiao B[3], Li J[1], Jia P[4], Zhang X[1], Ren J[2], Cheng X[2], Li X[1], Robinson N[5], Fei Y[1], Liu J[1]

[1] Center for Evidence-Based Chinese Medicine, School of Chinese Medicine, Beijing University of Chinese Medicine, China; [2] School of Chinese Materia Medica, Beijing University of Chinese Medicine, China; [3] School of Chinese Medicine, Beijing University of Chinese Medicine, China; [4] School of Life Sciences, Beijing University of Chinese Medicine, China; [5] School of Health and Social Care, London South Bank University, UK

Objective: To assess the incidence of adverse effects/events of genetically modified (GM) food consumption by an evidence-based evaluation of the safety of GM food.

Methods: We searched seven databases from 1 January 1983 to 8 October 2019. We included in vivo animal studies and human studies in multiple languages. We included animal and human participants, with no prespecified limitations such as age, population, species/race or health status. However, we excluded studies that focused on the effects of GM food on secondary or multilevel consumers in the food chain where GM food was only consumed by primary consumers. Interventions/exposures to the genetically modified animal/plant/microorganism food included referred to GM foods, GM food ingredients and GM feed, regardless of their dosage, duration, or whether the food was prepared or has been approved for marketing.

This study focuses on the incidence of adverse effects/events of GM food consumption and is ongoing. Data synthesis on adverse events has not been completed yet. This abstract focuses on adverse events on carcinogenesis. Researchers independently review the retrieved articles by titles and abstracts and the full text to identify the studies meeting eligibility criteria and independently extracted data from the included studies according to a predesignated extraction table. We resolved disagreement by consensus and, if necessary, arbitration by another researcher. We performed statistical analyses in Microsoft Excel 2010 and SPSS 20.0.

Results: Of the 9328 citations, we included 173 articles with 22 kinds of GM food after reviewing the full text of 432 articles. However, no human clinical study met the inclusion criteria. Finally, only two mouse/rat feeding studies have been reported to trigger cancer/tumor. Seralini GE 2012, which has been retracted but republished, did a long-term toxicity study on a Roundup-tolerant GM maize (from 11% in the diet), cultivated with or without Roundup, and Roundup alone for two years in rats. It found that females in the treatment groups almost always developed large mammary tumors more often than and before controls. As for males, four times more large palpable tumors than controls were presented which emerged up to 600 days earlier. Velimirov A (2008) revealed a life term study where mice in the three groups were fed with transgenic maize NK603xMON810 (from 33.0% in the diet), control isoline diet and GM free Austrian corn reference diet respectively. The survival rate was not significantly different while cancer (leucosis) was the common cause of death.

Conclusion: A majority of studies failed to detect adverse events of carcinogenesis while animal studies occupy the lowest hierarchy of evidence. Further clinical studies such as cohort studies are still warranted.

Identifying ineffective interventions from Cochrane Reviews through the interpretation of the main 'Summary of findings' table

Franco JVA[1], Vergara-Merino L[2], Verdejo C[2], Schuller-Martínez B[2], Kirmayr M[2], Kopitowski K[3], Madrid E[4], Meza N[4], Pérez-Bracchiglione J[4], Muscat Meng D[5], Johansson M[6]

[1] Research Department, Cochrane Associate Centre, Instituto Universitario Hospital Italiano, Argentina; [2] Escuela de Medicina, Universidad de Valparaíso, Centro Asociado Cochrane Chile, Chile; [3] Family and Community Medicine Division, Hospital Italiano de Buenos Aires, Argentina; [4] Interdisciplinary Centre for Health Studies (CIESAL), Cochrane Associate Centre, School of Medicine, Universidad de Valparaíso, Chile; [5] Cochrane Sustainable Healthcare, Denmark; [6] Cochrane Sustainable Healthcare, Sweden

Background: Cochrane Reviews (CR) of interventions summarise the evidence on the comparative effectiveness of interventions in health care by collating all available randomized trials and performing, when possible, meta-analysis. CR have highlighted the importance of the implementation of highly effective interventions (e.g. corticosteroids for women at risk of premature birth), however, their impact on de-implementation is unclear. Currently, there is not a comprehensive summary of recent high-quality research on interventions that might be candidates for de-implementation due to their lack of efficacy or their associated harms. Furthermore, more information is needed in terms of how those interventions can be defined (via the author's conclusions or 'Summary of findings' [SoF] tables).

Objectives: To identify CR in which the main SoF table described moderate-to-high quality evidence of no effect or suggested harm.

Methods: We analysed CR interventions with SoF tables. We then identified those reviews in which: the estimates of at least one outcomes with moderate or high quality of evidence in the main SoF table included no effect (e.g. significant difference thresholds for dichotomous and continuous outcomes), and/or the estimates of adverse events suggested harm (e.g. beyond significant difference thresholds for dichotomous and continuous outcomes). We presented a descriptive summary of the characteristics of the patients, interventions, comparisons and outcomes (PICO) questions covered by these reviews. We also analysed the correspondence between the author's conclusions and confidence in the outcomes of SoF tables.

Results: We screened 597 CR and found 36 reviews (see Figure 1) with moderate or high-quality evidence of no beneficial and/or harmful effect for at least one outcome in the main SoF table. In five of these reviews, authors did not conclude that the interventions should potentially be de-implemented; in two reviews this was based on outcomes not presented in the main SoF table and in three reviews on uncertainties in the body of evidence. Our main limitations included that we did not assess secondary SoF tables and the variable report of SoF tables (not systematically assessing harms or clinically relevant outcomes). Additionally, since we included reviews with non-statistically significant measures, we might have missed reviews that reported statically significant but clinically irrelevant effects (that would also be candidates for de-implementation). We could not screen reviews systematically because thresholds for clinical significance across reviews vary substantially.

Conclusions: We provide a structured process to identify interventions that may be considered ineffective and/or harmful. Whereas we found limitations in our screening process, we believe that this process is sustainable and can serve as input for a knowledge translation (KT) strategy for the Sustainable Healthcare Field. We will pilot a Knowledge Translation package for these 36 reviews and present it to relevant stakeholders.

Patient or healthcare consumer involvement: None.

Additional file: Figure

International comparison of the job satisfaction of nurses and gross domestic product (GDP): a systematic review

Chiang L[1], Chen S[2]

[1] School of Nursing, National Defense Medical Center, Taiwan; [2] School of International Business, College of Management, National Taiwan University, Taiwan

Background: Job satisfaction of nurses is the degree to which a registered nurse is satisfied with the individual's job performance, job adjustment, and the practice environment in the nursing profession. Faced with an aging global population, the World Health Organization estimates that human resources for health care will be confronted with a considerable threat. Nurses have gradually migrated to work in richer countries, which has resulted in a shortage of manpower in poor countries.

Objectives: To conduct a systematic review of the global job satisfaction of nurses around the world in this decade, and the relationship with per capita gross domestic product (GDP) in 2019.

Methods: We found a total of 100 quantitative research results on the job satisfaction of nurses in various countries around the world. The means of job satisfaction of nurses were standardized for global comparison, and correlation with the per capita Gross Domestic Product (GDP) of each country was analysed.

Results: This systematic review and meta-analysis of nursing job satisfaction in the previous ten years (from 2009 to 2019) over 38 countries indicated that overall average job satisfaction of nurses is 64.185 (standard deviation (SD) 8.29) among a total of 71,429 nurses. Correlation with the national per capita GDP showed a coefficient 0.374 (P = 0.021). Deleting higher scores in Laos, 37 countries also showed a higher significant correlation coefficient of 0.440 (P = 0.006). The results indicated that the higher the national per capita GDP, the higher the job satisfaction of nurses in the country. Nineteen Taiwan studies showed an overall average job satisfaction of 67.82.

Conclusions: Nurses around the world showed medium job satisfaction, with nurses in Taiwan showing a slightly higher average satisfaction. This study found that nurses with higher overall economic status in the country have higher job satisfaction. Improvement in the country's overall economic strength, development of health policies, and a high-quality working environment for nurses can increase their job satisfaction, increase their willingness to stay and reduce turnover rates, thereby ensuring the universal health coverage and quality of health care for people internationally.

Patient or healthcare consumer involvement: No

Additional files: Figure 1; Figure 2

Interventions to optimize the care cascade for migrants living with chronic viral hepatitis: a mixed methods systematic review

Hung J[1], Magwood O[2], Welch V[3], Pottie K[4]

[1] Department of Family Medicine, McMaster University; School of Epidemiology and Public Health, University of Ottawa, Canada; [2] C.T. Lamont Primary Health Care Research Centre, Bruyère Research Institute, Canada; [3] The Campbell Collaboration, Oslo, Norway; Centre for Global Health, Bruyère Research Institute; School of Epidemiology and Public Health, University of Ottawa, Canada; [4] Centre for Global Health, Bruyère Research Institute; Department of Family Medicine and School of Epidemiology and Public Health, University of Ottawa, Canada

Background: Over 350,000 deaths are attributed to viral hepatitis infection annually, a high proportion of which are among migrants from high endemic regions. Socio-economic barriers and the lack of an effective viral hepatitis model of care for migrants contribute to hepatitis mortality and morbidity. Given the multi-faceted nature of the hepatitis care cascade for migrants and the dynamic change of migrant care policy, there is a need to synthesize evidence on the effectiveness of the hepatitis care cascade, as well as barriers and facilitators at each step along the continuum.

Objectives: (1) To assess the effectiveness of interventions along the hepatitis care cascade on screening, referral, treatment uptake and completion among migrant populations; (2) To identify barriers and facilitators along the care cascade for migrants living with viral hepatitis.

Methods: We searched five databases and grey literature published until January 2020. We conducted a mixed-method synthesis by developing a conceptual framework (Fig.1), undertaking qualitative comparative analysis (QCA), and meta-analysing outcomes where appropriate. The certainty of the evidence was assessed using the GRADE and GRADE CERQual approaches.

Results: We screened 1448 studies by title and abstract, followed by full-text. We Included 52 studies providing information for our QCA and 84 quantitative studies measuring the effect of interventions. Our preliminary results highlighted the importance of access to screening and treatment. Facilitators included trust in the provider, and peer and navigation support. Language differences, health literacy, hepatitis-related stigma, and a lack of migrant resources were identified as barriers to successful implementation of the hepatitis care cascade. The estimated yield of hepatitis C screening was 31.20 cases per 1000 screened (95% CI 25.65 to 37.86; I^2 = 35.83%). The comparison of the screening uptake of physician-delivered clinic-based programs versus community outreach programs was 80.30% versus 20.51% (95% CI 57.59% to 92.44% vs. 9.66% to 38.37%; P < 0.001) (Fig.2). Effects of interventions on linkage to viral hepatitis treatment are uncertain due to the variation between the results of the studies. Interventions that engaged parents, community members, and healthcare providers might be associated with successful implementation. Evidence is insufficient to show whether sustained viral suppression is affected by specific features of interventions.

Conclusions: Our study provides evidence on effective interventions to optimize screening uptake along the care cascade for migrants living with chronic viral hepatitis. We identified the facilitators and barriers of successful implementation of viral hepatitis care for migrants from the perspective of healthcare providers and patients to inform future implications.

Patient or healthcare consumer involvement: The input of the clinicians specialized in migrant health and hepatitis and the patients had been incorporated into our outcome choice process along with data interpretation.

Additional files: Figures

Monitoring process barriers and enablers towards universal health coverage within the sustainable development goals

Doshmangir L[1], Derakhshani N[1], Ahmadi A[2], Gordeev VS[3]

[1] Health Services Management Research Center, School of Management and Medical Informatics, Tabriz University of Medical Sciences, Tabriz, Iran; [2] Knowledge Utilization Research Centre, University Research & Development Center, Tehran University of Medical Sciences, Tehran, Iran; [3] Pragmatic Clinical Trials Unit, Centre for Primary Care and Public Health, Queen Mary University of London, London, United Kingdom

Background: This study builds on previous successes of using tracer indicators in tracking progress towards universal health coverage (UHC) and complements them by offering a more detailed framework that would allow identifying potential factors that impede or advance such progress. This tool was designed accounting for possibly available data in low- and middle-income counties.

Methods: We carried out a systematic review of relevant studies using PubMed, ISI Web of Science, Embase, Scopus and ProQuest databases with no time restriction. The search was complemented by a scoping review of grey literature, using the World Bank and World Health Organization official reports depositories. Inductive content analysis identified determinants influencing the progress towards UHC and its relevant indicators. We explored the conceptual proximity between indicators and categorized themes through three focus group discussion with 18 experts in UHC. Finally, we converted a comprehensive list of indicators into an assessment tool and refined it following three consecutive expert panel discussions and two rounds of email surveys.

Results: We extracted 416 themes (including indicators and determinants factors) from 170 eligible articles and documents. Based on conceptual proximity, the number of factors was reduced to 119. These were grouped into seven domains: social infrastructure and social sustainability, financial and economic infrastructures, population health status, service delivery, coverage, stewardship/governance, and global movements. The final assessment tool included 20 identified subcategories and 88 relevant indicators.

Conclusions: Despite various challenges, public–private partnerships in primary health care can facilitate access to healthcare services, especially in remote areas. Governments should consider long-term plans and sustainable policies to start public–private partnerships in primary health care and should not ignore local needs and context.

Patient and healthcare consumer involvement: Patients and healthcare consumers were not involved in this research.

Partnership for research and capacity building in evidence-informed decision-making to improve occupational health

Gong W[1], Morata T[2]

[1] National Institute for Occupational Safety and Health (NIOSH), Centers for Disease Control and Prevention (CDC); Jiangsu Provincial Center for Disease Control and Prevention, USA; China; [2] National Institute for Occupational Safety and Health (NIOSH), Centers for Disease Control and Prevention (CDC), USA

Background: Traditionally, occupational health programs have been informed by research evidence, professional experience, and stakeholder preferences. Time and lack of skills are challenges to accessing recent evidence and updating programs designed to promote health and prevent disease. The high rates of work-related health conditions and injury indicate that the effectiveness of preventive programs has been limited. This suggests that that rigorous evaluation is not being conducted to verify impact. Internationally, support available to build capacity for evidence-informed decision-making among the public health workforce varies widely. There is a need to develop capacity in intervention effectiveness research and systematic reviews in the area of occupational health.

Objectives: To describe a strategy to promote evidence-informed decision-making to improve occupational health.

Results: The participation of Chinese authors in Cochrane has been focused in the areas of clinical interventions and medical devices. In recent years, the Jiangsu Provincial Center for Disease Control and Prevention identified the need to adopt intervention effectiveness research to evaluate and expand activities directed to the prevention of leading work-related health conditions. One of these conditions is work-related hearing loss. Two approaches have been taken towards building capacity in evidence-informed decision-making to improve occupational health interventions:

1) With the availability of new technologies to measure the attenuation provided by hearing protection devices, field studies were conducted to examine their effectiveness in offering needed protection against noise effects (Liu et al., Ear Hear. 2019; Gong et al., Int J Audiol. 2019).

2) The work described above facilitated a research partnership with the National Institute for Occupational Safety and Health, of the Centers for Disease Control and Prevention in the US. That consultation led to the opportunity of building capacity for participation in Cochrane Reviews in the area of occupational health. The effort included training of staff, the translation of reviews into Chinese, and co-authorship of a planned review.

Conclusions: The co-ordination of these efforts has the potential to:

- advance the knowledge and stimulate the practice of evaluating the effectiveness of interventions;
- expand the dissemination of the results to inform decision-making; and
- help identify approaches for international efforts to improve quality of research and the uptake of evidence into practice.

Ultimately, we expect it will contribute to Cochrane's Strategy to 2020 efforts towards building a diverse, inclusive and transparent international organization.

Disclaimer: The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention.

Prevalence of depression and depressive symptoms among postgraduate students: a systematic review and meta-analysis

Guo L[1], Li Z[1], Yang K[1], Fan H[2]

[1] School of Basic Medicine, Lanzhou University, China; [2] School of Education, Bohai University, China

Background: Suicide events were frequently reported among postgraduate students, which has aroused widespread concern about their mental health, especially depression or depression symptoms. Although researchers conducted numerous studies on the prevalence of depression in postgraduates, the effect sizes ranged from 4.1% to 67.9%.

Objectives: To summarize the information of publications on depression, estimate the prevalence of depression with a meta-analysis and identify factors associated with depression among postgraduates with systematic review.

Methods: The meta-analysis was adopted to estimate the prevalence of depression among postgraduates, and the systematic review was adapted to synthesize related factors. We retrieved the primary studies from English electronic databases, including articles published from 1980 to 2019. We used R 3.6.1, CMA V2 and SPSS 22 to perform the statistical analyses.

Results: We included 16 primary studies (with 17 reports and 11,386 individuals) in the meta-analysis, and 15 primary studies in the systematic review (totalling 7800 individuals). The random-effects meta-analysis showed that: (1) The prevalence of mild depression was 31% (95% CI 22% to 40%), moderate depression was 18% (95% CI 15% to 23%), and severe depression was 6% (95% CI 4% to 10%); (2) PhD students have a higher prevalence of depression than masters except in severe depression; The effect size was influenced by the measurements and sampling methods. (3) There was no difference both in genders and regions.

In addition, the prevalence of depression among postgraduates has increased since 2003; this change was not statistically significant. The systematic review showed that the factors related to depression in postgraduates included suicidal behaviour, negative emotion, substance abuse, and eating disorder.

Conclusions: Postgraduate students, especially PhD students, have a high risk of depression. The prevalence of depression has been increasing, but this change was not statistically significant.

Patient or healthcare consumer involvement: Postgraduate students.

Additional files: Tables; Figures

The association between dairy intake and risk of mortality: protocol for a dose–response meta-analysis of prospective cohort studies

Hui X[1], Li J[1], Lu ZX[2], Dai H[2], Ma XY[3], Li XJ[3], He GG[3], Feng XW[3], Rong RX[2], Yan PJ[4], Yao L[5], Yang KH[6]

[1] School of Public Health, Lanzhou University, Lanzhou, China; [2] First Clinical Medical College, Lanzhou University, Lanzhou, China; [3] Second Clinical Medical College, Lanzhou University, Lanzhou, China; [4] Institution of Clinical Research and Evidence Based Medicine, Gansu Province People's Hospital, Lanzhou, China; [5] Health Research Methodology/Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, Ontario, Canada; [6] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China

Background: Dairy products are globally consumed and recommended in dietary guidelines around the world, but their effect on mortality is still controversial. Limited sample size and inconsistent control of adjustment factors make the evidence far from robust. Moreover, the use of relative risk (RR) as a merge indicator might result in information loss of survival analyses in prospective cohort studies.

Objectives: To examine the association of dairy intake and risk of mortality.

Methods: The target question was being assessed at PROSPERO: the association between dairy intake and risk of mortality: a dose–response meta-analysis of prospective cohort studies (CRD161807). A comprehensive search will be performed in the following databases without time limit for publication: PubMed and Embase. Prospective cohort studies of the association between dairy intake and risk of mortality will be included. Two review authors will independently select citations for inclusion, extract data and assess the methodological quality. The third review author will resolve any disagreement. Stata 15 will be used for data synthesis and exploring the dose–response relationship between dairy intake and risk of mortality. We will use a piecewise linear regression model and a restricted cubic spline model for linear and nonlinear trend estimation respectively. We will undertake subgroup analysis on gender, types of dairy intake and risk of mortality. Also, sensitivity analysis will be performed to evaluate the stability of the research results.

Results: This study is ongoing and will be submitted to a peer-reviewed journal for publication.

Patient or healthcare consumer involvement: There was no patient or healthcare consumer in this project.

The effectiveness of technology delivered mental health interventions among forcibly displaced migrants: protocol for an equity-focused systematic review

Saad A[1], Magwood O[2], Alkhateeb Q[3], Pottie K[2]

[1] School of Epidemiology and Public Health, University of Ottawa, Canada; [2] C.T. Lamont Primary Healthcare Centre, Bruyere Research Institute, Canada; [3] Department of Medicine, Faculty of Medicine, Western University, Canada

Background: Many migrant communities in Canada have been forcibly displaced from their homes and lands due to nature or man-made disasters. These disadvantaged populations experience trauma and violence prior to and during their migration journey, which increases the odds of developing trauma-specific mental health disorders, such as depression and post-traumatic stress disorder PTSD. Technology-delivered interventions have emerged as a possible solution to reach patients that would otherwise not have access to mental health treatment due to personal or system-level barriers.

Objectives: To systematically review best available evidence on the effectiveness and implication on health equity of technology-delivered interventions in the management of mental health conditions among forcibly displaced migrants.

Methods: We plan to conduct an equity-focused systematic review conforming to the PRISMA-E checklist. We will develop a comprehensive search strategy in consultation with a health sciences librarian and will use this strategy to systematically search multiple online databases. Two authors will independently screen the records yielded by our search. We will extract relevant data from included studies and analyze them as appropriate. Cochrane's Risk of Bias 2 tool will be used to critically appraise the quality of primary studies, and GRADE methodology will be used to assess certainty of evidence. Moreover, to examine possible implications on health equity, we will stratify reported effect estimates by migrants' place of residence, race and ethnicity, occupation, gender, religion, education, socioeconomic status, social capital, and any characteristics associated with discrimination.

Relevance and added value: To the best of our knowledge, this systematic review is the first of its kind to provide a robust and equity-focused analysis of evidence on technology-delivered mental health interventions among forcibly displaced migrants. Moreover, the scope of this review conforms to the Colloquium's research focus, as it examines the use of reachable and interactive interventions in improving the experiences and health outcomes of one of the most disadvantaged populations around the world.

The landscape of clinical trial activity focusing on Indigenous health in Australia from 2008 to 2018

Hunter KE[1], Xu G[1], Modi D[1], Askie LM[1], Jamieson L[2], Brown A[3], Seidler AL[1]

[1] National Health and Medical Research Council (NHMRC) Clinical Trials Centre, University of Sydney, Australia; [2] Australian Research Centre for Population Oral Health, The University of Adelaide, Australia; [3] South Australian Health and Medical Research Institute, Australia

Background: There are major disparities in health outcomes between Indigenous and non-Indigenous Australians. To address this, it is vital to understand key health priorities and knowledge gaps in the current landscape of Indigenous trial activity.

Objectives: To provide an overview of trial activity in Australia focusing on Indigenous health, and compare this to overall Australian trial activity.

Methods: We extracted data from all trials registered between 2008 and 2018 on the Australian New Zealand Clinical Trials Registry or ClinicalTrials.gov, which only recruited participants in Australia. We identified trials with an Indigenous focus by searching titles, eligibility criteria, study summary, intervention description and ethics committee name for relevant terms such as 'Indigenous', 'Aboriginal' and 'Torres Strait Islander'. We compared Indigenous versus non-Indigenous trials and Australian trials overall by sample size, conditions studied, intervention type, study design, sponsorship and funding.

Results: Of the 9206 included trials, 139 (1.5%) focused on Indigenous health. The three disease areas that had the greatest number of Indigenous trials were 'Public Health' (n = 69, 50%), 'Mental Health' (n = 35, 25%) and 'Cardiovascular' (n = 25, 18%) (see Figure 1). Compared to other Australian trials, Indigenous trials more frequently studied ear conditions (odds ratio (OR) 16.47, 95% confidence interval (CI) 8.43 to 29.99), public health (OR 4.87, 95% CI 3.65 to 6.41) and infection (OR 2.51, 95% CI 1.53 to 3.89), and were more likely to focus on early detection/screening (OR 3.57, 95% CI 2.10 to 5.70) and preventive interventions (OR 2.24, 95% CI 1.61 to 3.08) rather than treatment (OR 0.40, 95% CI 0.30 to 0.52). The majority of Indigenous trials were randomized (n = 97, 70%) and this proportion was similar for other Australian trials (n = 6763, 76%). Indigenous trials were less likely to be blinded (OR 1.72, 95% CI 1.20 to 2.49). Only 16 (12%) Indigenous trials had industry involvement compared to 2271 (25%) of other Australian trials (OR 2.52, 95% CI 1.54 to 4.43).

Conclusions: Trials with an Indigenous health focus differ from other Australian trials in terms of health conditions studied, intervention focus, blinding and industry involvement. Understanding these differences can help inform research prioritization to address the high burden of disease in Indigenous Australians. Relative to population size and burden of disease, the number of trials focusing on Indigenous health is low, and therefore limited Indigenous trial data are available for incorporation into systematic reviews and clinical guidelines. Trial registries are a freely available and valuable resource for evaluating trends in trial activity, which can be used to inform future research and ensure it addresses diverse and vulnerable populations such as Indigenous Australians.

Patient or healthcare consumer involvement: This research was conducted in collaboration with an Indigenous researcher (Alex Brown).

Figure 1. Top 14 conditions by number of trials, for Australian interventional trials with Indigenous health focus, registered 2008-2018
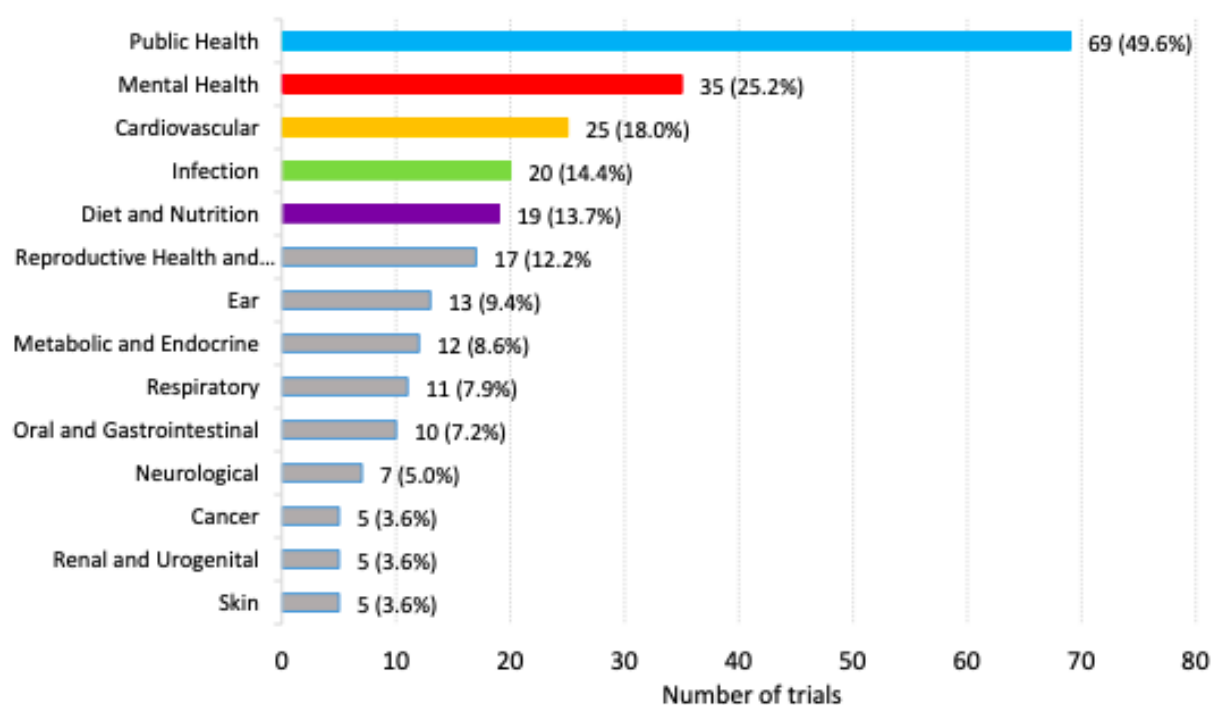
EVIDENCE ADVOCACY

An initiative to provide evidence-based support for judicial decision on healthcare in Brazil

Riera R[1], Pachito D[2], Bagattini Â[2]

[1] Centre of Health Technology Assessment (HTA), Hospital Sírio-Libanês, Universidade Federal de São Paulo, Brazil; [2] Centre of Health Technology Assessment (HTA), Hospital Sírio-Libanês, Brazil

Background: Between 2008 and 2017, judicial demands related to health care increased 130% in Brazil [1]. In 2016, the 10 most expensive drugs that were legally demanded comprised 91% of the resources guaranteed for the acquisition of technologies by the Brazilian Government. In this scenario, it is necessary to adopt strategies to enhance the judicial process, ensuring that the jurisdictional decision is informed by the best evidence.

Objectives: To present the implementation and initial results of an initiative that comprises a set of strategies to provide evidence-based support during the judicial process on healthcare in Brazil.

Methods: Descriptive study conducted at the Centre of Health Technology Assessment (HTA), Hospital Sírio-Libanês (São Paulo, Brazil).

Results: Through a broad project of the Brazilian Ministry of Health, formally named PROADI-SUS, a partnership has been established between Hospital Sírio-Libanês and the National Council of Justice since 2016. This partnership was consolidated with the implementation of an initiative composed of a set of seven strategies aimed at supporting the judicial process in health care. The strategies comprise:

1) three editions of an online course (300 hours) for 403 health and law professionals addressing topics related to evidence-based medicine (EBM) and HTA;
2) seven workshops, in all Brazilian geographic regions, for training 110 health professionals who provide technical support to the judiciary (NATJus), in the preparation of evidence reports on the efficacy and safety of healthcare technologies;
3) management and technical support to validate the evidence synthesis produced about the most legally requested technologies (69 drugs, devices or procedures to date) in Brazil;
4) creation and content maintenance of an online open access repository (e-natjus platform) for consulting evidence reports (487 to date) and evidence synthesis (69 to date) produced along this initiative;
5) online tutoring for the NATJus team, via the e-natjus platform, to support scientific questions that arise during the production of short evidence reports;
6) blog with 3 to 4 weekly posts focused on events and supporting material related to judicial process on healthcare, EBM and HTA (www.redenatjus.com); and
7) three editions of a 2-day symposium addressing judicialization issues and promoting debates among law and health professionals, health managers, consumers and journalists.

Conclusion: We hope that the initiative, by the adoption of EBM practices, will contribute to 1) enhancing the judicial process on healthcare in Brazil so the best possible option could be available within the context of a public and universal health system, 2) the sustainability of the Brazilian public health system.

Patient or healthcare consumer: Conducting judicial process under the premises of EBM may contribute to the population receiving the best possible healthcare in a scenario of a public and universal system with scarce resources.

Support: Brazilian Ministry of Health / PROADI-SUS, Hospital Sírio-Libanês.

Associations of hand-washing frequency with the incidence of illness: a systematic review and meta-analysis

Xun Y[1], Shi Q[2], Yang N[2], Li Y[3], Yang N[1], Si W[3], Liu X[3], Chen Y[1]

[1] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, Lanzhou, China; [2] Children's Hospital, Chongqing Medical University, China; [3] School of Public Health, Lanzhou University, Lanzhou, China

Background: Hand hygiene is one of the most effective ways to prevent the spread of disease. In health education, people are advised to wash their hands frequently. But there is a lack of recommendations on the frequency of handwashing, and the current research conclusions about the frequency of handwashing and disease prevention effects are not uniform.

Objectives: To explore the relationship between hand washing times and disease prevention effects by a systematic review and meta-analysis.

Methods: We searched PubMed, Cochrane Library, Web of Science, Embase, China National Knowledge Infrastructure, Wanfang and China Biology Medicine disc without search date and language restriction using "handwashing", "hand hygiene", "prevent*", "frequency" and "times" etc. We used Review Manager version 5.3. software to measure intervention effect. We performed subgroup analysis on study design. We assessed the quality of evidence with respect to each outcome indicator using the GRADE approach.

Results: We included eight studies. Meta-analysis results showed that there was no statistical significance in the effect of disease prevention on whether handwashing was more than four times/day (odds ratio (OR) 0.61, 95% confidence interval (CI) 0.37 to 1.01). The results of a case-control study showed that compared with handwashing ≤ 4 times/day, handwashing 5 to 10 times/day and handwashing > 10 times/day could prevent disease infection, and the results were respectively OR 0.75, (95% CI 0.63 to 0.91) and OR 0.65 (95% CI 0.53 to 0.80). There was no statistical significance between handwashing 10 times/day and 5 to 10 times/day (OR 0.86, 95% CI 0.70 to 1.06). Compared with handwashing ≤ 10 times/day, handwashing > 10 times/day was a protective factor against infection (OR = 0.59, 95% CI 0.36 to 0.97). Patients assigned to the intensive hand-washing intervention group washed their hands more frequently compared with the control group (seven vs four times a day) and developed fewer episodes of diarrheal illness (weighted mean difference (WMD) -1.68, 95% CI -1.93 to -1.43). All the above the quality of evidence was low.

Conclusions: The higher the frequency of handwashing, the better the effect of disease prevention, but so far there is no high-quality evidence indicating the best range of handwashing times for disease prevention. In the future, large-scale trials will be required to explore this. It is necessary for health workers to increase publicity on hand hygiene education, which is a low-cost and high-efficiency health measure.

Patient or healthcare consumer involvement: None.

Development of an evidence-based curriculum using Cochrane systematic reviews and tools in a learning healthcare system

Gupta A[1], Ironside KR[1], Koster MA[1], Sim JJ[1]

[1] Kaiser Permanente, USA

Background: Evidence-based medicine (EBM) education is an integral component of a rapid-learning system. With access to electronic health records and integrated data systems, Kaiser Permanente physicians and researchers are positioned to take advantage of unique opportunities to conduct research, develop best practices, and improve healthcare quality. However, access to data alone is not enough. Researchers and physicians also need guidance on how to design optimal studies, critically appraise results, and interpret findings. Cochrane tools are a critical resource in developing these skills and supporting the continuous development of an evidence-based, rapid-learning health system.

Objectives: To describe an EBM mini-course designed to train physicians in critical appraisal and EBM methodology.

Methods: The objectives of the EBM mini-course are to:

1) provide physicians with skills necessary to conduct high-quality evidence-based research;
2) encourage the incorporation of EBM principles into daily clinical practice; and
3) accelerate implementation of best practices across the organization.

To develop a curriculum that would achieve these goals, we conducted a comprehensive search of the literature to identify key components of an effective EBM curriculum. We also reviewed online EBM resources, including the Cochrane Interactive Learning Modules.

Results: Based on the results of research findings, we developed a six-session, interactive EBM course curriculum. The curriculum

1) introduced participants to the 5As of EBM methodology;
2) provided training on how to conduct a comprehensive literature search using public and Kaiser Permanente research resources;
3) included instruction on how to critically appraise research publications; and
4) provided training on interpreting a systematic review or meta-analysis.

The course included lectures, group exercises, and knowledge checks. The Cochrane Handbook and Cochrane risk-of-bias tool served as reference material for critical appraisal exercises. Cochrane systematic reviews were used as examples and standards. While the course was originally offered only to staff physicians, it was later expanded to nursing researchers, clinical trial managers, research project managers, and staff supporting clinical teams. Using skills gained from the course, class participants are currently in the process of conducting or supporting research in the areas of sepsis, cardiovascular disease, obstetrics, and patient safety.

Conclusions: Access to electronic health records and integrated data systems puts Kaiser Permanente in a unique position to rapidly conduct evidence-based research and spread best practices. Training in EBM methodology, using Cochrane as well as other EBM tools, provides physician and other clinical researchers with the skills needed to efficiently harness the opportunities provided by access to data. Prioritizing evidence-based teaching can accelerate the production of high-quality research and promote rapid learning of best practice.

Patient or healthcare consumer involvement: None.

Enhancing partnerships and collaboration in times of change

Clark E[1], Dobbins M[1]

[1] National Collaborating Centre for Methods and Tools, Canada

Background: In 2019, the government of the Canadian province of Ontario announced major changes to the structure, governance and funding of public health services. Throughout these changes, publicly funded local and regional public health organizations are still expected to provide effective, evidence-informed programming to their communities. The National Collaborating Centre for Methods and Tools (NCCMT) has facilitated collaborations to support Ontario public health organizations through this transition.

Objectives: The NCCMT sought to meaningfully support evidence-informed public health processes in Ontario during a time of system change by fostering partnerships and collaboration with key stakeholders.

Methods: We reached out to current and potential partners, looking beyond the public health sector for a multidisciplinary approach. We conducted a needs assessment survey to inform need for and use of a rapid evidence review repository, which would allow public health practitioners to share and adapt each other's work. Finally, demonstrating the value of public health to policymakers can be inherently challenging; return on investment in public health is often very long term. We partnered with health units in varying capacities to find and synthesize evidence to make the case for public health programs overall.

Results: This initiative has provided several important lessons in developing and maintaining strong partnerships and collaborations. Looking beyond the public health sector can find partners and allies in other disciplines. We also learned that there is a need to establish an infrastructure to support collaboration and sharing of resources. Finally, we learned that big picture questions like demonstrating the value of public health require many different perspectives, inputs and areas of expertise.

Conclusions: Through this initiative, we have developed a multidisciplinary, collaborative approach to supporting evidence-informed public health through times of major restructuring. This approach can be applied to smaller- or large-scale changes such as future changes to public health in Ontario or in other provinces.

Evidence mapping: a tool for synthesis and presentation of evidence

Li Y[1], Li X[1], Lu C[2], Chen N[1], Yang K[2]

[1] School of Public Health, Lanzhou University, China; [2] School of Basic Medical Science, Lanzhou University, China

Background: Evidence mapping is a comprehensive evidence research method that systematically collects, evaluates, and synthesizes existing evidence to clarify research status and gaps, thereby promoting scientific research and decision-making. There is relatively little theoretical and practical research, and the concept definition is inconsistent.

Objectives: To systematically introduce the methodology of evidence mapping production and reporting for researchers to use for reference.

Methods: Literature review and analysis.

Results: At present, the production steps and definitions of the evidence mapping are not uniform. On the whole, the evidence mapping is mainly divided into two types, one is the evidence map and the other is the gap map. The production steps mainly include: develop scope, set exclusion criteria, searching, assessing inclusion, critical appraisal, and producing.

Conclusions: Evidence maps are versatile and deserve the attention of researchers and decision-makers.

Patient or healthcare consumer involvement: Not applicable.

Rede NATJus: a blog providing content on EBM and J-HTA= for professionals involved with judicialization in healthcare in Brazil

Riera R[1], Bagattini Â[2], Pachito D[2]

[1] Centre of Health Technology Assessment, Hospital Sírio-Libanês; Universidade Federal de São Paulo (Unifesp), Brazil; [2] Centre of Health Technology Assessment, Hospital Sírio-Libanês, Brazil

Background: Between 2008 and 2017, the number of health-related lawsuits in Brazil increased 130% accompanied by an increase of 50% in all-cause lawsuits [1]. Several strategies have been implemented to understand the scenario and to identify measures to improve the judicial process in healthcare in Brazil.

Objectives: To present the implementation and the initial accesses data of a blog focused on evidence-based medicine (EBM) and health technology assessment (HTA) content for health and law professionals involved in the judicialization of healthcare in Brazil

Methods: We conducted a descriptive and analytic study at the Centre of Health Technology Assessment (C-HTA), Hospital Sírio-Libanês (HSL, São Paulo – Brazil).

Results: In view of this new judicial scenario, one of the strategies was the creation of the Rede NATJus blog (www.redenatjus.org.br) with EBM/HTA content and aimed at NATJus members. NATJus (Núcleos de Apoio Técnico ao Judiciário) are technical nuclei, composed of health professionals, who offer support to the judiciary during legal proceedings, providing evidence on efficacy and safety of health technologies requested from the government through the courts. The Rede NATJus blog, maintained by C-HTA of Hospital Sírio-Libanês, was officially launched on 3 October 2019 and besides EBM and HTA, addresses issues related to incorporation of technologies by the Brazilian public health system (SUS), health judicialization itself and related courses/events as well. The posts are made available weekly and the most relevant ones have also been sent as newsletters to a mailing list of 4000 users (previously identified or spontaneously registered on the blog's homepage). By 12 March 2020, 76 posts had been published and according to data from Google Analytics, 88% of accesses was from Brazil, 13.4% from USA and 2.1% from Argentina. The main access channels were email (newsletter) 50.8%, direct 21% and organic search 13.2%. Among the five most accessed posts, three addressed judicialization and two addressed training courses in EBM/HTA. It is important to note that in the week before 12 March, the most accessed post was about the special series of Cochrane Reviews about COVID-19 (23.7% of all accesses, http://redenatjus.org.br/cochrane-disponibiliza-colecao-especial-de-revisoes-sistematicas-relacionadas-a-covid-19/).

Conclusions: We hope that the Rede NATJus blog can be a reliable and an easy-to-understand source of information on EBM/HTA for professionals with interest in health judicialization in Brazil.

Patient or healthcare consumer involvement: Rede NATJus can contribute so that the Brazilian judicial decision-making can be more efficacious, providing only technologies that clearly work and are safe to the detriment of those ineffective, harmful or presenting uncertain efficacy/safety. Support: Brazilian Ministry of Health/PROADI-SUS.

Reference:

1. Instituto de Ensino e Pesquisa, INSPER. Judicialização em saúde no Brasil perfil das demandas, causas e propostas. Available from: https://encurtador.com.br/gRT07

US-PCG consumer involvement and outreach efforts

Kumar N[1], Kariyawasam D[2], Hill M[2], Jahanfar S[2], Haas D[1]

[1] Indiana University School of Medicine, USA; [2] Central Michigan University, USA

Background: The US Satellite of the Cochrane Pregnancy and Childbirth Group (US-PCG) was launched in March 2019. One of the main goals of the US-PCG was to have a greater consumer involvement with satellite activities.

Consumer Workshop: Four consumers attended our first half-day workshop and expressed valuable advice and interest in consumer engagement. In particular, we now have a standing relationship with Indianapolis Healthy Start, a Health Department initiative offering education and support services to pregnant women to eliminate disparities and improve infant survival rates. In 2020, we plan to be present at their events to garner more consumer involvement in an expanding area. Moreover, we submitted an NIH R13 grant to hold this workshop annually for consumers and engage them in knowledge translation efforts. **Prioritization:** The US-PCG underwent a priority-setting Delphi process that engaged with both clinicians and consumers. At the end of the process, four out of the top five most prioritized review titles were shared between both groups, and thus we were able to reach consensus on six titles in total. Three have been updated or are in the editorial process. Two will have US-PCG members assisting one of the original authors and the last one was relinquished to us. Of note, we have consumers present on reviews, which helps diversify the author team and emphasize patient-centred outcomes.

Other efforts: We have established a relationship with the Indiana University National Center of Excellence in Women's Health, which holds dinners, lectures, and an annual Women of Influence event that draws over 300 women to talk about current healthcare topics. Furthermore, we have attended or plan to participate in local conferences held by the State and County Health Departments to keep our consumer base growing (e.g. Labor of Love Infant Mortality Summit, Breastfeeding Conference). We also have created two consumer-facing fact sheets on breastfeeding and opioid use during pregnancy for distribution among community settings.

Conclusion: As a newly formed satellite, impact and sustainability are central to the US-PCG's goals and initiatives. Indiana has one of the highest infant and maternal mortality rates in the United States, and we are in a prime position in Indianapolis at the School of Medicine to help improve outcomes and deliver informational material to consumers by reaching out to stakeholder organizations. In addition, we have built new relationships with the American College of Obstetricians and Gynecologists to engage their networks of providers and stakeholder organizations to improve consumer involvement in processes.

Consumer involvement: By building these partnerships that are already dedicated to improving maternal and child health care, the US-PCG hopes to establish itself as an essential resource for evidence-based health care decision-making as well as provide dissemination materials to be publicized from a variety of sources.

INNOVATIVE SOLUTIONS FOR EVIDENCE PRODUCTION

Accuracy of the Hospital Anxiety and Depression Scale – Depression subscale (HADS-D) for detecting major depression: individual participant data meta-analysis

Wu Y[1], Levis B[2], Benedetti A[2], Thombs BD[1]

[1] Department of Psychiatry, McGill University, Canada; [2] Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Canada

Background: The Depression subscale of the Hospital Anxiety and Depression Scale (HADS-D) is recommended for depression screening in people with physical health problems. A score of ≥ 11 is considered a clinically significant depressive disorder, whereas a score between 8 and 10 suggests a mild disorder. Existing conventional meta-analyses, however, were conducted 10 years ago and were limited by excluding studies that did not report the results for standard cut-offs (≥ 8 or ≥ 11), incomplete cut-off reporting within included studies, and by not examining accuracy by reference standard.

Objectives: To evaluate HADS-D accuracy for detecting major depression in medically ill patients using individual participant data meta-analysis (IPDMA).

Methods: Eligible studies compared PHQ-9 scores with major depression diagnoses from validated diagnostic interviews. We synthesized primary study data and study level data extracted from primary reports. For HADS-D cut-off scores 5 to 15, we used bivariate random effects meta-analysis to estimate pooled sensitivity and specificity, separately, among studies that used semi-structured diagnostic interviews, which are designed for administration by clinicians; fully structured interviews, which are designed for lay administration; and the Mini International Neuropsychiatric (MINI) diagnostic interviews, a brief fully structured interview. We examined sensitivity and specificity using one-stage meta-regression by reference standard categories, considering all participant characteristics in a single model.

Results: We obtained individual participant data from 82 of 141 eligible studies (17,176 participants, 2100 cases) (see Table 1). We maximized combined sensitivity and specificity at cut-off ≥ 7 for semi-structured interviews, and fully structured interviews, and maximized at cut-off ≥ 6 for the MINI (See Table 2). Among studies with a semi-structured interview (44 studies, 6614 participants, 754 cases), sensitivity and specificity (95% CI) were 0.83 (0.77 to 0.88) and 0.78 (0.73 to 0.82) for a cut-off of ≥ 7, 0.74 (0.68 to 0.80) and 0.84 (0.80 to 0.87) for a cut-off of ≥ 8, and 0.45 (0.38 to 0.53) and 0.95 (0.93 to 0.96) for a cut-off of ≥ 11. Accuracy was similar across reference standards and subgroups, including for participants in different recruiting settings.

Conclusions: A HADS-D cut-off of ≥ 7 maximized combined sensitivity and specificity; the standard cut-offs of ≥ 8 and ≥ 11 were less sensitive but more specific.

Patient or healthcare consumer involvement: There was no direct patient or healthcare consumer involvement in this study. However, we will develop a web-based knowledge translation tool which will help clinicians considering screening for depression with the HADS-D to estimate the expected numbers of positive screens and the true and false screening outcomes based on results from the present IPDMA.

Additional files: Tables

Accuracy of the PHQ-9 for screening to detect major depression: an updated systematic review and meta-analysis of individual participant data

Negeri ZF[1], Levis B[2], Benedetti A[1], Thombs BD[2], DEPRESSD PHQ Collaboration N

[1] Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Canada; [2] Lady Davis Institute for Medical Research, Jewish General Hospital, Canada

Background: Depression accounts for more years of "healthy" life lost than any other medical condition. Major depressive disorder (MDD) is present in 5% to 10% of primary care patients and 10% to 20% of patients with chronic medical conditions. The Patient Health Questionnaire-9 (PHQ-9) is the most commonly used tool for screening major depression.

Objectives: This study aimed to determine the accuracy of the PHQ-9 for detecting major depression.

Methods: Individual participant data meta-analysis (IPDMA) was used to synthesize primary data obtained from several search engines (January 2000 to May 2018). We used a bivariate generalized linear mixed-effects model to estimate overall sensitivity and specificity for PHQ-9 cut-off scores 5 to 15, separately, among studies that used semi-structured, fully structured, and the Mini International Neuropsychiatric (MINI) diagnostic interviews. Meta-regression was used to examine potential associations between participant characteristics and the accuracy of the PHQ-9.

Results: We obtained data from 100 of 123 eligible studies (81%), for a total of 44,503 participants and 4541 major depression cases. We maximized sensitivity and specificity at a cut-off score of 10 or above among studies using a semi-structured interview (47 studies, 11,234 participants; Sensitivity = 0.85, 95% confidence interval (CI): 0.79 to 0.89; Specificity = 0.85, 95% CI 0.82 to 0.87), and a cut-off score of 8 or above among studies using both the fully structured and MINI interviews. The sensitivity (95% CI), specificity (95% CI) of the PHQ-9 at cut-off score of 8 among fully structured and MINI interviews were 0.77 (0.66 to 0.86), 0.81 (0.74 to 0.86) and 0.85 (0.79 to 0.89), 0.8 (0.76 to 0.83), respectively. Meta-regression showed that the age and sex of participants were significantly associated with the specificity of the PHQ-9. The median specificity for older participants was greater by 2% to 6% across cut-offs compared to younger participants across reference standards. The median specificity of the PHQ-9 for female participants was less than for male participants by 3% and 4% for the fully and semi-structured interviews, respectively.

Conclusions: The diagnostic accuracy of the PHQ-9 is higher for semi-structured interviews compared to fully structured and MINI interviews. Cut-off scores of 10 for semi-structured and 8 for fully and MINI studies yielded optimal sensitivity and specificity. Older age was associated with higher specificity for all three reference standard categories, and female participants tend to have lower specificity compared to male participants for semi-structured and fully structured interviews.

Patient or healthcare consumer involvement: There was no direct patient or healthcare consumer involvement in this study. However, we will update a web-based knowledge translation tool (depressionscreening100.com/phq) to help clinicians considering screening for depression with the PHQ-9 to estimate the expected numbers of positive screens and the true and false screening outcomes based on results from this study.

Evaluating risk of bias and applicability in individual participant data meta-analyses

Levis B[1], Snell K[1], Takwoingi Y[2], Collins GS[3], Moons KG[4], Reitsma JB[4], Hooft L[4], Stewart LA[5], Tierney JF[6], Riley RD[1]

[1] Centre for Prognosis Research, School of Primary, Community and Social Care, Keele University, UK; [2] Test Evaluation Research Group, Institute of Applied Health Research, University of Birmingham, UK; [3] Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, UK; [4] Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, The Netherlands; [5] Centre for Reviews and Dissemination, University of York, UK; [6] London Hub for Trials Methodology Research, MRC Clinical Trials Unit, University College London, UK

Background: Assessing risk of bias and applicability (RoB) of included studies is critical for interpreting meta-analysis (MA) results. RoB tools for diagnostic, prognostic, and prediction studies include the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) and PROBAST. However, individual participant data meta-analyses (IPD-MAs) differ from aggregate-data MAs in that in IPD-MA, datasets may include additional information, eligibility criteria may differ from the original publications, and definitions for index tests/predictors and reference standards/outcomes can be standardized across studies. Thus, tailored RoB tools may be needed.

Objectives: To review how RoB is currently assessed in IPD-MAs, and to examine QUADAS-2 and PROBAST, with the goal of developing IPD-MA extensions for each tool.

Methods: We reviewed RoB assessments in IPD-MAs published in the last 12 months. We then examined how QUADAS-2 (and in-progress extensions) and PROBAST items might be evaluated in an IPD-MA context; noting which items might be removed, edited, or added; and hypothesized how results may be incorporated into IPD-MA analyses.

Results: We observed that current IPD-MAs rarely and inconsistently evaluate RoB, and most do not incorporate RoB judgements into analyses. Our findings indicate using QUADAS-2 and PROBAST to assess RoB of IPD datasets themselves, rather than study publications. Certain items may need to be coded at the participant level (e.g. timing between index test/predictor and reference standard/outcome), whereas others (e.g. quality of diagnostic tool) may apply uniformly to an included study. Most analysis items (e.g. pre-specification of thresholds and variables for analysis) may not be relevant, as IPD-MA researchers perform the analyses themselves. RoB results may be incorporated into analyses by conducting subgroup analyses among studies and participants with overall low RoB or by conducting formal interaction analyses with item-level RoB responses.

Conclusions: Development and dissemination of IPD-MA extensions for QUADAS-2 and PROBAST will lead to improved RoB assessments in IPD-MAs of diagnostic, prognostic, and prediction studies.

Patient or healthcare consumer involvement: There was no direct patient or consumer involvement in this project. However, IPD-MAs are considered the gold-standard of evidence synthesis. Understanding and improving how risk of bias and applicability are assessed, reported and incorporated in IPDMAs will result in better understanding of IPD-MA evidence and thus lead to more informed health policies and better patient care.

Considerations for establishing a new, scholarly peer-reviewed journal: a partially completed protocol

Ng JY[1], Ashbury FD[2], Cooley K[3], Haddad PS[4], Guyatt GH[1], Levine MA[1], Busse JW[1]

[1] Department of Health Research Methods, Evidence and Impact, Faculty of Health Sciences, McMaster University, Hamilton, Canada; [2] Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Canada; [3] Department of Research, Canadian College of Naturopathic Medicine, Toronto, Canada; [4] Department of Pharmacology and Physiology, Faculty of Medicine, University of Montreal, Montreal, Canada

Background: Estimates suggest the existence of approximately 30,000 scholarly journals, and raise questions regarding the wisdom of starting new journals. New journals lack key metrics of legitimacy, such as an impact factor, or Directory of Open Access Journals (DOAJ) and Committee on Publication Ethics (COPE) memberships, by the simple virtue of being new. Nevertheless, there may be instances when starting a new journal meets a legitimate unmet need. However, little guidance exists for academic publishers to identify whether a need exists for a newly launched peer-reviewed journal.

Objectives: We propose to develop a two-step mechanism for academic publishers to identify whether a need exists for a newly launched periodical.

Methods: Step 1) By searching publicly available data online, including faculty webpages and publications, we have created an international database of researchers pertinent to the scope of a newly launched publication focussed on natural health products. We then manually extracted the following items for pertinent researchers: researcher name, affiliation, academic rank, research interest(s), faculty webpage URL, and title and journal of recently published research articles. Step 2) Upon database completion following search saturation, we designed and are currently administering a survey that includes questions that capture researchers' preferences towards open access publishing and the desirability for a new field-specific journal.

Results: Survey collection is still ongoing. Upon completion, this will serve two major purposes: 1) to provide publishers with a clear understanding of what researchers and types of research exist (or alternatively, are lacking) in the discipline (i.e. natural health products) pertaining to their contemplated publication, and 2) to afford the opportunity to directly gather data from potential future journal contributors, allowing publishers to specifically tailor their publication to their authors' needs and preferences.

Conclusions: Academic publishers have a responsibility to establish novel journals based on a properly identified need expressed by the researcher community. We propose a unique and novel model that provides academic publishers with actionable steps to identify whether a need exists for a newly launched peer-reviewed journal. Future work standardizing this protocol may result in the development of a guideline offering academic publishers greater guidance in establishing new periodicals positioned for greater long-term success.

Patient or healthcare consumer involvement: Besides researcher and publisher participation, patients with a particular interest in reading peer-reviewed literature should also be promoted to contribute to improving this protocol.

Core outcome sets and systematic reviews: untapped potential?

Saldanha I[1], Gorst S[2], Williamson P[2]

[1] Brown University School of Public Health, USA; [2] University of Liverpool, UK

Background: Outcomes are fundamental to systematic reviews. Core outcome sets (COS), which are agreed standardized sets of outcomes within topic areas, can help primary researchers as well as systematic review authors consistently measure outcomes tha t are most important and relevant to decision-making. More than 330 published COS are available in an online, publicly searchable database maintained by the Core Outcome Measures for Effectiveness Trials (COMET) Initiative. The utilization and impact of COS on systematic reviews have not been examined.

Objectives: For a sample of recently conducted systematic reviews, we examined the proportion of reviews:

1) for which a relevant COS exists;
2) where the authors mentioned using a COS in choosing outcomes for the review;
3) which identified problems with outcome inconsistency in included studies; and
4) which noted the need for development of COS.

Methods: We examined all Cochrane intervention reviews addressing any topic and published by any Cochrane Review Group between 1 January 2019 and 8 March 2019 and all US Agency for Healthcare Research and Quality (AHRQ)-funded systematic reviews published between 1 January 2018 and 15 March 2020. We extracted the information relevant to our Objectives that may have been reported in any section of the report of the review.

Results: We included 100 Cochrane Reviews and 61 AHRQ systematic reviews that addressed a range of diseases. Although a potentially relevant COS exists for 33/100 Cochrane Reviews (33%) and 18/61 AHRQ reviews (30%), only 7/100 Cochrane Reviews mentioned (or cited) a COS in relation to choosing outcomes for the review. This is despite 40/100 Cochrane Reviews explicitly noting problems of inconsistency in outcomes in included studies. Among the 40 reviews that identified outcome inconsistency, only six reviews (15%) noted the need for development of a COS. At the Colloquium, we will present detailed findings about both Cochrane Reviews and AHRQ reviews in terms of their use of COS.

Conclusions: This examination of recent systematic reviews suggests that although relevant COS exist for about a third of systematic reviews, very few reviews use COS to help inform their choice of outcomes. This is despite outcome inconsistency across included studies continuing to be a widespread problem. COS can help ensure that the most relevant outcomes are evaluated and that more included studies are incorporated into meta-analyses. We recommend that, when a relevant and well-developed COS exists, systematic review authors should at least consider it to inform their process of choosing outcomes for their review. The COMET Database is online, free, searchable, and covers a range of health fields. No relevant COS exist for approximately two-thirds of the systematic reviews we examined. Authors of systematic reviews should push their respective fields forward by explicitly noting the need for COS, recommending COS development, and participating in COS development.

Patient or healthcare consumer involvement: No, but COS are usually developed with consumer involvement.

Developing a technological platform for collection and utilization of patient reported data in Ontario Health Teams

Ghadban M[1], Ankita A[1]

[1] University of Toronto, Canada

Background: Ontario Health Teams (OHTs) were recently introduced as an integrated healthcare delivery model that brings together different healthcare providers as one co-ordinated team. With 24 confirmed OHTs to date, this new model aims to eliminate the gaps and redundancies in our current model, and ultimately improve patient care. The OHT is designed to deliver a full and co-ordinated continuum of care in a defined geographical region which offers 24 hours access of care with an aim to improve performance. The aim of OHT is to provide better patient and population health outcomes, better user and provider experience and better value. It provides direction to integrate existing healthcare delivery models to teams. The introduction of a new model of healthcare delivery brings with it questions on how to best evaluate this new model. Typically, this is done through patient reported data. However, we face a unique challenge: An integrated delivery model is novel in Ontario and there are no established tools to evaluate the performance of such a model. While there are many barriers faced when collecting and utilizing patient reported data, the lack of adequate technology is at the root of these barriers.

Objectives: To develop a technological platform that will allow OHTs to collect patient reported data, and make use of it in improving care delivery and service model evaluation within an OHT and across all OHTs.

Methods: Our approach uses an integrated knowledge translation process. First, we identify gaps in existing technological platforms by conducting a scoping review. We then validate these findings through a series of interviews with the stakeholders involved in collecting and using patient-reported data. These stakeholders include healthcare providers such as physicians, nurse practitioners, allied service, patients, caregivers and authorities. Finally, we will synthesize these data to propose the features that a new technological platform must have. The end product will be beta tested to ensure regulatory compliance and user friendliness.

Patient or healthcare consumer involvement: Patients will be engaged through interviews and as design partners in developing the features required in our proposed patient reported data platform.

Evidence collection during disease outbreaks: living systematic reviews

Counotte MJ[1], Imeri H[1], Meili K[2], Low N[1]

[1] Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland; [2] Department of Epidemiology and Global Health, Umeå University, Umeå, Sweden

Background: Early in disease outbreaks, evidence is often scarce but accumulates rapidly. Living systematic reviews (LSR) are systematic reviews that are updated as soon as new information becomes available and provide a solution to keep up with the evidence. However, LSRs in an emerging outbreak setting face unique challenges compared to other LSRs: Relevant information first becomes available in preprint publications, that are replaced by their peer-reviewed version and as outbreaks emerge, search terms may change.

Objectives: Here, we describe how we built an LSR system to cope with the rapidly evolving evidence of emerging outbreaks. We describe the challenges we faced conducting LSRs during the Zika virus and the COVID-19 outbreak. We focus on the methods used to retrieve unique citations from different sources and creating updatable data output.

Methods: We use application programming interfaces (API) to collect citation data from the preprint servers BioRxiv and MedRxiv, and from the medical bibliographic databases Embase, PubMed (Figure 1). We verify and clean the data. We apply deduplication algorithms to retain unique citations. We compare a rule-based algorithm using similarity scores with thresholds, a logistic regression model predicting duplicate status and a blended approach, where both were combined. We calculate similarity scores between titles, authors, journal names and other properties, using Longest Common Subsequence and other similarity indices. In a test-set of 2500 records from Embase and PubMed, with 220 duplicate and 2280 non-duplicate pairs, both the logistic regression prediction and the blended approach had a sensitivity of 100% and a specificity of 100%. Unique citations are retained and imported into a central database in 'Research Electronic Data Capture' (REDCap). Although primarily designed for data collection for clinical trials, the database allows tracking changes, has a conflict-resolution workflow, and allows application programming interface access. The database allows flexible data output, formatted as Research Information Systems (RIS) or Extensible Markup Language (XML), compatible with all citation managers.

Figure 1.
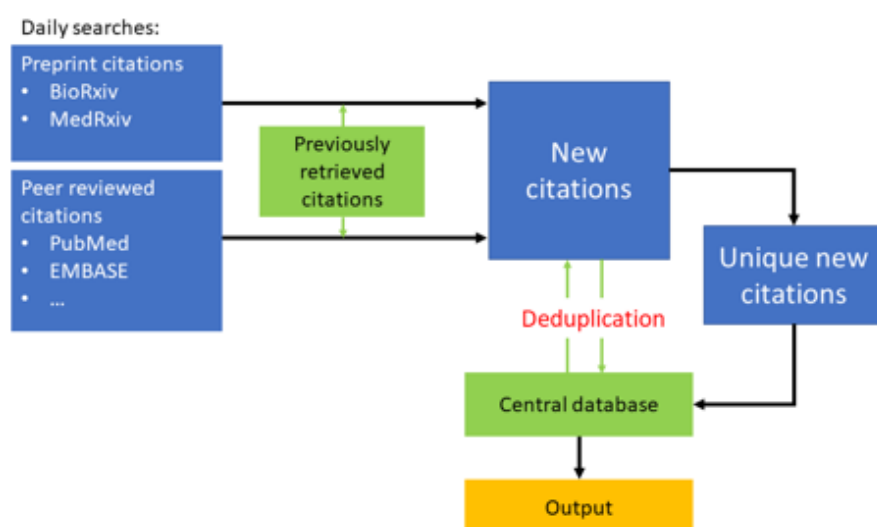
Results: The LSR system allows us to optimize our workflow. We receive daily new deduplicated citations from different sources. Using a central database, in which we screen and extract data, we can create dynamic content that allows rapid updating of figures, tables and other output. Indexed citations are distributed for screening and verification to a crowd; online tools are used to allow rapid screening and verification.

Conclusions: LSRs in an emerging outbreak allow us to keep up with the evidence, however, they pose unique challenges.

Give me the 'C'! How to define the comparator in a prognostic factor systematic review

Stallings E[1], Jimenez D[2], Muriel A[3], Zamora J[3], Fernandez-Felix B[3], Quezada C[2], Lopez-Alcalde J[4]

[1] Clinical Biostatistics Unit, Hospital Universitario Ramón y Cajal; Cochrane Madrid; CIBERESP, Spain; [2] Respiratory Department, Ramón y Cajal Hospital and Universidad de Alcalá IRYCIS, Madrid, Spain; [3] Clinical Biostatistics Unit, Hospital Universitario Ramón y Cajal (IRYCIS); CIBER Epidemiology and Public Health (CIBERESP), Spain; [4] Cochrane Madrid / Universidad Francisco de Vitoria, Madrid; Instituto Ramón y Cajal de Investigación Sanitaria (IRYCIS), Madrid; Institute for Complementary and Integrative Medicine, University Hospital Zurich and University of Zurich; CIBERESP, Spain

Background: Our team is working on several prognostic factor systematic reviews (SRs). A SR aiming to determine the independent prognostic significance of a factor must come from a structured question (population, index factor, comparator, outcomes, timeframe and study design). The definition of the comparator requires identification of additional prognostic factors for which the prognostic association should be adjusted for. Ideally, these key additional prognostic factors should be defined at the protocol stage of the systematic review.

Objectives: To describe the procedure that we followed to select the additional prognostic factors in the Cochrane protocol titled "Sex as a prognostic factor in patients with acute symptomatic pulmonary embolism".

Methods: We carried out a bibliographic search in PubMed and Embase to identify prognostic factors in acute pulmonary embolism (PE). This search retrieved six factors, which we compiled in GRADEpro GDT. We sent the list to the systematic review team. In this first stage, the team commented on the factors already listed or added new ones. At this stage, a total of 24 factors were compiled. In the second stage, the review team prioritized these factors, ranking them from 1 to 9 (with 1 being of least importance and 9 the highest). There was also the option to choose "Unknown". Once all the team members had finished ranking, they were asked to confirm the list of factors and the order in which they had been prioritized.

Results: The additional prognostic factors were classified into three groups: high priority (5 factors), low priority (9 factors), and excluded (10 factors). The high priority factors chosen were: immobilization history, history of surgery, history of recent bleeding, PESI score, and simplified PESI score.

Strengths: (1) our approach is transparent; (2) the process is straightforward in GRADEPro-GDT and doesn't require presential meetings; and (3) this process highlights the need to define an evidence-based procedure to define the list of additional confounders, which may be applied to any SR including non-randomized designs.

Limitations: (1) the criteria to define the relevance of the additional prognostic factors relied on clinical judgement only; however, there should be an evidence-based approach in place: for example, the additional adjustment factors should be statistically associated with both the outcome and the prognostic factor, and should not lie on a direct pathway between the prognostic factor and the outcome; and (2) the maximum number of additional prognostic factors is not defined and to limit the number is not simple.

Conclusions: We applied a transparent procedure for selecting additional prognostic factors to consider in a prognostic factor SR. This procedure can be applied to SRs of prognostic factor studies in the future. However, more research is needed to define the criteria on which to base the decisions to select the additional prognostic factors or the maximum number of factors to select.

Patient or healthcare consumer involvement: None foreseen

How the Evidence Synthesis Hackathon supports the efficient and rigorous production of systematic reviews

Haddaway N[1], Westgate M[2]

[1] Mercator Research Institute on Global Commons and Climate Change, Germany; [2] Australian National University, Australia

Background: Systematic reviews are resource-intensive, requiring several years to complete on average. Over recent decades, there has been an exponential increase in the volume of published primary research, challenging the efficiency and timeliness of systematic reviews. Rigorous systematic reviews must spend more and more resources in order to synthesise a given topic as evidence continues to expand. Technology may support and improve evidence syntheses in the face of this evidence avalanche by: simplifying review conduct; assisting transparency; supporting group working; increasing efficiency by supplementing or taking over certain tasks; and, increasing rigour. However, Evidence Synthesis Technologies (ESTech; innovative tools and techniques designed to support evidence synthesis conduct) are typically produced in isolation by small groups of researchers and coders working within specific disciplines. ESTech softwares typically have little to no long-term support, and many are hidden behind paywalls. A concerted effort to foster an active community of linked researchers and coders working on ESTech is therefore needed to reduce redundancy, increase interoperability, and ensure needs-driven approaches to software development.

Objectives: To create a community and provide a platform for continued collaboration on ESTech.

Methods: The Evidence Synthesis Hackathon (ESH) is a series of interactive workshops bringing together experts in systematic review and programming to brainstorm and code new technologies to support evidence synthesis. The ESH is guided by Open Synthesis (Open Science in evidence synthesis), producing tools to increase transparency, efficiency, repeatability, rigour and accessibility.

Results: The ESH has been run three times since January 2018, and has produced a range of outputs to support evidence synthesis across disciplines. Some of these are traditional softwares, whilst others are conceptual papers discussing fundamental issues relevant to evidence synthesis infrastructure and methodology. The ESH has received broad support from across and within the Collaboration for Environmental Evidence, the Campbell Collaboration and Cochrane.

Conclusions: We demonstrate that the ESH is a fit-for-purpose network for needs-driven development of tools and techniques to increase the rigour and efficiency of evidence synthesis. We call for continued engagement from the evidence synthesis community in suggesting areas where technology is needed and in joining and funding future ESH events.

Patient or healthcare consumer involvement: ESH projects have not yet interfaced with patients or healthcare consumers, focusing rather on evidence commissioners, producers and decision-makers. Specific projects may need to engage with end users, however, for example when visualizing evidence. This will be a required part of any relevant future ESH project.

Identifying predatory journals in systematic reviews

John D[1], Polani Chandrasekar R[2], Lohmann J[3], Dazy A[3]

[1] Campbell South Asia, India; [2] ICMR-National Institute of Epidemiology, India; [3] London School of Hygiene & Tropical Medicine, UK

Background: This paper aims to provide a systematic and transparent process for authors of systematic reviews to identify predatory journals prior to data synthesis or meta-analysis of included studies which have not benefitted from standard quality checks.

Methods: We used two systematic reviews from low- and middle-income countries as case studies, one on snakebites and one on the mental health of health workers. We removed articles published in duplicate journals and identified journals indexed in PubMed. We critically appraised the non-PubMed indexed journals using a checklist derived from https://thinkchecksubmit.org. We gathered information on the characteristics of the journal such as name and country of the publisher, latest papers/articles and contact details of the publisher, peer-review policy and process, article processing charges/publication fees and editorial board and members from the journal/publisher webpage. We scrutinized the journal names or publishers listed as predatory in the list of predatory journals and publishers on the stoppredatoryjournals.com webpage. We recorded responses regarding the journal and its editorial board with colleagues sharing our office space. We searched Zotero/Endnote libraries to check whether we had read or cited any other articles from the journal earlier. We conducted a search for listings of the journals in Web of Science, ProQuest, and Science direct databases. We checked the individual and affiliated institutes webpage of the editorial members to check whether the journal name was mentioned in their bio. We also checked the existence of the journal/Publisher in the Code of Publication Ethics (COPE), Directory of Open Access Journal (DOAJ), Open Access Scholarly Publishing Association (OASPA), Journals online project website (INASP Journals Online Platform) and STM (the global voice of scholarly publishing).

Results: After removing the duplicates we identified 68 journals, with 36 journals indexed in PubMed. Of the 32 non-PubMed indexed journals, two were listed in Web of Science. Most of the journals had a publisher based in India. Thirteen journals were listed in Beall's list or stoppredatoryjournal.com. Nine journals had mentioned the publication fees transparently on their website. None of the journals were listed with COPE, one journal was listed with DOAJ and one in OASPA. Six journals had clear peer-review guidelines, and five journals had phone/email details on its website. Based on the ThinkCheckSubmit checklist, 14 studies were identified as 'more likely predatory' and four articles as 'likely predatory'. Two of the 'likely predatory' articles were also indexed in Web of Science.

Conclusion: Review authors need to be aware of the possibility of including in their systematic review an article published in a predatory journal. The process described here could be used by review authors to identify predatory journals as part of the systematic review process.

LOCATE: a prospective evaluation of the value of Leveraging Ongoing Citation Acquisition Techniques for living Evidence syntheses

Gates M[1], Elliott SA[2], Gates A[1], Sebastianski M[3], Bialy L[3], Pillay J[1], Hartling L[1]

[1] Alberta Research Centre for Health Evidence (ARCHE), Department of Pediatrics, University of Alberta, Canada; [2] Alberta Research Centre for Health Evidence (ARCHE), Department of Pediatrics, University of Alberta; Cochrane Child Health, Canada; [3] Alberta Strategy for Patient-Oriented Research (SPOR) SUPPORT Unit Knowledge Translation Platform, Department of Pediatrics, University of Alberta, Canada

Background: Because they are time- and resource-intensive to produce, decision-makers often rely on systematic reviews (SRs) that are out of date. Living SRs (LSRs) offer a potential solution by locating and incorporating new evidence in real time; however, the expedited methods needed to maintain LSRs are not well established.

Objective: To explore the value of three complementary search approaches in the context of LSRs in terms of search performance, screening workload, and feasibility compared to the reference standard.

Methods: A research librarian developed three complementary search approaches for a SR on five pharmacologic treatments for bronchiolitis: the automated full search, PubMed Similar Articles, and Scopus Citing Articles. Beginning in October 2018, we automated the searches to run monthly (except PubMed Similar Articles, which was performed manually). Pairs of review authors (one per search approach) independently screened records retrieved via each search and commented on feasibility monthly. After one year, we conducted a full update search in four online databases, selected conference proceedings and clinical trials registers (reference standard). For each complementary approach, we calculated search performance (proportion missed, number needed to read (NNR)) and review author workload (number of records screened, time required) compared to the reference standard. We summarized comments about feasibility. We will investigate the impact of the newly located trials on the effect estimate and certainty of evidence for the two primary outcomes.

Results: Via the reference standard, review authors screened 505 titles/abstracts and 24 full texts and identified 4 new trials (NNR 126.6; 12.4 hours). Of the complementary approaches, only the automated full search located all four trials; these were located 5 to 12 months sooner than via the reference standard. The automated full search was the most resource-intensive approach (NNR 204.1; 17.0 hours). Compared to the reference standard, review authors screened more records (811 titles/abstracts and 21 full texts), due to duplicates. The PubMed Similar Articles and Scopus Citing Articles approaches located far fewer candidate records (244 and 451, respectively), thereby requiring less screening time (5.9 and 8.7 hours, respectively); however, each approach located only one of the four new trials (75% missed). Review authors found it feasible and convenient to conduct monthly screening for searches of this yield (median 15 to 65 records/month).

Conclusions: The automated full search located relevant trials sooner, but required more screening time than the reference standard. Although the monthly screening time for the PubMed Similar Articles and Scopus Citing Articles was far less, most relevant records were missed. Exploration of the impact of locating the four trials on the results and conclusions of the SR will provide additional insight into the value of each approach.

Patient or consumer involvement: Consumers were not directly involved but will benefit from LSR methods that ensure the availability of up-to-date evidence.

Open Synthesis: the application of Open Science principles to improve the rigour of systematic reviews

Haddaway N[1], Lotfi T[2]

[1] Mercator Research Institute on Global Commons and Climate Change, Germany; [2] McMaster University, Canada

Background: The Open Science movement can be broadly summarized as aiming to promote integrity, repeatability and transparency across all aspects of research, from data collection to publication. To date, evidence synthesis has only partially embraced Open Science, typically striving for Open Methodology and Open Access, and occasionally providing sufficient information to be considered to have Open Data for some published reviews.

Objectives: To launch the Open Synthesis Working Group (OSWG) that advocates to increase reliability, trust and reuse of information collected and synthesized within a review.

Methods: The OSWG is a cross-disciplinary, diverse group of experts in evidence synthesis, evidence ecosystems and evidence synthesis technology. The group aims to develop a definition of Open Synthesis along with recommendations of how communities of practice can make evidence syntheses more Open, in terms of ways of working with stakeholders, methods, data, software, analytical code, publication access, peer-review and educational materials. The founders of this Group built a list of potentially interested researchers and other stakeholders, along with the criteria for joining.

Results: The group is voluntary and open, and aims to produce relevant working papers, definitions and supporting materials through ongoing collaboration and discussion.

Conclusions: It is hoped that the group will help to develop clear pathways to more transparent, rigorous and truly open evidence synthesis activities, leading to greater impact and stronger legacy across organizations such as Cochrane, the Campbell Collaboration and the Collaboration for Environmental Evidence.

Patient or healthcare consumer involvement: Open Synthesis would promote meaningful and fair engagement with patients and healthcare consumers. In advocating for Open Synthesis, this group is protecting patients indirectly by making evidence available and accessible: for example, if a patient (or other stakeholder) wishes to re-run an analysis that led to a decision favouring one treatment over another.

Our Cochrane protocol is under peer-review: what can we do before its publication?

Stallings E[1], Lopez-Alcalde J[2], Álvarez-Díaz N[3], Monge D[4], Zamora J[5]

[1] Clinical Biostatistics Unit, Hospital Universitario Ramón y Cajal, Cochrane Madrid, CIBERESP, Spain; [2] Cochrane Madrid; Universidad Francisco de Vitoria, Madrid; Instituto Ramón y Cajal de Investigación Sanitaria (IRYCIS), Madrid; Institute for Complementary and Integrative Medicine, University Hospital Zurich and University of Zurich; CIBERESP, Spain; [3] Library, Hospital Universitario Ramón y Cajal (IRYCIS), Madrid, Spain; [4] Cochrane Associate Center of Madrid; Universidad Francisco de Vitoria, Madrid, Spain; [5] Clinical Biostatistics Unit, Hospital Universitario Ramón y Cajal (IRYCIS), CIBER Epidemiology and Public Health (CIBERESP), Spain

Background: The process to have a Cochrane protocol published is lengthy, sometimes taking up to seven or eight months. This can delay the review project and some researchers may avoid performing Cochrane Reviews for this reason. The process involves external reviewers assessing the protocol, sending their comments to the authors, the authors revising these comments and making the necessary changes and then returning the protocol to Cochrane with corrections implemented ready for publication. This process is carried out before the review can begin. We propose an approach to speed up the review process before the protocol is published.

Objectives: To describe the approach we used to speed up the review process before publication of a Cochrane protocol.

Methods: Description of the approach followed at Cochrane Madrid to speed up the review process before publication of the protocol "Sex as prognostic factor for mortality in adults with acute symptomatic pulmonary embolism".

Results: Our approach involved the following tasks: (1) To carry out a preliminary search in PubMed; (2) To request the EPPI-Reviewer team to have our project ready to use before the protocol publication; (3) To screen through the preliminary search results to identify eligible studies; (4) To screen conferences for eligibility; (5) To create and pilot the data extraction template in EPPI-Reviewer; (6) To create and pilot the risk-of-bias tool (QUIPS) in EPPI-Reviewer; (7) To request the Cochrane Editorial Group to approve and execute the final search strategy before the protocol publication; (8) When we received the results from the final Cochrane search, we combined them with our preliminary search results and we only screened those references that had not been previously screened. We used Endnote to facilitate this process of combining the two sets of references of search results.

Our preliminary search retrieved 18,672 references. We screened 2673 references and found 2 studies to include in the SR. We created the data extraction template and QUIPS tool in EPPI-Reviewer and piloted this process. These included studies were then available for the SR team to work with three months before the protocol was published. We estimate that these preliminary tasks saved around two months for our project.

Conclusions: Carrying out certain review tasks before the protocol publication saves time in the review project. We suggest that Cochrane publish guidelines on the steps that can be taken while waiting on the protocol publication. These steps could include the screening of preliminary search, combining the preliminary search results with the Cochrane search results, receiving access to EPPI-Reviewer and Covidence before the protocol publication, and piloting the data extraction and risk of bias with several included studies.

Patient or healthcare consumer involvement: None foreseen.

Presentation approaches for enhancing interpretability of patient-reported outcomes measures in meta-analyses: a systematic survey of Cochrane Reviews

Zeng L[1], S Devji TSD[2], Wang Y[2], Yao L[2], Han M[3], K Nampo F[4], Granholm A[5], Tadayon B[2], Manja V[6], El-Khechen H[2], C Johnston B[2], Bhandari M[2], Thabane L[2], L Patrick D[7], H Guyatt G[2]

[1] West China Second University Hospital, Sichuan University, China; [2] McMaster University, Canada; [3] Chosun University, South Korea; [4] Federal University of Latin-American Integration, Brazil; [5] Copenhagen University Hospital, Denmark; [6] University of California Davis, USA; [7] University of Washington, USA

Background: Interpreting the magnitude and importance of pooled effects in meta-analysis can be challenging, particularly for patient-reported outcomes (PROs) measured using questionnaires with which clinicians have limited familiarity.

Objectives: We are evaluating Cochrane systematic review authors' approaches to calculation, presentation and interpretation of PRO measures (PROMs) results in meta-analyses.

Methods: Our methodological survey includes Cochrane systematic reviews that report at least one statistically significant result for at least one PRO measured by a PROM reported as a continuous outcome. We are including 200 eligible Cochrane systematic reviews as planned in our published protocol. We started from Cochrane systematic reviews published in April 2019 and screened backwards, stopping at October 2015 when we achieved the target sample size. We are documenting authors' approach to calculating pooled effects when studies used different PROMs for the same construct (standardized mean difference, natural units of the most widely used instrument, dichotomization with relative effects, dichotomization with absolute effect, minimally important difference (MID) units) and their approach to interpreting the importance or magnitude of the pooled effects (e.g. reference to the MID). We are assessing whether a MID was applied in the calculation, presentation and interpretation of PRO measures.

Results: We plan to finish data analysis in June.

Discussion: Our methodological survey will inform the systematic review community regarding the current practice of summarizing and presenting effect estimates for PROMs presented as continuous variables in Cochrane systematic reviews. We anticipate an underuse of the most informative available methods, and possible deficiencies in interpretation. We anticipate our results will influence recommendations on reporting, conduct and interpretation of PROMs and subsequent practice.

Patient or healthcare consumer involvement: Not involved.

Probability of major depression classification based on different diagnostic interviews: a synthesis of individual participant data meta-analyses

Wu Y[1], Levis B[1], Ioannidis JPA[2], Benedetti A[1], Thombs BD[1], DEPRESSD Collaboration

[1] McGill University, Canada; [2] Stanford University, USA

Background: Three previous individual participant data meta-analyses (IPDMAs) reported that, compared to the Structured Clinical Interview for DSM (SCID), alternative reference standards, primarily the Composite International Diagnostic Interview (CIDI) and Mini International Neuropsychiatric Interview (MINI), tended to misclassify major depression status, controlling for depression symptom severity. However, there was important imprecision in the results.

Objectives: To synthesize results from three previous IPDMAs and compare performance of the most commonly used diagnostic interviews for major depression classification, the SCID, CIDI, and MINI.

Methods: We used databases from published IPDMAs that used the Edinburgh Postnatal Depression Scale (EPDS) and Hospital Anxiety and Depression Scale – Depression subscale (HADS-D) to control for depressive symptoms and an updated Patient Health Questionnaire-9 (PHQ-9) database. We standardized screening tool scores across IPDMA databases. For each IPDMA, separately, we fit binomial generalized linear mixed models to compare adjusted odds ratios (aORs) of 1) major depression classification, controlling for symptom severity and participant characteristics, and 2) the interaction between interview and symptom severity. We synthesized results using DerSimonian-Laird random-effects meta-analysis.

Results: We included 69,405 participants (7574 (11%) with major depression) from 212 studies (Table 1). As reported in Table 2, controlling for symptom severity and participant characteristics, the MINI (74 studies; 25,749 participants) classified major depression more often than the SCID (108 studies; 21,953 participants; aOR (95% confidence interval (CI) 1.46 (1.11 to 1.92)). Classification odds for the CIDI (30 studies; 21,703 participants) and SCID did not differ overall (aOR (95% CI) 1.19 (0.79 to 1.75)), but as screening scores increased, aOR increased less for the CIDI than the SCID (interaction aOR (95% CI) = 0.64 ([0.52 to 0.80)).

Conclusions: Compared to the SCID, the MINI classified major depression more often. Odds of depression classification with the CIDI increased less as symptom levels increased. Interpretation of research that uses diagnostic interviews to classify depression should consider interview characteristics.

Patient or healthcare consumer involvement: Patients and healthcare consumers were not directly involved in the study. However, better understanding differences in performance of different reference standards for major depression classification will lead to better quality evidence syntheses and thus more informed health policies for mental health.

Additional files: Tables

Screening for depression in pregnant and postpartum women: an individual participant data meta-analysis of the Edinburgh Postnatal Depression Scale

Levis B[1], Negeri Z[1], Sun Y[1], Benedetti A[1], Thombs BD[1]

[1] McGill University, Canada

Background: The Edinburgh Postnatal Depression Scale (EPDS) is recommended for depression screening in pregnancy and postpartum. Cut-offs of ≥ 10 or ≥ 13 are commonly used to detect possible depression, but the only previous meta-analysis found that a cut-off of ≥ 12 maximized combined sensitivity and specificity. That meta-analysis, however, was conducted over 10 years ago and was limited by a small number of included studies, by incomplete cut-off reporting within included studies, and by not examining accuracy by reference standard or participant subgroups, including pregnancy versus postpartum. Individual participant data meta-analysis (IPDMA), which synthesizes participant-level data from primary studies rather than summary results from study reports, has the potential to overcome these challenges.

Objectives: To evaluate EPDS accuracy for screening to detect major depression in pregnancy and postpartum using IPDMA.

Methods: We searched MEDLINE, MEDLINE In-Process & Other Non-Indexed Citations, PsycINFO, and Web of Science from inception to 3 October 2018 for datasets that compared EPDS scores to major depression classification based on a validated diagnostic interview. We used bivariate random-effects meta-analysis to estimate EPDS sensitivity and specificity compared to semi-structured, fully structured (Mini International Neuropsychiatric Interview [MINI] excluded), and MINI diagnostic interviews, separately, using individual participant data. We used one-stage meta-regression to examine accuracy by reference standard categories and participant characteristics (age, pregnant versus postpartum status, and country human development index).

Results: We obtained individual participant data from 58 of 83 eligible studies (15,557 participants, 2069 major depression cases). EPDS ≥ 11 maximized combined sensitivity and specificity (81%, 88%). For commonly used cut-offs, sensitivity and specificity were 85% and 84% for EPDS ≥ 10 and 66% and 95% for EPDS ≥ 13. Accuracy was similar across reference standards and subgroups, including for women in pregnancy and postpartum.

Conclusions: An EPDS cut-off of ≥ 11 maximized combined sensitivity and specificity; a cut-off of ≥ 13 was less sensitive but more specific. To identify women in pregnancy and postpartum with higher symptom levels, a cut-off of 13 or greater could be used. Lower cut-offs could be used if the intention is to avoid false negatives and identify most patients who meet diagnostic criteria.

Patient or healthcare consumer involvement: There was no direct patient or consumer involvement in this project. However, clinicians considering screening for depression with the EPDS can refer to our web-based knowledge translation tool: depressionscreening100.com/epds, which estimates expected numbers of positive screens and true and false screening outcomes based on results from the present IPDMA.

Screening for depression with the Patient Health Questionnaire-2 (PHQ-2) alone and in combination with the PHQ-9: individual participant data meta-analysis

Levis B[1], Sun Y[1], He C[1], Wu Y[1], Krishnan A[1], Bhandari PM[1], Neupane D[1], Imran M[1], Brehaut E[1], Negeri ZF[1], Fischer FH[2], Benedetti A[1], Thombs BD[1]

[1] McGill University, Canada; [2] Charité – Universitätsmedizin Berlin, Germany

Background: The Patient Health Questionnaire-2 (PHQ-2) depression screening tool includes items that assess frequency of depressed mood and anhedonia in the past two weeks. It can be used alone or as a first step to identify patients for evaluation with the full Patient Health Questionnaire-9 (PHQ-9). Meta-analyses on PHQ-2 accuracy have been limited by including only published data and by not examining accuracy for different reference standards, in participant subgroups, or in combination with the PHQ-9, as it is commonly used in practice. Individual participant data meta-analysis (IPDMA), which synthesizes participant-level data from primary studies rather than summary results from study reports, has the potential to overcome these challenges.

Objectives: To use IPDMA to evaluate the accuracy of the PHQ-2 alone and in combination with the PHQ-9 for screening to detect major depression.

Methods: We searched MEDLINE, MEDLINE In-Process & Other Non-Indexed Citations, PsycINFO, and Web of Science from 1 Jan 2000 to 9 May 2018 for datasets that compared PHQ scores to major depression classification based on a validated diagnostic interview. We used bivariate random-effects meta-analysis to estimate sensitivity and specificity compared to semi-structured, fully structured (Mini International Neuropsychiatric Interview [MINI] excluded), and MINI diagnostic interviews, separately, and in participant subgroups based on age, sex, country human development index and recruitment setting.

Results: We obtained individual participant data from 100 of 136 eligible studies (44,318 participants, 4572 major depression cases). Among studies that used semi-structured interviews, PHQ-2 sensitivity and specificity were 0.91 and 0.67 for cut-off ≥ 2 and 0.72 and 0.85 for cut-off ≥ 3. Sensitivity was significantly greater for semi-structured versus fully structured interviews. Specificity was not significantly different across interviews. There were no significant differences in accuracy across subgroups. For semi-structured interviews, sensitivity for PHQ-2 ≥ 2 followed by PHQ-9 ≥ 10 was not significantly different than for PHQ-9 ≥ 10 alone (0.82 versus 0.86); specificity was significantly but minimally higher (0.87 versus 0.85). The combination reduced the number of participants needing to complete the full PHQ-9 by 57%.

Conclusions: PHQ-2 ≥ 2 followed by PHQ-9 ≥ 10 had similar accuracy as PHQ-9 ≥ 10 alone and reduced the proportion of participants needing to complete the full PHQ-9 by 57%.

Patient or healthcare consumer involvement: There was no direct patient or consumer involvement in this project. However, clinicians considering screening for depression with the PHQ alone or in combination with the PHQ-9 can refer to our web-based knowledge translation tool: depressionscreening100.com/phq-2, which estimates expected numbers of positive screens and true and false screening outcomes based on results from the present IPDMA.

Step-by-step guidance on conducting prospective meta-analyses

Seidler AL[1], Hunter K[1], Cheyne S[1], Ghersi D[2], Berlin J[3], Askie L[1]

[1] National Health and Medical Research Council Clinical Trials Centre, University of Sydney, Australia; [2] National Health and Medical Research Council, Australia; [3] Johnson & Johnson, USA

Background: Prospective meta-analyses (PMA) can reduce many of the problems associated with traditional meta-analyses by specifying study selection criteria, hypotheses and analyses before the results of included studies are known. This can reduce risk of publication bias and selective outcome reporting, and enables researchers to harmonise their ongoing research efforts to answer important questions with greater certainty. However, despite rising numbers of PMA, the terminology and definitions used to date have lacked clarity and consistency, and there is little guidance on how to conduct PMA. This threatens successful implementation of PMA.

Objective: To develop step-by-step guidance on how to conduct PMA.

Methods: We, the Cochrane PMA Methods Group, developed step-by-step guidance based on:

1) a scoping review of methodology papers;
2) a scoping review of existing PMA;
3) expert opinions;
4) experiences with previous PMA.

We illustrate each step with a recent case study.

Results: We describe seven steps for PMA (Figure). First, a protocol needs to be developed, including details on collaboration policies (Step 1). Next, a systematic search for planned/ongoing studies should be conducted, including a search of trial registries, medical databases and contacting stakeholders (Step 2) and eligible studies need to be identified for inclusion (Step 3). Importantly, only studies for which the results are not known can be included in a PMA. These studies are then invited to form a collaboration (Step 4), ideally including a steering and data analysis committee, with representatives from each study. Next, core outcomes, common intervention features and a statistical analysis plan are agreed upon within the collaboration (Step 5). This can be particularly useful for rare but important outcomes such as adverse side effects, that individual studies would not have enough power to test statistically. There is usually a waiting period while all studies are being completed, before the evidence is synthesized. Certainty of evidence is assessed by adapting tools such as GRADE (Step 6). Results should be reported using adapted versions of reporting tools such as PRISMA (Step 7).

Conclusion: PMA allow for greatly improved use of data, while reducing bias and research waste. PMA could be integral to rapid learning health systems since evidence gaps are identified, and ongoing studies are initiated, tracked and meta-analysed as soon as their results become available, which can then inform policy and practice. Adaptive trial methodology can be used to adapt ongoing PMA to emerging evidence. With rising trial registration compliance and new technical advances in machine learning and data processing, we see new horizons for PMA. This step-by-step guidance will improve the understanding of PMA in the research community and enable more researchers to conduct successful PMA.

Patient or healthcare consumer involvement: We will invite healthcare consumers to comment on this research project, to increase its accessibility from their perspective.

Additional file: Figure

Systematic reviews of case reports and patient level analysis as a new tool to study rare adverse events of medications: the case of neuroleptic malignant syndrome

Gastaldon C[1], Guinart D[2], Misawa F[3], Rubio JM[2], de Filippis R[4], Schoretsanitis G[2], Barbui C[1], Kane JM[2], Correll CU[5]

[1] WHO Collaborating Centre for Research and Training in Mental Health and Service Evaluation Department of Neuroscience, Biomedicine and Movement Sciences Section of Psychiatry, University of Verona, Italy; [2] Division of Psychiatry Research, Zucker Hillside Hospital, Northwell Health, New York; Center for Psychiatric Neuroscience, Feinstein Institute for Medical Research, Manhasset; Department of Psychiatry, Zucker School of Medicine at Nort., USA; [3] Yamanashi Prefectural KITA Hospital, Yamanashi, Japan; [4] Psychiatric Unit, Department of Health Sciences, University Magna Graecia of Catanzaro, Catanzaro, Italy; [5] Division of Psychiatry Research, Zucker Hillside Hospital, Northwell Health, New York; Department of Child and Adolescent Psychiatry, Charité-Universitätsmedizin Berlin, Berlin, USA; Germany

Background: Neuroleptic malignant syndrome (NMS) is a rare, potentially fatal condition associated with dopamine receptor blocking agents, such as antipsychotics (APs). Due to its relatively low incidence (~1% of antipsychotic users), epidemiological and experimental studies are almost absent. High severity low incidence conditions, such as NMS, may therefore require alternative research approaches. Using NMS as a paradigm, we propose to conduct a systematic review and assessment of published case reports, supplemented by an individual patient case report meta-analysis.

Objectives: To pilot the feasibility of developing a patient-level database based on a systematic literature review of NMS case reports/series, and to explore the applicability of individual patient case report data analytical approaches. In particular, we aim to address association of NMS with specific medications, and the clinical characteristics and outcomes of NMS.

Methods: At least two authors conducted a systematic and independent search in MEDLINE, Embase, Cochrane, CINAHL and PsychINFO databases. We included all case reports describing NMS that occurred during ongoing AP treatment or within one injection interval of a long-acting injection (LAI) in adults. Each included case report was reviewed and extracted by at least two physicians independently to extract demographic, clinical, treatment and outcome data. NMS severity was coded using the Francis-Yacoub scale.

Results: The database was created 10/2018 and last updated 01/2020. To date, it includes 690 case reports of NMS. Clinical characteristics and outcomes of NMS (frequency of complete recovery, incomplete recovery or death, duration of NMS, and length of hospital stay) were compared between different types of AP formulations and class as well as between monotherapy and polytherapy. The analyses were adjusted for between-group differences and potential confounders in a multivariable regression model.

Conclusions: To our knowledge, this is the first systematic review and individual patient case report meta-analysis of a severe rare adverse event, such as NMS. Since NMS is so infrequent, it is unlikely to observe a significant number of cases in alternative designs such as epidemiological or experimental studies. Hence, our approach provided more information than what could have possibly been made available by epidemiological studies. For example, the largest 11-year longitudinal register case-control study of NMS (based on the Psychiatric Danish Register) gathered only 83 cases of NMS. We retrieved 690 cases and compared characteristics and outcomes of NMS occurring during treatment with different AP classes and formulations. Therefore, the results from our study are clinically relevant, innovative and could be one of the few possible strategies to generate relevant clinical and treatment evidence for rare adverse events. Individual patient case report meta-analyses, although sensitive to reporting bias, can arguably be informative to researchers and clinicians, and can help guide the design of future research studies.

The 22nd anniversary of the Cochrane Back and Neck Group

Furlan A[1], Harbin S[1], Pardo Pardo J[2], Chou R[3]

[1] Cochrane Back and Neck Group, Canada; [2] Cochrane Musculoskeletal group, Canada; [3] Cochrane Back and Neck Group, USA

Background: Low back pain (LBP) affects 80% of people at some time in their lives. The Cochrane Back and Neck (CBN) Group has been housed in Toronto at the Institute for Work & Health since 1996 and has 85 reviews and 32 protocols published in the Cochrane Library. With the ending of the external funding in 2015, the CBN has had to find another institution to continue its activities.

Objectives: To provide an update of CBN activities.

Methods and Results: In the past three years, CBN conducted priority setting with organizations that develop clinical practice guidelines for LBP. CBN editors and associate editors published key methodological articles in the field of back and neck pain research. The methodological quality of CBN reviews has been assessed by external groups in a variety of areas, and the conclusions were that CBN reviews had higher methodological quality than non-Cochrane reviews. CBN reviews are included in 35 clinical practice guidelines for back and neck conditions. The 2018 journal impact factor of CBN is 11.154, which is higher than the 2018 impact factor for the Cochrane Database of Systematic Reviews (7.755). CBN reviews ranked fourth among all 53 Cochrane review groups in terms of Cochrane Library usage data. The most accessed CBN review was "Yoga treatment for chronic non-specific low-back pain" which had 9689 full-text downloads. CBN is active on Twitter and has 3958 followers.

Conclusions: As of 1 April 2020, due to a continued funding shortfall, the editorial base of CBN will be transferred to the Cochrane Musculoskeletal Group in Australia. CBN published many systematic reviews and made important methodological contributions to the field of spine research over the past 22 years of work with Cochrane.

Patient or healthcare consumer involvement: The Cochrane Back and Neck (CBN) Group has an editorial board composed of internationally renowned scientists, clinicians and consumers.

The research progress of open-label placebo and implications for future studies: a systematic survey

Ma T[1], Zhang Y[2], Zi M[3], Fang H[4], Jin Q[5], Lu C[2], Yang M[2], Liu J[2]

[1] Beijing University of Chinese Medicine, Xiyuan Hospital, Beijing, China; [2] Center for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine; Beijing GRADE Center, Institute for Excellence in Evidence-Based Chinese Medicine, Beijing, China; [3] Xiyuan Hospital, First Clinical Medical College of China Academy of Chinese Medical Sciences, China; [4] Department of Traditional Chinese Medicine Pulmonary Diseases, China-Japan Friendship Hospital, China; [5] Beijing University of Chinese Medicine, Dongzhimen Hospital, Beijing, China

Background: Single- or double-blinding techniques are often used in randomized controlled trials (RCTs). However, post-1960s transparency and honesty in research became a concern from the ethical perspective. Several trials introduced open-label placebo (OLP) in trial design. The participants who were allocated to the placebo group had been clearly informed and definitely knew what they were taking, i.e. OLP design. Certain therapeutic effects were found in the OLP group and a growing number of trials covering quite different disease types showed similar results.

Objectives: This study aims to systematically review OLP trials, and explore the P (patients), I (interventions), C (comparisons), O (outcomes), and Setting etc., meanwhile summarize methodological aspects for OLP design. Afterwards, we will discuss the potential use of OLP for future research, especially traditional Chinese medicine (TCM).

Methods: We performed systematic searches of CNKI and PUBMED (from its inception to 31 December 2019). We will include original articles only. We summarized the characteristics and the methodological elements.

Results: We retrieved 34 RCT trials, including six cross-over RCTs and eight parallel ones. Among them, 23 explicitly mentioned the process of informed consent. The sample size ranged from 9 to 539, and there were 32 experiments with a sample size of less than 160 people. Thirty-three trials were observed for at most three months, and only one study lasted for a year. Thirteen trials found the positive effects of OLP, while two failed to show the effectiveness. Nineteen trials didn't give any conclusive results about the effects as OLP was just one of the comparison groups. A total of 21 trials were conducted on diseases with symptoms such as depression, pain, itching, test anxiety, rhinitis subjective symptoms etc.

Conclusions: So far, most of the OLP design still mainly focuses on the short-term and subjective outcomes with relatively small sample size. OLP might be effective for a variety of diseases with temporarily no specific drugs and no side effects, which can avoid the nocebo phenomenon in controlled trials and ethically guarantee the right of subjects to know. It could be a new advance in the field of placebos. The mechanism of OLP positive reaction may be related to positive hints, doctor–patient relationship, prediction and error processing (PEP), neurobiological factors, spontaneous remission, regression mean effect, Hawthorne effect, etc. For RCTs of TCM, doubling blinding is always challenged by preparing perfect stimulator for placebo. Besides, some argue that TCM lacks specific effects beyond placebo effects. In these circumstances, OLP could be one solution.

Patient or healthcare consumer involvement: No patients participated in this study. The OLP conceptions relates to patients' welfare.

Using overview of systematic reviews to identify topics for health systems living reviews

Mansilla C[1], Lavis J[1]

[1] McMaster University, Canada

Background: Rapid-learning health systems need to rely strongly on existing research evidence to make decisions about health system design, implementation and planning. Living reviews offer an outstanding opportunity to keep evidence up to date, which is a key issue for making decisions about health systems. Despite the relevance of having living evidence to make decisions, no living review on health systems research has been published.

Objectives: To summarize the existing evidence syntheses describing and evaluating a health system topic, in order to identify a topic for a living health systems review.

Methods: We conducted an overview of systematic reviews on payment mechanisms to organizations. We selected this topic because of the relevance of financial arrangements for health systems research, as well as the number of existing evidence syntheses already published. We searched four databases as well as grey literature. We extracted data to characterize the existing literature and used explicit criteria to assess quality. Finally, each one of the topics was assessed to analyze its potential to conduct a living review. Based on the existing publications of the Living Systematic Review network, we categorize each topic across the three criteria defined to when a living review is appropriate (namely a priority for decision making, certainty in the existing evidence, and the likelihood that new evidence becomes available).

Results: The preliminary results of this overview show that the main payment mechanisms used to fund organizations (fee-for-service, capitation, case-based and global budgets) are a priority for decision-making, and its certainty in the existing evidence is generally low or very low. However, different payment mechanisms differ in the probability that new evidence could be published, changing the existing conclusion.

Conclusions: An essential component of rapid-learning health systems is to count on living evidence to make constant transformations. This presentation shows a concrete method to identify topics where living health systems reviews might be needed.

Patient or healthcare consumer involvement: To incorporate consumers and patients' perspectives, the results of this overview of systematic reviews need to be presented to relevant stakeholders to receive important feedback. This brings special insights on how health systems living reviews can be an essential tool for rapid-learning health systems.

Using the Response-adaptive Randomization (RAR) platform to explore potentially effective treatments during the outbreak of COVID-19 in China

Jin Q[1], Zhang Y[2], Jiang F[3], Wang J[4], Lu C[2], Yang M[2], Liu J[2]

[1] Dongzhimen Hospital, First Clinical Medical College of Beijing University of Chinese Medicine, China; [2] Center for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, Beijing GRADE Center, Institute for Excellence in Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, China; [3] Department of Science and Technology, Ruikang Hospital of Guangxi University of Chinese Medicine, China; [4] National Institute of TCM Constitution and Preventive Medicine, Beijing University of Chinese Medicine, China

Background: COVID-19 has spread all over the world and become a major public health issue since the end of 2019. Although some patients have recovered, this novel coronavirus still showed strong infectivity and, for some patients, rapid progress and high fatality. There are currently no specific drugs that target the virus. First-line clinicians found that some generic drugs that focused on the host symptoms may be helpful to treat patients with moderate or severe illness. Many clinical trials were registered and have undertaken to evaluate the effectiveness of different interventions. The Response-adaptive randomization (RAR) platform allows the investigation of multiple treatments within the same time period and the screening of optimal treatment(s) during the trial.

Objectives: To explore the feasibility and applicability of the RAR platform in clinical trials of COVID-19.

Methods: We searched ClinicalTrials.gov (clinicaltrials.gov) and the Chinese clinical trial Registry (www.chictr.org.cn). We included all trials of COVID-19 registered before 6 March 2020. We developed a design plan for the RAR platform in co-operation with an IT corporation.

Results: We identified 328 trials. Interventions included antiviral drugs such as Remdesivir and Arbidol, hormones such as glucocorticoid, convalescent plasma, and traditional Chinese medicine such as decoction or patent medicine. Most of these drugs are generic drugs tested by previous safety data. The primary outcomes focused on clinical symptoms, progress rate, hospital discharge rate or conversion rate of viral nucleic acid, etc. Most of the primary outcomes can be obtained within 14 days. So far, investigators were not sure which drugs would bring benefit to their patients. Their decisions may be based on the clinical outcomes of a given treatment. All these conditions make the application of RAR platform possible.

Conclusions: There are currently no specific and effective treatments for COVID-19. Most of the registered clinical trials involve the selection of multiple treatment regimens or multiple doses. Based on the interactive web-based randomization system (IWRS), the application of RAR makes it possible to evaluate the effectiveness and safety of multiple treatments in a clinical trial at the same time. Different roles in the platform will include principal investigator, sub-investigator, data manager, drug dispensers, statistician and trial manager etc. However, the application of RAR requires that the speed of the primary outcome responses is relatively faster than that of the patient enrolment, which remains a major challenge.

Patient or healthcare consumer involvement: There was no patient or consumer participation in these research activities. The proposed platform is designed to fulfil the goals of maximum benefits for patients by screening drugs simultaneously.

What is the most efficient de-duplication software for use in systematic reviews?

Jubb A[1], Carr E[1], Sanderson A[1], Baragula E[1], McCool R[1], Glanville J[1]

[1] York Health Economics Consortium, UK

Background: Systematic reviews search several databases. Search results then need to be de-duplicated. Generally, review authors de-duplicate using bibliographic software. Using software with the most efficient automatic de-duplication algorithm would save time.

Objectives: To identify the most efficient de-duplication options from six commonly used and/or free of charge software packages. Our criteria for efficiency was 100% specificity and the highest possible sensitivity. 100% specificity is critical to avoid potentially relevant studies being incorrectly de-duplicated – we can compensate for poor sensitivity by manual de-duplication.

Methods: We manually de-duplicated the search results of two literature searches (1578 mammography records and 2239 atopic dermatitis records) within Endnote X9 to produce two sets of known duplicates against which to compare the performance of the selected software. We ran these results through the automatic de-duplication options available in Citavi, Endnote, EPPI-Reviewer 5 (beta), Mendeley, the Systematic Review Assistant-DeDuplication Module (SRA-DM) and Zotero. We calculated the specificity and sensitivity of the software.

Results: The results are in Tables 1 and 2. According to our criteria, the most efficient software is Endnote (100% specificity). Mendeley and Citavi exceeded Endnote's sensitivity in both result sets, but had lower specificity in the atopic dermatitis set (99.8% and 99.3% respectively). SRA-DM was less sensitive and less specific than Endnote, and EPPI-Reviewer did not achieve 100% specificity in either record set.

Conclusions: Endnote was the most efficient software package according to our specifications with 100% specificity. Endnote's slightly lower sensitivity compared to some other software can be redressed with manual de-duplication, whereas records incorrectly de-duplicated by other software may not be easily recovered. Zotero seemed the least safe option, de-duplicating all records from the same abstract book. Efficiency is not the only benefit of de-duplication software. When sensitivity is lower than 100% manual de-duplication is also required. However, the relative advantages of conducting manual de-duplication in the different software packages was beyond our scope. Ancillary approaches to manual de-duplication such as published algorithms were not assessed for this project. Further, we did not assess software features that mitigate low specificity (e.g. Citavi's manual review of identified duplicates) since they introduced a second step to de-deduplication. Finally, each software has default de-duplication algorithms, and we did not assess the impact of choosing a different default, where available. Due to resource constraints we only used two test sets, which means generalisability may be impacted. A larger number of result sets would help with generalisability.

Patient or healthcare consumer involvement: This research was methods focused and no patients were involved. We are the consumers of the software products.

Additional files: Tables

LIVING META-ANALYSIS

Addressing interdependency of data when conducting systematic reviews and meta-analyses

Lietz M[1], Sieben W[1], Sauerland S[1]

[1] Institute for Quality and Efficiency in Healthcare, Germany

Background: Some clinical study designs require that interdependency of data (IoD) be addressed in the reporting of study results. In self-controlled and cross-over trials, patients form part of both the intervention and control groups. In other studies, multiple disease locations are measured in the same patient. Because clinical studies often fail to account for IoD, one must apply specific statistical methods to the data of those studies when conducting a systematic review (SR) with a meta-analysis (MA). Otherwise, the results may be misleading.

Objectives: This presentation describes different options for addressing and solving problems with IoD in MAs.

Methods: We performed three SRs with MAs in medical fields where data clustering is common: ophthalmology (paired data on both eyes), dentistry (multiple data on single teeth), and sleep medicine (paired cross-over data in patients with chronic diseases). Before conducting the MAs, we assessed the studies included for IoD and corrected the data when IoD adjustment was necessary. If possible, we adjusted the results by using a correlation coefficient (CC) for the standard deviation of the effect estimates of the studies included.

Results: In the ophthalmology review (on keratokonus therapy), only 3 of 7 randomized controlled trials (RCTs) adjusted for IoD. Obtaining individual patient data (IPD) from a fourth RCT with a large data set allowed for IoD adjustment, so the MA included four rather than only three RCTs. In the dentistry review (on periodontitis), several RCTs erroneously inflated the sample size by entering up to more than 100 values per patient when calculating group means, instead of aggregating the measurements and entering a single value per patient. Using a CC based on data from two epidemiological studies, we adjusted the results for the standard deviation of the effect estimates. In the sleep medicine review (on therapy for obstructive sleep apnea), correction for IoD was made for studies with a cross-over design. Only a few studies had properly accounted for IoD. Using their data, we estimated the CCs for the MAs and conducted sensitivity analyses with the set of CCs to test the robustness of our results.

Conclusions: There are several ways to solve problems with IoD when performing MAs. Ideally, IPD are available for re-analysis of study results and estimation of CCs. Alternatively, CCs from other studies included in the MA can be used to adjust the results of studies with IoD problems. Finally, study results based on multiple measurements per patient can be adjusted by applying CCs from other studies, even if these studies are not included in the SR.

Patient or healthcare consumer involvement: For the described methodological procedures regarding meta-analyses, patient or healthcare consumer involvement is not applicable.

Conducting multiple living systematic reviews rapidly without compromising quality is possible with an innovative approach: L·OVE COVID-19 initiative

Rada G[1], Acuña MP[2], Alvares C[3], Araneda G[4], Ávila C[5], Baladia E[6], Bohorquez S[7], Bravo R[8], Carrasco C[9], Carvajal-Juliá N[10], Chahuan J[11], Corbalán J[1], Pérez M[5], Pérez Gaxiola G[12], Pérez-Bracchiglione J[9], Pesce F[13], Pizarro AB[14], Poloni D[15], Prieto P[16], Ragusa M[17], Rojas MX[18], Santillan-Garcia A[19], Sepúlveda J[5], Torres López LA[18], Urrea G[10], Vargas M[9], Verdejo C[10], Verdugo F[5], Vergara Merino L[10], Villar J[18], Cuadrado C[20], Ferrada C[20], Peña E[16], Ortiz L[21], Olguin P[10], Ojeda P[21], Neumann J[16], Morel-Marambio M[21], Meza N[9], Madrid E[9], Izcovich A[17], González C[16], Goez-Mogollon L[18], Gempeler A[22], Garnham R[10], Franco JVA[23], Flores I[24], Escobar-Liquitay C[23]

[1] Cochrane Chile, Chile; [2] Unidad de Infectología, Hospital Dr Sótero del Río, Santiago, Chile; [3] Fundación Sanitas, Colombia; [4] Escuela de Salud Publica Universidad de Chile, Chile; [5] Epistemonikos Foundation, Chile; [6] Red de Nutrición Basada en la Evidencia, Academia Española de Nutrición y Dietética, Pamplona, Spain; [7] Universidad de Jaen, Spain; [8] UC Evidence Center, Cochrane Chile Associated Center, Pontificia Universidad Católica de Chile, Chile; [9] Interdisciplinary Centre for Health Studies (CIESAL), Cochrane Associate Centre, School of Medicine, Universidad de Valparaíso, Chile; [10] School of Medicine, Universidad de Valparaíso, Chile; [11] Pontificia Universidad Católica de Chile, Chile; [12] Hospital pediátrico de Sinaloa, México; [13] Living Knowledge, Chile; [14] Facultad de Enfermería, Pontificia Universidad Javeriana, Coordinadora de red de consumidores Cochrane Colombia, Colombia; [15] Department of Internal Medicine, School of Medicine, Universidad de Valparaíso, Chile; [16] Escuela de Salud Publica, Universidad de Chile, Chile; [17] Servicio de medicina interna del Hospital Alemán, Argentina; [18] Departamento de investigaciones, Fundación Cardioinfantil-IC, Colombia; [19] Hospital Universitario de Burgos, Spain; [20] School of Public Health, Universidad de Chile, Chile; [21] Centro Evidencia UC, Centro asociado Cochrane Chile, Chile; [22] Centro de Investigaciones Clínicas, Fundación Valle del Lili, Cali, Colombia; [23] Departamento de Investigación – Instituto Universitario Hospital Italiano, Argentina; [24] Universidad de Antioquia, Centro Asociado Cochrane Colombia, Colombia

Background: The evidence about COVID-19 is being produced at high speed, so it is very difficult for decision-makers to keep up. It seems appropriate, then, to put into practice a novel approach able to provide the scientific community and other interested parties with up-to-date, quality evidence that is actionable, and rapidly and efficiently produced.

Objectives: To systematically assess the evidence for multiple questions relevant to COVID-19, and to update them using a living approach.

Methods: We designed a protocol for multiple parallel living systematic reviews in line with PRISMA. We searched for primary studies which answer different questions related to COVID-19 using both a centralized repository (Epistemonikos database) and a manual search in MEDLINE, Embase and the Cochrane Central Register of Controlled Trials. We also searched for literature in several other sources, including grey literature and trial registries. All the evidence was organized in an open platform (L·OVE – Living Overview of Evidence) that is continuously updated using artificial intelligence and a broad network of experts. At least two researchers independently selected studies, extracted data, and assessed the risk of bias of included studies. We synthesized data for each question using meta-analysis, when possible, and prepared 'Summary of findings' tables according to the GRADE approach.

Results: We compiled a list of questions by liaising with local stakeholders and consulting with clinical experts. We set a team of 56 researchers from 14 organizations, who selected their questions according to their areas of expertise. We established a central team composed of methods experts, software engineers, information specialists, project managers, professional writers and journalists. A common protocol was written for all the reviews, and individual protocols were adapted to each individual review and made public. We released a short article in plain language and a preliminary report of the review as soon as all the data had been analysed. These were widely disseminated through social networks and sent to relevant authorities. Then, we submitted the full review for publication. Twenty days after the kickoff of the project, at the moment of submission of this abstract, 17 systematic reviews have been initiated. Six are already completed and six are finishing data extraction. Most of the reviews have needed reassessment of new evidence after completing the initial screening, and one review has needed

two updates since its release. It is likely that many more reviews will be initiated in the next weeks as the COVID-19 pandemic evolves.

Conclusions: A production model of multiple living systematic reviews in the same topic is feasible with a large team of researchers, a central co-ordinating team, and the appropriate technological tools to streamline and manage the process.

Patient or healthcare consumer involvement: No

Development and pilot of a framework using automation and crowd-sourcing to identify and classify randomized controlled trials for rheumatoid arthritis drug therapy

Kamso MM[1], Thomas M[1], Lee C[2], Ejaredar M[1], Whittle S[3], Buchbinder R[4], Deardon R[1], Tugwell P[5], Pardo J[5], Kelly S[5], Wells G[5], Hazlewood G[1]

[1] University of Calgary, Canada; [2] University of Alberta, Canada; [3] University of Adelaide, Australia; [4] Monash University, Australia; [5] University of Ottawa, Canada

Background: The treatment landscape for rheumatoid arthritis (RA) continues to evolve, and timely, high-quality evidence syntheses are of high interest to clinicians, patients, and policy-makers. To accomplish this, rapid approaches for identifying and classifying trials are required.

Objectives: To develop and pilot a novel approach combining automation and crowd-sourcing to identify and classify all randomized controlled trials (RCTs) of disease-modifying anti-rheumatic drugs (DMARDs) for RA to inform living network meta-analyses (LNMA).

Methods: We searched MEDLINE, Embase and the Cochrane Central Register of Controlled Trials (CENTRAL) using filters for "RA" and "RCTs". The search results were then uploaded to an RCT Classifier (Wallace 2017), available in the Cochrane Register of Studies (CRS-Web), which uses machine learning algorithms to assign a probability of each citation being a true RCT. Citations with a < 1% probability of being an RCT were excluded. The remaining records were then uploaded to Cochrane Crowd, a citizen science platform, to further exclude non-RCTs. From here, we uploaded records to a custom online Cochrane PICO annotator tool we developed for this project. Trained reviewers were asked a series of questions to first confirm the record was an RCT in patients with rheumatoid arthritis, then to classify the intervention (DMARD or non-DMARD). Reviewers were trainees and invited through an invitation from authors, or relevant organizations (e.g. trainees in undergraduate or graduate programs; clinical trainees through rheumatology associations).

Results: After removing duplicates, the literature search identified 32,068 records, of which 14,682 were excluded through the RCT classifier (probability RCT < 1%). Of the remaining 17,206 records, a further 8003 were excluded through Cochrane Crowd for not being an RCT. To date, 2111 of the 9203 potentially eligible records have received at least 2 judgements through the custom PICO annotator tool. Of the 2111 classified records, 560 (26.53%) were excluded as all reviewers agreed that the study was either not an RCT (n = 232), not in RA (n = 120) or not a DMARD (n = 208). The remaining 1373 (64%) were either rated as 'unclear' for 1 or more questions by all reviewers (n = 62) or had a disagreement between reviewers regarding eligibility (n = 1311). Disagreements were between unclear and 'yes' or unclear and 'no' for 710 records, and between 'no' and 'yes' in 601. Amongst the 601 'yes'/'no' disagreements, 384 occurred for assigning the record as an RCT, 59 for assigning the population as RA, and 158 for assigning the intervention as a DMARD.

Conclusions: An approach combining automation and crowdsourcing is a promising method for rapidly identifying and classifying RCTs for RA. Further work is required to validate the approach and investigate methods to reduce disagreements between reviewers.

Patient or healthcare consumer involvement: Identifying and classifying trials for RA treatments are needed for quality evidence synthesis for patients and clinicians to make informed choices.

Glycated hemoglobin and risk of mortality in dialysis patients with diabetes mellitus: a systematic review and dose–response meta-analysis

Wang W[1], He Q[1], Sun X[1]

[1] Chinese Evidence-based Medicine Centre and CREAT Group, State Key Laboratory of Biotherapy, West China Hospital, Sichuan University and Collaborative Innovation Centre, Chengdu, China, China

Background: Evidence suggested that either excessively high or low HbA1c level could result in adverse outcomes. The non-linear relationship between HbA1c level and adverse outcomes among patients receiving dialysis, however, has not been established.

Objectives: To conduct a dose–response meta-analysis aiming to explore if there is a non-linear relation between HbA1c level and mortality in diabetic patients receiving dialysis.

Methods: We searched PubMed, Embase and the Cochrane Central Register of Controlled Trials (CENTRAL) from inception to November 2018. Eligible studies were randomized controlled trials (RCTs), non-RCT and observational studies that assessed the relationship between HbA1c level and mortality in diabetic patients receiving dialysis. We used a modified version of Cochrane's tool and the Newcastle-Ottawa Scale (NOS) to assess risk of bias for RCTs and observational studies. We performed a dose–response meta-analysis to investigate the possible relationship between HbA1c level and mortality in dialysis patients with DM. We used adjusted hazard ratio (HR) as the effect measure and a one-stage robust error meta-regression (REMR) model to fit the potential non-linear trend.

Results: We included 19 studies involving 113,119 participants in the data analysis, of which, 13 were prospective cohort studies, 5 were retrospective cohort studies, and the other one was an RCT. All the 18 cohort studies selected participants from the same population, had confident in ascertaining exposure and control, and well adjusted for prognostic factors. Fifteen studies had adequately followed up and only five studies reported similar co-intervention between groups. The one RCT adequately generated their randomization sequence, concealed treatment allocation, blinded participants, caregivers and outcome assessors, and free from selective reporting. The dose–response meta-analysis analysis showed J-shaped association between HbA1c and mortality ($P < 0.05$ for non-linear test). The HbA1c level at about 7% had the lowest all-cause mortality. Both low and high HbA1c level were associated with increased risk of all-cause mortality compared with the reference group of 7%: (HbA1c at 5%: HR 1.03, 95% CI 0.99 to 1.07; HbA1c at 6%: HR 1.01, 95% CI 0.99 to 1.02; HbA1c at 8%: HR 1.02, 95% CI 1.01 to 1.02; HbA1c at 9%: HR 1.05, 95% CI 1.04 to 1.06; HbA1c at 10%: HR 1.10, 95% CI 1.08 to 1.13; HbA1c at 11%: HR 1.15, 95% CI 1.09 to 1.20). Fourteen studies investigated the association between the HbA1c level and all-cause mortality in hemodialysis patients, and the dose–response analysis also showed a J-shape association between HbA1c level and mortality, and HbA1c at 7% had the lowest mortality.

Conclusions: Current evidence suggested a J-shaped relationship between HbA1c level and mortality in diabetic patients with dialysis, and patients with HbA1c at about 7% had the lowest mortality.

Living systematic reviews and meta-analysis: when, why, and how do we need to use them?

Yu X[1], Zhang H[2]

[1] WHO Collaborating Centre for Guideline Implementation and Knowledge Translation, China; [2] Evidence-Based Medicine Center, Lanzhou University, China

Background: In the current systematic review and meta-analysis, there are many alternative review methods, such as scoping review, evidence map, rapid review and so on. Living systematic review is another one and it could be continuously updated and more efficient. However, when it is suitable to use a living systematic review has become an important topic that needs further study and discussion.

Objectives: Through the evidence synthesis of current living systematic reviews, to find the current use situations and use rules. To find the difference between living systematic reviews and other alternative reviews.

Methods: We searched six electronic databases (MEDLINE via PubMed, Embase, Web of Science, CBM, CNKI, Wanfang) and retrieved all Chinese and English literature related to living systematic review and meta-analysis published from inception to 31 March 2020. We also searched Google Scholar, Baidu Xueshu, gray literature, and references included in the studies. The research team is an interdisciplinary research team. Two review authors independently screened literature according to the inclusion-exclusion criteria formulated in advance and consulted a third party when there was a dispute.

Results: The results will be presented at the Cochrane Colloquium.

Conclusions: Through our review, we can find the status of current living systematic reviews, the difference between living systematic review and other reviews, and understand how to use living systematic reviews.

Patient or healthcare consumer involvement: None.

Maintaining currency in the rapidly evolving world of diabetes research – the living evidence approach

White H[1], Tendal B[1], Millard T[1], Elliott J[1]

[1] School of Public Health and Preventative Medicine, Monash University, Australia

Background: Living systematic reviews represent a novel approach to maintaining currency in the world of rapidly evolving evidence. As part of the broader Australian Living Evidence Consortium, the Living Evidence for Diabetes program is focused on developing living systematic reviews and associated clinical guidelines within two priority areas of diabetes care: the use of medical device technology in type 1 diabetes, and therapeutics for blood glucose control in type 2 diabetes.

Objectives: To develop and implement a system in which two systematic reviews, developed as part of the Living Evidence for Diabetes demonstration project, are rapidly updated through the integration of new and relevant evidence.

Methods: Traditional systematic reviews were developed for each of the priority topics. Implementation of a living system involves monthly searching of PubMed, Embase and the Cochrane Central Register of Controlled Trials, pre-screening of citations using an RCT classifier and study selection within Covidence. These methods use an iterative approach to determine the potential impact of new data on the existing evidence base, and analyses are updated where required.

Results: We established a platform in which new and relevant evidence is identified and incorporated frequently as and when it becomes available. This process is facilitated through the application of an RCT classifier, use of crowd sourcing to screen potential studies and use of an expert panel for assessing the likely relevance and impact of new evidence. Results from both living systematic reviews form the basis for the living Australian Evidence-Based Clinical Guidelines for Diabetes

Conclusions: This project is an exemplar of the methods used to maintain currency of systematic reviews through the frequent searching, analysis and incorporation of new and relevant evidence. The utility of these reviews as a foundation for clinical guidelines ensures that patients receive the best possible care based on available evidence.

Patient or healthcare consumer involvement: Patients and consumers are involved in determining the clinical questions, interpreting the evidence and assist in generating clinical guideline recommendations using the evidence.

Thorough debridement reduces infection rates in open fracture care not time to index surgery: addressing the Global Commission of Surgery's call to improve care for the Bellwether Procedures – a systematic review and meta-analysis of observational studies

Foote C[1]

[1] McMaster University, Canada

Background: Open fractures are one of the leading causes of disability worldwide. The treatment of these injuries has been commissioned by the Lancet Commission for Global Surgery as one of the central three surgical priorities to reduce mortality and morbidity globally.

Objective: To identify the optimal timing of the treatment of open fractures.

Methods: In February 2020, we searched PUBMED, Cochrane Wounds Specialised Register, the Cochrane Central Register of Controlled Trials (CENTRAL), Ovid MEDLINE (including In-Process & Other Non-Indexed Citations), Google Scholar, Orthoevidence.com, Ovid Embase, and EBSCO CINAHL Plus. We conducted manual searches of retrieved articles and previous systematic reviews along with a gray literature search. We included observational studies, both prospective and retrospective, as well as randomized trials that evaluated the association between the timing of irrigation and debridement and the development of surgical site infection in open fractures. We then conducted an extensive meta-analysis of observational studies using raw and adjusted outcomes to determine if there was any plausible association between the timing of surgery and infection.

Results: The search resulted in 1148 studies. After screening the titles, we conducted a thorough review of 316 studies, which resulted in 35 observational studies, including 8459 fractures. In unadjusted analyses, there was no association between irrigation and debridement and surgical site infection. If anything, the estimate showed a protective effect of later debridement (OR 0.88, 95% CI 0.72 to 1.07, I^2 = 24%, P = 0.20, 35 studies). This effect was consistent across most subgroups. Adjustment for confounding was only performed in six studies. Aggregating adjusted estimates using the inverse-variance technique with time to debridement as a continuous variable, the risk did not increase significantly for each hour of delay to debridement (OR 1.02, 95% 1.00 to 1.04, I^2 = 32%, P = 0.08, 6 studies). Adjusted estimates using a 6 to 8 hour cut-off and mostly very low-quality evidence, showed no significant increase in the odds of infection with delay past 6 to 8 hours (random effects, OR 1.08, 95% CI 1.00 to 1.18, 6 to 8 hour cut-off, I^2 = 72%, P = 0.07, 6 studies and fixed effects, OR 1.02, 95% CI 1.01 to 1.03, I^2 = 45% P = 0.005, 6 studies).

Conclusions: This complete review of the evidence consisting of 35 observational studies did not find an association between irrigation and debridement and surgical site infection in open fractures. This was consistent across all subgroups and when deep infection was considered. There is currently no available time point where irrigation and debridement is associated with increased rates of infection. Future studies must improve methodological quality to validate this finding.

Patient or healthcare consumer involvement: None

Volatile anesthetics versus total intravenous anesthesia in patients undergoing coronary artery bypass grafting: An updated meta-analysis and trial sequential analysis of randomized controlled trials

Jiao X[1], Ni X[1], Li H[1], Zhang L[1]

[1] West China Second University Hospital, Sichuan University; Cochrane China Center, China

Background: The benefits of volatile anesthetics in coronary artery bypass grafting (CABG) patients remain controversial.

Objectives: We aimed to conduct an updated meta-analysis to assess whether the use of volatile anesthetics during CABG could reduce mortality and other outcomes.

Methods: We searched eight databases from inception to June 2019 and included randomized controlled trials (RCTs) comparing the effects of volatile anesthetics versus total intravenous anesthesia (TIVA) in CABG patients. The primary outcomes were operative mortality and one-year mortality. The secondary outcomes included the length of stay in the intensive care unit (ICU) and hospital and postoperative safety outcomes (myocardial infarction, heart failure, arrhythmia, stroke, delirium, postoperative cognitive impairment, acute kidney injury, and the use of intra-aortic balloon pump (IABP) or other mechanical circulatory support). Trial sequential analysis (TSA) was performed to control for random errors.

Results: We included 89 RCTs comprising 14,387 patients. There were no significant differences between the volatile anesthetics and TIVA groups in operative mortality (relative risk (RR) 0.92, 95% confidence interval (CI): 0.68 to 1.24, P = 0.59, I^2 = 0%), one-year mortality (RR 0.64, 95% CI 0.32 to 1.26, P = 0.19, I^2 = 51%), or any of the postoperative safety outcomes. The lengths of stay in the ICU and hospital were shorter in the volatile anesthetics group than in the TIVA group. TSA revealed that the results for operative mortality, one-year mortality, length of stay in the ICU, heart failure, stroke, and the use of IABP were inconclusive.

Conclusions: Conventional meta-analysis suggests that the use of volatile anesthetics during CABG is not associated with reduced risk of mortality or other postoperative safety outcomes when compared with TIVA. TSA shows that the current evidence is insufficient and inconclusive. Thus, future large RCTs are required to clarify this issue.

MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE

A fully automated pipeline for a living review of methods to (semi)automate data extraction

Schmidt L[1], Olorisade BK[1], McGuinness LA[1], Higgins JPT[1]

[1] University of Bristol, UK

Background: Data extraction is one of the most time-consuming and complex tasks for the authors of systematic reviews (SR). It is an area that holds promise for the application of machine learning technology and text mining. The fields of information science and data science are constantly evolving, and there is a steady flow of new research in data mining. This situation supports the choice of conducting a living review in this topic area. However, living reviews are resource-intensive, and require the application of technological support for efficiency and maximization of their life cycles.

Objectives: (1) To conduct a living review of methods and tools for extracting specific items of information/data from reports of health research studies in order to (semi)automate parts of the systematic review process. (2) To develop fully automated, technology backed workflows to assist with this living review throughout its life cycle.

Methods: Publications for this living review are regularly retrieved from MEDLINE, Web of Science, IEEE, dblp and the computer science arXiv using database APIs, Python and R libraries to scrape and search data. Two review authors screen titles and abstracts every two months with the aid of machine learning algorithms. We screen eligible full texts and extract data related to design and quality of reporting for a cross-sectional analysis of the available evidence. Full review updates are planned in 6-month intervals if the amount of new evidence permits it. For machine learning we employ an ensemble of classical (SVM, LDA) and deep neural methods (BERT, XLM).

Results: The initial information retrieval is automated by the first two modules in our pipeline, using APIs and scraping of databases in order to automate systematic searches on grey literature and information science databases that do not offer advanced search techniques in their interfaces. The third module in our pipeline applies an ensemble machine leaning classifier based on our own, as well as on other previously published machine learning architectures.

Conclusions: We present a fully automated information retrieval pipeline with an integrated, active-learning abstract screening system to support a living review throughout its life cycle (Figure 1). By re-using and integrating previously published classifiers into one ensemble we strive to reduce duplication of efforts. The pipeline is modular, and parts related to the search strategy, searched databases, and training for the machine learning can be replaced when conducting a different living review.

Patient or healthcare consumer involvement: No patients were involved in this research. We involved fellow systematic reviewers as stakeholders and aimed to integrate already existing machine-learning infrastructures into this project in order to reduce duplication of efforts.

Additional files: pipeline

A new machine-learning powered tool to aid citation screening for evidence synthesis: PICOPortal

Agai E[1]

[1] The Fast Healthcare Interoperability Resources, Resources for Evidence-Based Medicine, Knowledge Assets project (EBMonFHIR), USA

Background: Identifying all evidence relevant to a systematic review remains a critical yet time-consuming step in the evidence synthesis process. Machine learning methods and collaborative screening software constitute potential means of reducing the workload necessary to perform citation screening, without sacrificing rigor. But while sophisticated web-based tools for performing citation screening have emerged (e.g. Covidence, Rayyan) they tend to feature relatively limited automation functionality. Conversely, open-source research prototypes (e.g. abstrackr, RobotReviewer) offer more sophisticated machine learning to aid synthesis but lack simple user experiences.

Objectives: We introduce PICOPortal, a new web-based tool for citation screening that facilitates systematic reviews that aspires to combine the strengths of a modern user interface and cutting-edge machine learning functionality. Users can create publicly available profiles within PICOPortal that include their areas of expertise and levels of experience, potentially facilitating collaboration across systematic review groups based in other labs or research organizations around the world. The tool is free for academic users.

Methods: PICOPortal provides project management and basic reference management functionality in a modern user interface along with state-of-the-art machine learning capabilities to facilitate an efficient review. It integrates a validated, state-of-the-art machine learning to optionally including only randomized controlled trials (RCT), it detects de-duplication of citations and automatically extracts snippets of text from titles and abstracts pertaining to the descriptions of trial Populations, Interventions/Comparators, and Outcomes (PICO elements), respectively (Figure 1). Additionally, extraction of these snippets facilitates automated topic-scope based exclusion of articles on the explicit basis of one or more elements such as Population that is inappropriate for the scope of the review at hand. This is in contrast to models that make an overall relevance determination without explicit reference to an underlying PICO criterion.

Conclusions: PICOPortal is a new web-based tool for collaborative citation screening for systematic reviews. It features cutting-edge machine learning models that are integrated into an intuitive interface, thus combining the respective strengths of existing commercial and academic citation screening tools. PICOPortal is designed to support a team's systematic review process through its entire life cycle; while maintaining an emphasis on academic rigor, workflow optimization and flexibility, and global collaboration. PICOPortal is free for academic users.

Patient or healthcare consumer involvement: Systematic reviews provide the best means of realizing the practice of evidence-based medicine. Citation screening, which the described tool facilitates, is a key component of such reviews. Patients, therefore, stand to benefit indirectly from researcher use of the PICOPortal tool described in this abstract.

Additional file: Figure

Automated extraction of adverse drug reactions from biomedical literature and FDA drug labels using machine learning

Nandal U[1]

[1] Elsevier, The Netherlands

Background: Elsevier is a global information analytics company leveraging its' rich tradition of curating and publishing leading scientific content to power clinical and research solutions. By combining content with cutting-edge technology in artificial intelligence, machine learning (ML) and natural language processing (NLP), we enable professionals to find the precise information they need to advance their research and make decisions that affect the lives of patients and whole societies. In pharmacovigilance, post-marketing drug safety surveillance is critical to the protection of public health and monitoring the diverse sources of information for cases of Adverse Drug Reactions (ADRs) is a critically important and time-consuming task. Automatic extraction of ADRs from both highly regularized and variably structured content could play an important role in augmenting the information about ADRs that is obtained during short-term clinical trials.

Objectives: To automatically extract ADRs from FDA structured product labels (SPLs) and scientific journal articles. Rule and dictionary-based approaches to this problem may yield excessive false positives due to the lack of context consideration, as adverse drug reaction terms may be indistinguishable from symptoms of diseases. Therefore, we aim to model the language surrounding adverse drug reaction mentions to provide more precise predictions.

Methods: We randomly selected FDA SPLs by Anatomical Therapeutic Chemical class and from Embase, journal articles containing mentions of drugs and ADRs. We then manually annotated mentions of drugs and ADRs in triplicate and harmonized annotations to create two separate gold standard data sets; one for SPL content and another for article content. We next used these manual annotations to train a number of ML models including CRF, BiLSTM, and spaCy for the prediction of drug and ADRs mentions.

Results: Our current models, trained on 6234 natural language sentences (5909 unique) containing 9796 ADR annotations (2687 unique) from SPL content yield mean 5-fold cross-validation precision (P), recall (R), and F1-scores (F) of 0.80, 0.78, and 0.79 respectively. For comparison, a dictionary-based method yields P, R, and F of 0.57, 0.70, and 0.63 respectively. Inter-annotator agreement ranges from 0.70 to 0.77 (Cohen's kappa), suggesting that model performance is comparable to human performance in this domain. Separate model development for table content (eg tables extracted from SPLs, as opposed to natural language) is ongoing. Gold set manual annotation for the journal article content is also still in progress.

Conclusions: Automatic extraction of ARDs from both highly structured SPLs and less structured journal articles is feasible and represents a viable methodology for fact extraction.

Patient or healthcare consumer involvement: This work has a number of applications to both patient safety and pharmacovigilance research. Automatic extraction of ADRs can save considerable time and effort in large-scale analyses and in integrating data from multiple diverse content sources.

Can artificial intelligence learn to identify systematic reviews on the effectiveness of public health interventions?

Read K[1], Husson H[1], Dobbins M[1]

[1] National Collaborating Centre for Methods and Tools (NCCMT), Canada

Background: Health Evidence aims to make it easier for public health professionals and decision-makers to use evidence in their programs and policies. We provide access to over 6000 critically appraised systematic reviews on the effectiveness of public health interventions. As the number of reviews published each year continues to grow, maintaining this repository is becoming more resource-intensive. On average, 8000 to 10,000 records are screened each month to identify around 50 relevant reviews that are critically appraised and uploaded to the registry. Artificial intelligence (AI) may be one way to ensure the maintenance of this registry continues to be feasible.

Objectives: To explore whether AI can be used to accurately and efficiently conduct monthly relevance screening for the Health Evidence registry.

Methods: Using the Distiller SR platform, we uploaded a large reference set (n = 4584 relevant reviews and 18,699 not relevant reviews) to train the Distiller Artificial Intelligence System (DAISY) with the Health Evidence relevance criteria. The team trained DAISY on 70% of the labelled training set and had the platform score the articles. We then established an exclusion threshold based on the lowest score to correctly identify not relevant reviews. To test this threshold, we used DAISY on the remaining 30% (n = 9,985) of uploaded articles and tested two additional sets from two monthly updates (month a = 7917, month b = 7848). We calculated the percentage of reviews DAISY automatically excluded and compared the predicted results to our manual screening of the monthly update to identify classification errors.

Results: Using 70% of the training set, the team identified an exclusion threshold of 0.47 (as 0.49 was the lowest score to correctly identify not relevant reviews). Applying DAISY to the additional test sets automatically excluded 24% of the records. When comparing to the manual screening results, the false exclusion rate with these sets was 0.02%. On average, the Health Evidence team will manually screen approximately 500 records an hour. Using this estimate, the AI functionality in Distiller could save up to four hours of manual screening per month. The next steps will be to test additional monthly update sets where manual screening results are available and explore both inclusion and exclusion threshold options.

Conclusions: The use of AI technology shows promise for helping to automate the Health Evidence monthly update process and improve the feasibility of maintaining a registry of public health relevant systematic reviews.

Decoding semi-automated title-abstract screening: an exploration of the review, study, and publication characteristics associated with Abstrackr's relevance predictions

Gates A[1], Gates M[1], Elliott S[1], Pillay J[1], DaRosa D[1], Rahman S[1], Vandermeer B[1], Hartling L[1]

[1] Alberta Research Centre for Health Evidence, University of Alberta, Canada

Background: Machine learning (ML) tools can reduce screening workloads in systematic reviews but their adoption has been slow. To build trust, review teams may benefit from a better understanding of how and when ML-assisted screening may be most safely and effectively applied.

Objectives: We evaluated the risks (missed records) and benefits (time saving) of using Abstrackr to semi-automate title-abstract screening, and explored whether Abstrackr's predictions varied by review or study-level characteristics.

Methods: For each of 16 reviews we uploaded the records to Abstrackr, screened a 200-record training set, and downloaded the predicted relevance of the remaining records. We then retrospectively simulated the liberal-accelerated screening approach, whereby the senior review author screened the records predicted as relevant and the second review author then screened those predicted as irrelevant and those excluded by the senior review author. We estimated the time savings (assuming 30 seconds per record) and calculated the proportion missed (records included in the final reports that were wrongly excluded) compared with dual independent screening. We compared the review and study-level characteristics of Abstrackr's 'correct' and 'incorrect' predictions using Fisher's Exact and unpaired t-tests.

Results: The median (interquartile range (IQR)) screening workload was 2123 (4641) records. Across systematic reviews our approach wrongly excluded 0 to 3 (0% to 14%) records in the final reports and saved a median (IQR) 26 (33) hours of screening time. Of 802 records in the final reports, 87% were correctly predicted as relevant. The correctness of the predictions did not differ by review type (systematic (88% correct) or rapid (84%), P = 0.37) or intervention type (simple (88%) or complex (86%), P = 0.47). The predictions were more often correct in reviews with multiple (89%) vs. single (83%) research questions (P = 0.01), and that included only trials (95%) vs. multiple study designs (86%) (P = 0.003). At the study level, trials (91%), mixed methods (100%), and qualitative (93%) studies were more often correctly predicted as relevant compared with observational studies (79%) or reviews (83%) (P = 0.0006). Studies at high or unclear (88%) vs. low risk of bias (80%) (P = 0.039), and those published more recently (mean (standard deviation (SD)) 2008 (7) vs. 2006 (10), P = 0.02) were more often correctly predicted as relevant. There was no difference in the mean (SD) journal impact factor for correctly included (4.91 (8.39)) and wrongly excluded (4.61 (9.14)) records (P = 0.74).

Conclusions: Our ML-assisted screening approach saved considerable time and may be suitable in conditions where the limited risk of missing relevant records is tolerable (e.g. rapid or scoping reviews). ML-assisted screening may be most trustworthy for reviews that seek to include only trials or more recent publications; however, as several of our findings are paradoxical further study is needed to understand the tasks to which ML-assisted screening is best suited.

Patient or healthcare consumer involvement: None.

Diagnostic accuracy of machine-learning-assisted detection for anterior cruciate ligament and meniscal injuries based on magnetic resonance imaging: protocol for a systematic review and meta-analysis

Lao Y[1], He G[1], Gao X[1], Guan Y[1], Li J[1], Yang K[2]

[1] Second Clinical Medical College, Lanzhou University, Lanzhou, China; [2] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China

Background: Magnetic resonance imaging (MRI) of the knee is the preferred method for diagnosing knee injuries. Anterior cruciate ligament (ACL) and meniscal injuries are two common knee sports injuries. Concomitant meniscal lesions are common in patients with anterior cruciate ligament (ACL) injuries and frequently involve the posterior horn. Although many machine learning algorithms have been developed to detect ACL and meniscal injuries based on MRI, the performance of different algorithms required further investigation.

Objectives: To evaluate the diagnostic accuracy of machine-learning-assisted detection for ACL and meniscal injuries based on MRI and find the current best algorithm.

Methods: We will conduct a comprehensive database search for clinical diagnostic tests in PubMed, Embase, the Cochrane Library, and Web of Science without restrictions on publication status and language. We will also check the reference lists of the included articles to identify additional studies for potential inclusion. Two review authors will independently review all literature for inclusion and assess their methodological quality using Quality Assessment of Diagnostic Accuracy Studies version 2 (QUADAS-2). We will consider for inclusion clinical diagnostic tests exploring the efficacy of machine-learning-assisted system for detecting ACL and meniscal injuries based on MRI. Another two review authors will independently extract data from eligible studies based on a pre-designed standardized form. Any disagreements will be resolved by consensus. We will use RevMan 5.3 and Stata SE 12.0 software for data synthesis. If appropriate, we will calculate the summary sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, and diagnostic odds ratio of machine-learning-assisted diagnosis system for ACL and meniscal injuries detection. We will plot a hierarchical summary receiver operating characteristic (HSROC) curve, and the area under the ROC curve (AUC) will be calculated using the bivariate model. If applicable, we will conduct subgroup analysis based on pre-set criteria to find more information on:

1) different type and degree of ACL and meniscal injuries;
2) different machine learning algorithms used in primary studies;
3) different MRI sequences and magnet intensities used in primary studies. If the pooling of results is considered inappropriate, we will present and describe our findings in diagrams and tables and describe them narratively.

Patient or healthcare consumer involvement: There was no patient or healthcare consumer in this project.

How many reviews are using automation? A cross-sectional analysis

Luo X[1], Lv M[2], Liu Y[2], Ren M[2], Zhang X[3], Wang L[2], Liu X[2], Yang K[4], Chen Y[5], Wang M[6]

[1] School of Public Health of Lanzhou University, Lanzhou; Cochrane China Network for GRADE and Guideline Working Group, Lanzhou, China; [2] School of Public Health of Lanzhou University, Lanzhou, China; [3] The First School of Clinical Medicine, Lanzhou University, Lanzhou, China; [4] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China; [5] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou; Cochrane China Network for GRADE and Guideline Working Group, Lanzhou, China; [6] Department of Radiology, the First Hospital of Lanzhou University, Lanzhou, China

Background: Automation technology has been proposed or used to accelerate most steps of the systematic review process, including search, screening, data extraction, and quality evaluation. However, the number of reviews using automation is unclear.

Objectives: To review the current situation and quality of reviews using automation.

Methods: We searched MEDLINE via PubMed, Embase, the Cochrane Library, Web of Science from inception to 1 March 2020 for evidence synthesis studies that used automation methods. We used the following search strategy: [Automation terms (such as RobotReviewer, machine learning, Artificial Intelligence, etc.)] AND [evidence synthesis terms (such as systematic review, meta-analysis, literature review, etc.)]. The search had no language restrictions and was limited to human subjects. We used AMSTAR (A MeaSurement Tool to Assess systematic Reviews) to assess the quality of included studies. Finally, we conducted a descriptive analysis of the characteristic of the included reviews.

Results: We included 61 reviews. More than half of the reviews (36; 59.0%) were published between 2019 and 2020. The topics of the included reviews varied. The most common type was systematic reviews (57; 93%), including 15 Cochrane Reviews (24%). Most reviews used Covidence (41; 77%), and others used Rayyan (8; 13%), EPPI-Reviewer 4 (5; 8%) and RobotReviewer (3; 5%). The application of automation in evidence synthesis mainly included duplication removal (52; 86%), study selection (49; 81%), data-extraction (21; 35%) and quality assessment (11; 18%). The quality of those reviews was moderate to high. More than half of the reviews (40; 66%) scored more than 8 (total score 11), and the rest scored between six and eight.

Conclusions: In recent years, the number of reviews using automation technology has been increasing, and the quality is relatively high. At present, automation is mainly used in the retrieval, screening, data extraction and evaluation stages. However, the reliability and validity of automation tools should be specified.

Patient or healthcare consumer involvement: None.

Implementation of artificial intelligence in healthcare: a scoping review

Tricco AC[1], Darvesh N[2], Thomas SM[2], Fennelly O[3], Brar R[2], Straus SE[4]

[1] Li Ka Shing Knowledge Institute, St. Michael's Hospital, Unity Health Toronto; Epidemiology Division, Dalla Lana School of Public Health, University of Toronto, Canada; [2] Li Ka Shing Knowledge Institute, St. Michael's Hospital, Unity Health Toronto, Canada; [3] Irish Centre for High End Computing (ICHEC), National University of Ireland (NUI) Galway, Ireland; [4] Li Ka Shing Knowledge Institute, St. Michael's Hospital, Unity Health Toronto; Department of Medicine, University of Toronto, Canada

Background: Artificial intelligence (AI) has the potential to lead to significant improvements in health care and public health. This includes using data analytics to personalize health care for patients, assist healthcare professionals, and tailor organizational and policy decision-making. However, to optimize the implementation of AI across health care in a safe, effective and sustainable manner, the current implementation strategies and outcomes at both the patient and population-level need to be examined.

Objectives: To conduct a scoping review to identify what strategies are used to implement AI interventions for health or within healthcare systems.

Methods: The Joanna Briggs Institute reviewer's manual will guide the conduct of this review and the review protocol has been published. The eligibility criteria for this review includes:

- Population: adults and children of any age.
- Intervention: implementation of AI tools for health or within a health system.
- Comparator: any.
- Outcome: any outcome at the patient, public, clinician, population or system level.
- Study design: all primary experimental, cohort and case-control studies.
- Year published: limited to 2008 onwards.

An experienced information specialist will develop a search strategy, which will be peer reviewed. We will search multiple databases as well as grey literature and reference lists of included studies. We will screen identified articles by titles and abstract and then by full text in pairs of review authors, with discrepancies resolved by a third review author. We will use a standardized charting form to extract data from the included studies by pairs of review authors independently and we will contact study authors where information is missing or unclear.

Results: Results will include the types of AI tools implemented for health or within a health system, the implementation strategies, the participants, and outcomes which may include sustainability, scalability, barriers and facilitators.

Conclusions: The findings from this review will identify areas where AI has been implemented in health care, how it has been implemented, and the outcomes at patient, public, clinician, population, and system levels. This will inform future research on AI implementation. In the long term, it will also help inform strategies for the safe, effective, and sustainable implementation of AI tools in health care in order to improve healthcare quality.

Patient or healthcare consumer involvement: Not applicable.

Is it time to trust the robots? The reliability and usability of machine learning tools for screening in systematic reviews

Gates A[1], Pillay J[1], Guitard S[1], Elliott S[1], Dyson M[1], Newton A[2]

[1] Alberta Research Centre for Health Evidence, University of Alberta, Canada; [2] Department of Pediatrics, University of Alberta, Canada

Background: Machine learning tools can expedite the completion of systematic reviews (SRs) by reducing manual screening workloads, yet their application has been minimal. Evidence of their benefits and enhanced usability may improve their acceptance within the SR community.

Objectives: We tested the performance of three tools when used to: (a) eliminate irrelevant records (Simulation A) and (b) replace one of two independent reviewers (Simulation B). We evaluated the usability of each tool.

Methods: We selected three SRs completed at our centre and subjected these to two retrospective screening simulations. Using each tool (Abstrackr, DistillerSR, and RobotAnalyst), we screened a 200-record training set and downloaded the predicted relevance of the remaining records. To test their performance, we calculated the proportion missed, workload savings, and estimated time savings compared to dual independent screening by two reviewers. To test usability, screeners undertook a screening exercise in each tool and completed a user experience survey, incorporating the System Usability Scale (SUS).

Results: Using Abstrackr, DistillerSR, and RobotAnalyst respectively, the median (range) proportion of records missed was 5 (0 to 28)%, 97 (96 to 100)%, and 70 (23 to 100)% in Simulation A and 1 (0 to 2)%, 2 (0 to 7)%, and 2 (0 to 4)% in Simulation B. The median (range) workload savings was 90 (82 to 93)%, 99 (98 to 99)%, and 85 (85 to 88)% for Simulation A and 40 (32 to 43)%, 49 (48 to 49)%, and 35 (34 to 38)% for Simulation B. The median (range) time savings was 154 (91 to 183), 185 (95 to 201), and 157 (86 to 172) hours for Simulation A and 61 (42 to 82), 92 (46 to 100), and 64 (37 to 71) hours for Simulation B. Based on the median (IQR) SUS scores (/100), Abstrackr fell in the usable (79 (23)), DistillerSR the marginal (64 (31)), and RobotAnalyst the unacceptable (31 (8)) usability range (n = 8). Participants indicated that usability was contingent on six interdependent properties: user friendliness, qualities of the user interface, features and functions, trustworthiness, ease and speed of obtaining predictions, and practicality of the export file(s).

Conclusions: Our findings support the cautious use of machine learning tools to replace the second reviewer (Simulation B); the workload savings were substantial and few, if any, records were erroneously excluded. Designing tools based on reviewers' self-identified preferences may improve their usability.

Patient or healthcare consumer involvement: None.

Machine learning methods for motor recovery prediction and prognosis in post-stroke rehabilitation: a systematic review

Mannini A[1], Campagnini S[1], Arienti C[2], Patrini M[2], Negrini S[3], Carrozza MC[4]

[1] Istituto di Biorobotica, Scuola Superiore Sant'Anna, Pisa; MARE Lab, IRCCS Fondazione Don Carlo Gnocchi, Florence, Italy; [2] IRCCS Fondazione Don Carlo Gnocchi, Milan, Italy; [3] Department of Biomedical, Surgical and Dental Sciences, University of Milan "La Statale", Milan; IRCCS Istituto Ortopedico Galeazzi, Milan, Italy; [4] IRCCS Fondazione Don Carlo Gnocchi, Milan; Istituto di Biorobotica, Scuola Superiore Sant'Anna, Pisa, Italy

Background: The rehabilitation field has always been characterized by the difficulty of conducting rigorous clinical trials and the need for individualized care for the patient. The recent framework of Rehabilomics addresses the gap between research and clinical treatment needs. It promotes a systematic collection of data from the patient and it uses it to generate a treatment protocol for personalized therapy. Machine learning techniques can be considered a primary tool for embracing this new framework.

Objectives: To develop a systematic review on machine learning algorithms trained and validated as predictive models for the clinical outcome of post-stroke patients after rehabilitation treatment.

Methods: We conducted a systematic review and included machine learning methods as predictive performance of motor recovery in all types of stroke. We conducted a comprehensive search of electronic databases such as PubMed, Web of Science, Scopus, CINAHL and the Cochrane Central Register of Controlled Trials (CENTRAL) using a Patient, Intervention, Comparison, Outcome (PICO) format, from inception to 7 February 2020. Data extracted included: health condition, intervention in the experimental and control group, dose, frequency and number of sessions, outcome assessed and how it has been measured, method for features extraction and selection, algorithm used for the model and validation approach. To assess the methodological quality of included reviews we used the prediction model risk-of-bias assessment tool (PROBAST), which assesses risk of bias over four domains, as well as applicability. We have provided a narrative description of the characteristics of the primary studies and performed a narrative data synthesis reporting the performance of individual prognostic models. We evaluated the opportunity of performing a meta-analysis on the level of heterogeneity of primary studies included.

Results: A total of 846 studies met the inclusion criteria and were included in systematic review. All participants were adults with stroke. The data analysis is ongoing and we will present the final results during the Cochrane Colloquium.

Conclusions: Our results will highlight the better performing models and next steps for their comparison, extension or implementation.

Patient or healthcare consumer involvement: Not applicable.

New approach to automated citation screening to improve reliability of review updates

Nowak A[1], Pieniak M[1], Borowiack E[1], Bąk A[1], Kunstman P[1], Brozek J[2]

[1] Evidence Prime Inc, Poland; [2] McMaster University, Canada

Background: A great deal is expected of semi-automation of screening of titles and abstracts, given the ever-increasing rate of scientific output. However, the reported experiences of users of the existing machine learning systems have been mixed, especially in terms of achieved sensitivity. Arguably, an important source of this variability is that the percentage of citations that need to be screened manually in order to achieve high sensitivity (the size of the training set) is different for every review. This parameter cannot be estimated a priori, only after exploration of a subset of the data, for example in the active learning scenario. It is unclear, however, how to improve the reliability of updates, especially in the context of living systematic reviews, which – due to their frequency – require a more hands-off approach.

Objectives: To identify which review-specific factors can lead to poor performance of the machine learning models used for screening and design improved models that would be free of these limitations.

Methods: We constructed a representative sample of 36 systematic reviews and performed a retrospective, simulated screening update using a re-implementation of the current state-of-the-art models, trained on 50% of the dataset and tested on the remaining half. For 10 of the reviews, we conducted error analysis to determine sources of poor performance.

Results: The performance of the baseline models varied greatly ranging from 33% to 100% in terms of sensitivity and 3% to 19% in precision. We attribute this variability to the fact that the training set size was insufficient for some of the reviews. The error analysis further showed that low recall was caused by a small number of included studies in the original review, complex inclusion criteria (e.g. indirect evidence) and topic drift (for instance, appearance of a new intervention in the updated review). Based on these findings, we constructed a formalized framework for definition of inclusion and exclusion criteria and designed a new machine learning model that can use this information. This model, inspired by the recent advancements in few-shot learning, differs from the previous approaches, as it is not trained individually for every review, but is pre-trained on a large set of existing reviews to meta-learn the specifics of the screening task. This direction proved to be very promising and while the work is still ongoing, we hope that our results will encourage other groups to pursue this direction.

Conclusions: Typical problems encountered while applying machine learning to screening may be caused by insufficient information supplied to the model. For reviews with very low inclusion rates, it may never be feasible to train the models on the included and excluded citations alone. Therefore, we propose to redefine the problem of screening automation to include other data, such as inclusion and exclusion criteria or the citation graph.

Patient or healthcare consumer involvement: None

Semi-automated data extraction workbench for environmental health

Howard B[1], Maharana A[1], Tandon A[1], Albert T[1], Phillips J[1], Taylor M[2], Thayer K[2], Shah R[1]

[1] Sciome LLC, USA; [2] Environmental Protection Agency, USA

Background: Systematic review, already a cornerstone of evidence-based medicine, has begun to gain significant popularity in several related disciplines including environmental health and evidence-based toxicology. A critical, time-consuming process that occurs during systematic review is the extraction of relevant qualitative and quantitative raw data from the text of scientific documents. The specific data extracted differs among disciplines, but within a given domain, certain data points are extracted repeatedly for each review that is conducted.

Methods: We have recently developed a semi-automated data extraction workbench for use in this context. Our research has focused on three specific goals. First, we are using deep learning to build novel data extraction components for items of interest within the domain of environmental health. Second, we have created web-based software specifically designed for extraction in the context of systematic review. Finally, we have introduced new protocols to standardize the inputs and outputs for data extraction software components.

Results: A beta version, currently under evaluation at EPA, includes more than 30 novel data extraction components relevant to environmental toxicology. Performance varies widely among data types with some tasks inherently more difficult than others. For certain simple data items, like sex of the experimental animal, we achieve F-scores in excess of 95%; for more difficult entities, we were still often able to achieve an F-score of 65% or more, given sufficient training data. Importantly, the design of our workbench makes it easy to include extraction components developed by other research groups. The workbench currently includes several such components, with new ones added regularly.

Conclusions: Because accurate data extraction is a challenging problem, and given that current methods rarely achieve 100% accuracy, we are integrating our methods into a "human-in-the-loop" system that combines machine and human intelligence in a manner that is superior to using either in isolation. The system will:

- highlight extracted terms in a pdf;
- automatically populate forms with extracted data;
- allow humans to intervene and correct the results; and
- learn from the corrections to continually update the model.

The resulting system will make systematic reviews both more efficient to produce and less expensive to maintain, greatly accelerating the process by which scientific consensus is obtained in a variety of health-related disciplines.

Patient or healthcare consumer involvement: The resulting system will make systematic reviews both more efficient to produce and less expensive to maintain, greatly accelerating the process by which scientific consensus is obtained in a variety of health-related disciplines.

The reliability and relative advantages of semi-automated approaches to title and abstract screening: making the case for machine learning

Gates A[1], Gates M[1], Sebastianski M[2], Guitard S[1], Elliott S[1], Hartling L[1]

[1] Alberta Research Centre for Health Evidence, University of Alberta, Canada; [2] Alberta Strategy for Patient-Oriented Research (SPOR) SUPPORT Unit Knowledge Translation Platform, University of Alberta, Canada

Background: Machine learning can expedite evidence synthesis by semi-automating title and abstract screening. There is a need for evidence of the relative advantages and reliability of semi-automated screening approaches to inform guidance on their integration into modern review processes.

Objectives: Compared to screening by a single experienced review authors in rapid reviews (RRs) and dual independent screening in systematic reviews (SRs), we investigated the reliability and relative advantages of using a machine learning tool to (a) automatically exclude irrelevant records and (b) replace one of two independent review authors. We evaluated the impact of erroneously excluded records on the primary outcome.

Methods: We selected 11 SRs and 6 RRs completed at our centre and subjected these to two retrospective screening simulations in Abstrackr, a machine learning tool. For each SR and RR, we screened a 200-record training set and downloaded the predicted relevance of the remaining records. We calculated the proportion missed, workload savings, and estimated time savings compared to single (RRs and SRs) and dual-independent screening (SRs only) by human reviewers. We performed a citing articles search in Scopus or Google Scholar to determine if the missed studies would be identified via reference list scanning. For SRs with pairwise meta-analyses, we removed the missed studies and compared the pooled estimates of effects for the primary outcome to those in the final reports.

Results: When Abstrackr was used to exclude irrelevant records, the median (IQR) proportion missed was 20 (21)% (i.e. 9 (10) studies) for the SRs and 6 (12)% (i.e. 2 (10) studies) for the RRs. When used to replace one of two reviewers in the SRs, the median (IQR) proportion missed was 0 (1)% (i.e. 0 (2) studies). This diminished to 0 (1) studies following the citing articles search (0 studies in 7 SRs, 1 study in 2 SRs and 2 studies in 2 SRs). The missed studies had no impact on the results of the SRs. When used to exclude irrelevant records, the median (IQR) workload savings was 83 (12)% for the SRs and 34 (11)% for the RRs, for an estimated time savings of 44 (67) hours and 3 (3) hours, respectively. When used to replace one of two reviewers in the SRs, the median (IQR) workload savings was 33 (12)%, for an estimated time savings of 20 (30) hours.

Conclusions: Too many relevant studies were missed when records were automatically excluded to consider this approach. Few (≤ 3), if any, relevant studies were missed when Abstrackr was used to replace the second reviewer in a pair; however, this amounted to up to 14% of the included studies in small SRs. The proportion missed diminished to ≤ 10% (≤ 2 studies) after scanning reference lists. In the context of SRs with comprehensive search strategies, the cautious application of machine learning to replace one review author in a pair could save considerable screening time without impacting the results.

Patient or healthcare consumer involvement: None.

Evidence-based decision making for health policy in HIRA

Baek SJ[1], Seo JY[1], Lee SM[1], Bae JW[1], Seo KH[1], Kim HO[2]

[1] Health Insurance Review and Assessment Service, South Korea; [2] College of Medicine, Ewha Womans University, South Korea

Background: The Health Insurance Review and Assessment Service (HIRA) is the organization in South Korea that is responsible for the benefit claims review, quality assessment, and the setting and management of benefit standards of the National Health Insurance System. Gastric cancer is the most common malignancy in Korea, and fluorodeoxyglucose (FDG)-positron emission tomography (PET) is selective intensive review item in HIRA. There was a demand for evidence-based review of clinical effectiveness due to the controversy of the value of FDG-PET in evaluation of gastric cancer.

Objectives: To define appropriate use of FDG-PET and to determine a claims review guideline for FDG-PET in gastric cancer.

Methods: We organized consensus development for claims review (CDCR) committee. The CDCR consists of physicians of internal medicine, general surgery, pediatrics, nuclear medicine, including chair. The CDCR developed a questionnaire about the staging (9 items), RT planning (1 item), re-staging (4 items), measures of treatment efficacy (4 items). The CDCR reviewed evidence and medical standard, i.e. textbook, review article, guidelines, health policies in foreign countries. CDCR also reviewed the characteristics of gastric cancer in Korea. The consensus was achieved with a modified RAM method.

Results: From April to June 2019, a total of six meetings were held for review evidence and medical standard about key questions. For evidence-based decision making, we established a decision algorithm (Figure 1) Finally, 11 items were agreed. The consensus was developed for review guideline by the President of HIRA. The review guideline was announced officially in December 2019.

Figure 1
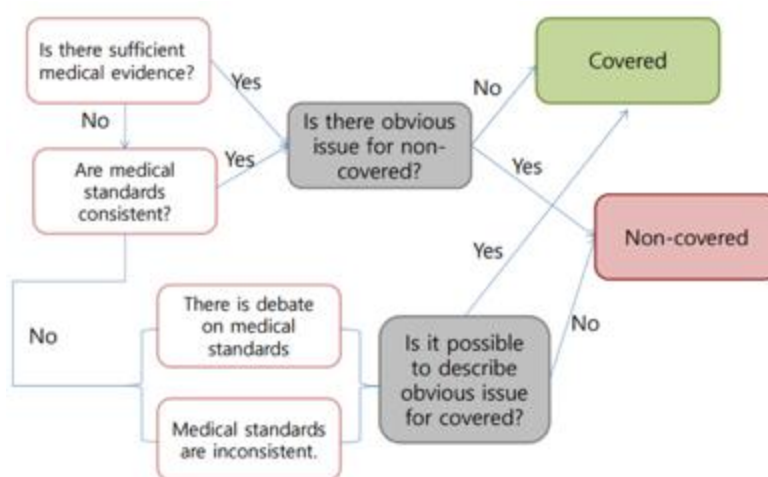
Conclusions: The review guideline of FDG-PET in gastric cancer is expected to improve review consistency. HIRA will continue to make efforts for the advancement of evidence-based decision-making systems. Evidence-based decision-making systems are a useful tool for making transparent health policy decisions, particularly when there is debate around the value of the treatment.

Patient or healthcare consumer involvement: Not applicable.

The use of nature language processing (NLP) for rapid literature screening in the update of systematic reviews: a comparison of four different models

Qin X[1], Liu J[1], Wang Y[1], Li L[1], Sun X[1]

[1] Chinese Evidence-Based Medicine Centre and Cochrane China Centre, West China Hospital, Sichuan University, Sichuan, China

Background: Updating systematic reviews represents an important mission of Cochrane. Literature screening accounts for a large proportion of efforts in the update of systematic reviews (SRs). The natural language processing (NLP) technology may have great potential for improving the efficiency of literature screening in the update of systematic reviews, particularly when the technology has learned from existing screened literature (i.e. the gold standard set).

Objectives: To compare the performance of different NLP models that are used for literature screening in the update of systematic review.

Methods: In our earlier systematic review of randomized controlled trials (RCTs) of SGLT2 inhibitors for treatment of type 2 diabetes (T2DM), we obtained 3460 de-duplicate reports by searching MEDLINE, Embase and the Cochrane Central Register of Controlled Trials (CENTRAL) from inception to June 2019. Two methods-trained review authors, using explicit eligibility criteria, manually screened titles and abstracts of these reports. We randomly divided these 3460 reports into training, development, test set at a ratio of 3:1:1. We firstly developed four supervised learning fitting models (i.e. NLP models) using the training and development set, including blueBERT-base uncased pretrained on PubMed (BlueBUP), blueBERT-base uncased pretrained pre-trained on PubMed abstracts and clinical notes (BlueBUPC), BERT-base cased (BBC), and BERT-base uncased (BBU). We then applied the test data set to evaluate the performance of four NLP models, including precision (i.e. the fraction of true positive samples among the retrieved positive samples), recall (i.e. sensitivity), f1 (i.e. harmonic mean of the precision and recall), and Areas Under the Receiver Operating Characteristics (AUROC).

Results: For the four NLP models – BlueBUPC, BlueBUP, BBC, and BBU – the precision scores were 0.767, 0.724, 0.724, and 0.728; the recall scores were 0.818, 0.869, 0.803, and 0.854; the f1 scores were 0.792, 0.815, 0.764, and 0.796; the accuracy scores were 0.915, 0.907, 0.902, and 0.919; and the AUROC were 0.96, 0.95, 0.96 and 0.96. The Receiver Operating Characteristics (ROC) curves were shown in Table 1 and Figures 1.

Conclusions: Our study showed that the NLP may be a useful tool to assist in literature screening when updating a systematic review, and the BlueBUP model may be a preferred method given the highest recall score and good f1 score, which are essential in literature screening. This approach is only usable for updating systematic reviews, and more validations studies are warranted.

Patient or healthcare consumer involvement: None.

Additional file: [Table and figure](#)

Using crowdsourcing and machine learning for study identification: a quantitative and qualitative evaluation of Cochrane's Screen4Me workflow

Noel-Storr A[1], Thomas J[2], Dooley G[3]

[1] Oxford University, UK; [2] University College London, UK; [3] Metaxis Ltd, UK

Background: In April 2019 Cochrane launched its Screen4Me workflow. The service, available to Cochrane author teams, has three components: 1) crowdsourcing, via Cochrane Crowd, 2) machine learning through the use of Cochrane's RCT classifiers, and 3) existing data – already known studies that have previously been classified by the Crowd. The aim of Screen4Me is to help reduce the number of studies to assess for potential inclusion.

Objectives: To provide quantitative data regarding the use of the new workflow, including number of times Screen4Me has been used, average mean reductions in size of search results set for each component of Screen4Me, as well as average time taken. We also plan to perform a qualitative analysis to better understand how Screen4Me is being used by two main user groups: 'the implementers', and 'the workers'.

Methods: For the quantitative analysis we downloaded the latest Screen4Me usage data from the Cochrane Register of Studies in March 2020. We collected data on: number of uses of Screen4Me for specific Cochrane Reviews, type of review (new or update), overall percentage reduction in number of search results, individual percentage reduction for each of the three components. For the qualitative analysis we will survey the Cochrane Information Specialists community and the Cochrane Crowd community. We will aim to understand how easy or not Cochrane Information Specialists have found using Screen4Me and what recommendations for improvement they have. For the Cochrane Crowd we will seek to ascertain satisfaction with participating in Screen4Me tasks in terms of task difficulty and task rewards.

Results: Screen4Me has been used 75 times by Cochrane Information Specialists across 15 Cochrane Review Groups. Sixty uses of it have involved the known assessment and RCT classifier components alone, while 15 have also used the Cochrane Crowd. Overall mean reduction in search results sets post Screen4Me was 54%. Average time taken where all three components were used was two weeks. We will present the results of the qualitative surveys at the Colloquium.

Conclusions: Screen4Me can reduce the number of search results for author teams to assess by a significant amount and is a robust and easy to use workflow available to Cochrane Review Groups.

Patient or healthcare consumer involvement: Anyone can join Cochrane Crowd and play a role in helping to identify studies for inclusion in Cochrane systematic reviews via Screen4Me.

HEALTH POLICY

A proposal of reporting items for evidence briefs for policy: RIGHT-EBP

Wang Q[1], Zhang JY[2], Wu SY[3], Ma YF[3], Yang N[3], Chen YL[3], Yang KH[3], Lavis J[1]

[1] Department of Health Research Methods, Evidence & Impact (HEI), Faculty of Health Sciences, McMaster University; McMaster Health Forum, McMaster University, Canada; [2] School of Public Health, Lanzhou University; Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Cochrane China Network, China; [3] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Cochrane China Network, China

Background: The evidence briefs for policy (EBPs) can provide potential policy options on a specific topic for health policymakers. An increasing number of organizations (e.g. McMaster Health Forum, the World Health Organization) already developed and are developing the policy briefs on different topics or fields. However, the reporting of policy briefs by different organizations and groups varies widely.

Objectives: To develop the reporting checklists for policy briefs for healthcare, to further help policymakers for better evidence-informed policymaking.

Methods: We will develop the reporting checklists for EBPs as an extension of the Reporting Tool for Practice Guidelines in Health Care (RIGHT) statement. ts design and implementation.

Results: We found 10,630 non-duplicate potentially eligible articles from our search. Of these, 11 have been included for data extraction and ten of which apply a difference in difference methodology. Initial findings from the 11 studies focus on two of the four quadruple aims, in particular on improving patient health outcomes and reducing per capita costs. In general, the findings indicate modest improvements in patient health and a reduction in per capita costs, most of which can be attributed to a reduction in discretionary services.

Conclusions: A key component of rapid-learning and improvement is ensuring that research evidence can be easily used by decision-makers to inform the design of a solution. The contextualization of findings from a systematic review can be used to support the rapid-learning cycle and to ensure that lessons learned can be applied to ongoing design of reforms elsewhere.

Contextualizing the findings from systematic reviews to inform health system transformations

Waddell K[1], Mansilla C[2], Hossain P[3], You J[4]

[1] McMaster Health Forum, McMaster University, Canada; [2] McMaster University, Chile; [3] McMaster University, Bangladesh; [4] McMaster University, South Korea

Background: While using evidence to inform decision making processes, policymakers often rely on evidence coming from other jurisdictions while planning high-level health system transformations. In these contexts, decision-makers often want to know what interventions are effective to achieve the quadruple aim. This presentation aims to show how contextualizing findings from a systematic review can address pressing issues to inform health system transformations in other contexts. Accountable care organizations (ACO) were implemented in the US in 2012, building on the successful elements of health maintenance organizations and managed care innovations. However, since their inception there has been a lot written in academic literature either touting their benefits or pointing out their flaws. Other jurisdictions have begun adopting this model with little synthesized information on its benefits and risks, including most recently significant health system reforms in Ontario, Canada

Objectives: To present how contextualizing the findings from systematic reviews can be used to inform health system transformation, showing an application on ACO.

Methods: We conducted a systematic review on accountable care organizations. We selected this topic because of the paucity of synthesized evidence addressing this reform and its relevance to ongoing reforms in Ontario and abroad. We searched three databases from 2010 to February 2020: MEDLINE, Embase and EconLit. To find gray literature, we supplemented the electronic database search with a handsearch of relevant websites. We extracted data related to the type of ACO, time period study, study design, relevance to the quadruple aim, and the key findings. We use Cochrane EPOC Risk of Bias tool and the Maryland Scientific Methods Scale to assess the quality of the included studies. We then map the findings from the systematic review on the components of the health system reform in Ontario to determine where lessons can be used to inform its design and implementation.

Results: We found 10,630 non-duplicate potentially eligible articles from our search. Of these, 11 have been included for data extraction and ten of which apply a difference in difference methodology. Initial findings from the 11 studies focus on two of the four quadruple aims, in particular on improving patient health outcomes and reducing per capita costs. In general, the findings indicate modest improvements in patient health and a reduction in per capita costs, most of which can be attributed to a reduction in discretionary services.

Conclusions: A key component of rapid-learning and improvement is ensuring that research evidence can be easily used by decision-makers to inform the design of a solution. The contextualization of findings from a systematic review can be used to support the rapid-learning cycle and to ensure that lessons learned can be applied to ongoing design of reforms elsewhere.

Creating a rapid learning primary healthcare system in Iran (developing a framework for implementation): study protocol

Ahmadi A[1], Doshmangir L[2], Majdzadeh R[1], Rashidian A[3]

[1] Knowledge Utilization Research Center, Tehran University of Medical Sciences, Iran; [2] Tabriz University of Medical Sciences, Iran; [3] Director of Information, Evidence and Research in the Eastern Mediterranean Region, WHO, Iran

Background: A rapid-learning (RL) health system makes optimal use of information technology and electronic health record capabilities. This system is a new concept in many countries especially in low- and middle-income countries. In such countries there is much more work to be done to build a RL health system, generate useful knowledge to improve health decisions and translate new evidence into better health for millions of people. Primary healthcare (PHC) is an approach to designing and delivering front line health services that lays a foundation for achieving universal health coverage (UHC). PHC in Iran is publicly funded and provided by a network of services more comprehensive in rural areas. On 5 May 2015, Iran's MoHME launched a major new reform, called Health Transformation Plan (HTP) in the Iranian health system. It included main interventions to: increase basic health insurance coverage in the total population, increase the quality of health care in public hospitals affiliated to MoHME, reduce out-of-pocket (OOP) payments, expand primary healthcare (PHC) networks, and revise medical tariffs services. However, the information process through the new plan needs to be enhanced. This study aimed to develop a framework to create a rapid learning primary healthcare system in Iran.

Objectives: (1) Provide a set of characteristics for a RL system in Iran's PHC; (2) Identify assets and gaps exist in Iran's PHC that can be leveraged or addressed, respectively, in creating a RL system; (3) Capitalize on or creating 'windows of opportunity' to stimulate the development and consolidation of a RL-PHC system in Iran; (4) Identify and adapt PHC plans or projects to RL system characteristics.

Methods: Step 1: Review of tools and mechanisms to establish and to support RL in health systems. Objective 1: Describe tools and mechanisms to improve RL system: engaged patients/citizen; digital capture, linkage and timely sharing of relevant data; timely production of research evidence; aligned governance, financial and delivery arrangements; appropriate decision supports; competencies for RL and improvement; culture of RL and improvement. Objective 2: Develop a tool/framework to improve RL in health systems -expert opinion/ qualitative study: consensus study; pilot/case studies

Step 2: Choosing one of the plans and projects of PHC system in Iran. Objective 1: Narrowing the focus of a subject. Objective2: Use of the results of this step for all PHC system.

Step 3: implementation of RL in Iran's PHC system Objective 1: identifying and defining components of RL health system approach in Iran's PHC system: assessment of Iran's health system characteristic; dentifying barriers and required interventions. Objective 2: defining well-suited intervention package to improve RL in Iran's PHC system. Objective 3: identifying and engaging key decision-makers to secure implement.

Patient or healthcare consumer involvement: PHC policy makers are involved in the project via participation in qualitative data/expert opinion and in arriving at consensus.

From evidence to policy on a national level: supporting the government's role in a learning healthcare system

Hultcrantz M[1], Hall U[2], Lagerstedt K[3], Richter Sundberg L[4], Larsson L[5], Hellberg L[2]

[1] Swedish Agency for Health Technology Assessment and Assessment of Social Services, Sweden; [2] Ministry of Health and Social Affairs, Sweden; [3] Health and Social Care Inspectorate, Sweden; [4] Umeå University, Sweden; [5] Lund University, Sweden

Background: In August 2018, the Swedish government appointed a special investigator to support a national ecosystem for evidence-based healthcare. In November 2019, the inquiry received an additional task now focusing on follow-up to enable a more strategic, evidence-based and long-term sustainable management of health care on a national level. Since healthcare systems around the world face similar challenges, we believe the results of the investigation are highly relevant for a broader audience.

Objectives: To analyze how governmental agencies can better support a comprehensive follow-up of health care and thereby create a learning system on a national level. This includes following the effects of the government's initiatives and reforms and analyzing where future governmental interventions are needed.

Methods: The investigator and her team have worked with an expert committee including representatives of governmental agencies, healthcare professions and healthcare providers, as well as a reference group of patient representatives. Additional information has been collected through questionnaires, workshops and meetings with stakeholders and other governmental inquiries working on related topics. Background information was collected from published research, governmental reports, existing regulations etc.

Results: Although roughly estimated more than 1000 full-time government employees work with follow-up in Sweden, the quality, effectiveness and equity in health care is not improving at the anticipated rate. Our findings suggest that a co-ordination of the different initiatives is a prerequisite for creating a learning system on a national level. Key challenges arise in the interface between evidence and policy. At the Colloquium, we will present possible solutions to these challenges using the Swedish healthcare system as an example.

Conclusions: Although a lot of efforts are made in conducting and developing methods for evidence generation, implementation and follow-up, it is apparent that the full value for patients is not reached. Facilitators are needed for a learning system on a national level, where the government's initiatives efficiently contribute to an increased quality, effectiveness and equity in health care. Sharing experiences from national efforts can be one way of increasing the understanding of what these facilitators are.

Patient or healthcare consumer involvement: We have received valuable input through regular meetings with a reference group of six representatives from different patient organizations throughout the work. The expert committee linked to the investigation also included a patient representative adding important perspectives to the discussions.

Future development through the past: analysis of financing dimension of Iran's health transformation plan

Doshmangir L[1], Najafi B[2], Mostafavi H[3], Moshiri E[4]

[1] Health Services Management Research Center, Iranian Center of Excellence in Health Management, School of Management and Medical Informatics, Tabriz University of Medical Sciences, Tabriz, Iran; [2] Health Services Management Research Center, School of Management and Medical Informatics, Tabriz University of Medical Sciences, Tabriz, Iran; [3] Shahid Beheshti University of Medical Sciences, Tehran, Iran; [4] Social Determinants of Health Research Center, Semnan University of Medical Sciences, Semnan, Iran

Background: On 15 Jun 2014, a significant reform called Health Transformation Plan (HTP) started in Iran to achieve Universal Health Coverage by 2025. To implement HTP, significant funding was allocated to the Ministry of Health and Medical Education (MoHME). However, before HTP, the budgeting model used for healthcare funding was affected by global oil prices, fluctuation in the exchange rates, and economic sanctions.

Objectives: To analyze the future of HTP in light of potential financial barriers and challenges.

Methods: We reviewed published and unpublished documents from social media, national policy-making organizations, including MoHME, parliament, medical universities, and some other related organizations. We also conducted face-to-face interviews with key informants, including health policymakers at different levels of the health system, health researchers, health insurance officers, and faculty members. We used both deductive and inductive approaches to analyze the data.

Results: Unsustainability in financial resources and continuing in the allocation of Gross Domestic Product (GDP) to the HTP is a challenge in continuing this policy reform. Also, the bureaucratic procedures in allocation and distribution of funds, discrepancy between health insurance funds and affiliated units of MoHME, especially Medical Sciences Universities (MSU), the lack of purchaser-provider split, inappropriate payments and compensation mechanisms, and dependence of MSU on individual revenues are the main challenges that threaten the future of HTP in Iran.

Conclusions: Given the current situation, where the country's funding resources are often unpredictable, it seems that such funding challenges will lead to poor HTP results. Having a holistic view, adopting and implementing evidence-informed policies and future-based strategies can prevent unexpected consequences.

Patient or healthcare consumer involvement: No patient or healthcare consumer was involved in the study.

How the Shenzhen government could become a pioneer in rapid-learning health systems

Fang Q[1], Wu LQ[1], Chen YL[2], Wang Q[3], Zhang JY[4], Ma YF[2], Wang JJ[4], Wu SY[2], Lavis J[3]

[1] Shenzhen Health Development Research Center, China; [2] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Cochrane China Network, China; [3] Department of Health Research Methods, Evidence & Impact (HEI), Faculty of Health Sciences, McMaster University; McMaster Health Forum, McMaster University, Canada; [4] School of Public Health, Lanzhou University; Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Cochrane China Network, China

Background: As China's first special economic zone (SEZ) that was set up in 1978 and a comprehensive reform pilot city, Shenzhen ranked fifth in Asia in terms of economic power. On 18 August 2019, China issued a guideline on supporting Shenzhen in building a pioneering demonstration zone for socialism with Chinese characteristics. Shenzhen is envisioned to be a "global benchmark city" that excels in competition, innovation and influence by the middle of this century, including to become a pioneer in health systems. In the aspect of strengthening health systems, it is important to learn from successful experiences, approaches, and mechanisms in other settings, and pilot and refine them in an iterative process. Based on the rapid growth development in Shenzhen (i.e. "Shenzhen Speed"), the development of a "rapid-learning and improvement" (RL + I) health system is suitable and will facilitate achieving Shenzhen's goal as a pioneer city, especially in health systems.

Objectives: To explore how the Shenzhen government could become a pioneer in RL + I health systems, in order to strengthen the health systems and promote evidence-informed policymaking (EIP).

Methods: We will conduct a rapid synthesis in a 60-business-day timeframe, including four steps: 1) proposing a question after the consultations with policymakers; 2) searching, selecting, assessing and synthesizing research evidence related to the question; 3) presenting findings of rapid synthesis; 4) finalizing synthesis with the input of reviewers. We will search MEDLINE, Web of Science, Chinese Biomedical Literature Database (CBM), Health Systems Evidence and grey literature.

Results: The rapid synthesis will cover a Shenzhen-appropriate definition of a RL + I health system, its characteristics, the assets that can be leveraged and the gaps that need to be addressed, and the potential 'windows of opportunity' for promoting the development of a RL + I health system in Shenzhen. For example, the expected characteristics of a RL + I health system in Shenzhen might be patient-centred, evidence and data-driven, Evidence-Informed Policy Network (EVIPNet) supported and so on. We will present the detailed and comprehensive findings of a RL + I health system in Shenzhen at the conference.

Conclusions: The development of a RL + I health system in Shenzhen will help to strengthen institutional capacities, create a municipality-level ecosystem of policy-supporting organizations, and collaboratively establish a community of practice to support evidence-informed policymaking about health systems. The Shenzhen government, as a pioneer in RL + I health systems, will achieve better population health and more sustainable development.

Patient or healthcare consumer involvement: One of the expected characteristics of a RL + I health system in Shenzhen is patient-engaged or empowered, which will put patient and public needs, preferences and perspectives at the center. We will embed patient and public involvement in everything we do.

Initiating population-based "5A" evidence-based clinical practice model: EBM anti-Coronavirus action in China

Fei Y[1], Liu J[1], Yu M[1], EBM AA[2]

[1] Beijing University of Chinese Medicine (BUCM), China; [2] Evidence-Based Chinese Medicine, China

Background: The COVID-19 pandemic is an unprecedented public health emergency. In the beginning of the epidemic in China, there was completely no direct experience and evidence that can be used as reference to guide medical actions. However, the experience and evidence gained from handling previous epidemic situations can be used as indirect evidence to support current practice, when waiting for direct scientific research evidence to be produced. At the same time, fast growing real-world first-hand experience of healthcare professionals needs to be collected, summarized, consented and disseminated as soon as possible to complement clinical decision making in a most efficient way. The "Ask-Acquire-Appraise-Apply-Assess"-"5A" evidence-based clinical practice (EBCP) model needs to be extended to a population-based "5A" EBCP model.

Objectives: To identify clinical needs for the medical prevention and control of COVID-19 infection, and to prioritize them and collaborate with evidence teams all over China to provide reliable information and evidence to frontline workers.

Methods: Evidence-Based Medicine (EBM) anti-Coronavirus Action was initiated calling on all Chinese EBM professional teams to collaborate to provide evidence for urgent clinical questions or information needs, which were collected by online PICO structured questionnaires sent out via public media WeChat to clinicians and nurses, mainly working in Wuhan city. We developed structured evidence summary tables to facilitate the work of the collaborative teams.

Results: EBM anti-Coronavirus Action was initiated on 20 February 2020, by Institute of Excellence in Evidence-Based Chinese Medicine. Twenty-four EBM teams, plus one public health team and one psychiatry team from 13 provinces/autonomous regions/ municipalities in China, and one health economics team from Canada, joint the action and contributed until 30 March 2020, 85 evidence summaries. The topics of those summaries were selected by each team from 184 questions we collected from the online survey, in which 1343 clinicians participated. The questions covered a very broad range, including transmission of the virus, identification and control the source of infectious pathogenesis, diagnosis, treatment, prognosis, social and psychological impacts, and prevention of secondary disasters, etc. The 58 evidence summaries that were sent out via WeChat were read 16,448 times. Fourteen media in China reported on EBM anti-Coronavirus Action.

Conclusions: The successfully initiated EBM anti-Coronavirus Action was based on the rapid responses of both the clinicians asking important queries and the collaborative evidence-providing teams answering the queries. This action serves adequately as a pilot for the theoretical population-based "5A" EBCP model, which might benefit the world in the continuing COVID-19 pandemic in clinical medicine area, as well as other health related areas.

Patient or healthcare consumer involvement: Clinicians in Wuhan were consulted and involved in the designing of the online questionnaires and the initiation of the action.

Lessons learnt from a rapid evidence synthesis service for health policymakers hosted by the World Health Organization Regional Office for Europe

Kruja K[1], Scarlett J[2], Sy TR[3], Sharma T[3], Kuchenmüller T[3]

[1] World Health Organization Regional Office for Europe, UK; [2] World Health Organization Regional Office for Europe, USA; [3] World Health Organization Regional Office for Europe, Denmark

Background: Health Evidence Network (HEN) synthesis report series is the World Health Organization Regional Office for Europe's (WHO/Europe) information service for public health decision-makers. Operating since 2003, it supports rapid learning in health systems by summarizing the best available evidence, on average within a nine-month publication timeline, to respond to policy questions. Sixty-seven reports have been published, covering high-priority topics in WHO/Europe. To better understand their impact on policy-making, standardized methods have been used to monitor their use and uptake since 2017.

Objectives: To better understand how HEN supports evidence-informed policymaking and identify lessons learnt from the HEN process to inform the improvement of HEN and other similar services.

Methods: A multi-method qualitative approach was used. HEN commissioners and authors were asked standardized questions after reports were published about the HEN process and how the reports were used in policymaking. It also drew on findings from HEN uptake summaries, through citation searches and website analytics, that track how HEN synthesis reports have been used and shared for one year from publication.

Results: Feedback from 10 authors and 8 commissioners on 16 reports and from 12 HEN uptake summaries identified that HEN reports have contributed to the development of technical guidelines, strategies, and other policy documents, served as a foundation for technical assistance and capacity building, and contributed to relationship building and opportunities for collaboration with other stakeholders. Commissioners felt that HEN's well-known brand and established structure added legitimacy to synthesis reports. The findings identified how HEN currently facilitates evidence-informed policymaking and areas where HEN can do so more effectively. The following lessons can be for improvement of HEN, and other evidence synthesis services:

- A co-ordinating team is crucial to maintain focus by facilitating communication between authors, commissioners and reviewers, and ensuring adherence to timelines.
- Author team diversity (expertise, country of origin, and language) contributes to a better grasp of the evidence, but may pose difficulties for working cohesively.
- Create consensus between authors and commissioners early on about the project scope, methods and limitations and provide tools, training and support to ensure that authors can derive policy considerations from the evidence.
- A platform for authors to stay involved in HEN and WHO processes longer term.
- Technical editing maintains a consistent style for reports.
- Evidence synthesis use should be evaluated 1 to 2 years post-publication.

Conclusions: Effective evidence synthesis can improve health policies which affect people's experiences, health outcomes, and costs. HEN synthesis reports offer valuable lessons for ensuring that evidence syntheses in the future are being carried out in the most effective and efficient way possible to meet the needs of policymakers and promote evidence uptake.

Methodological criteria for regulating orphan drugs for rare diseases: have EBM concepts been adopted? Zolgensma case study

Riera R[1], Pachito D[2], Bagattini Â[2]

[1] Centre of Health Technology Assessment (HTA), Hospital Sírio-Libanês, Universidade Federal de São Paulo, Brazil; [2] Centre of Health Technology Assessment (HTA), Hospital Sírio-Libanês, Brazil

Background: In 2019, Zolgensma was approved in the United States for treating spinal muscular atrophy (SMA), based on two open-label (unmasked), non-comparative studies. Some methodological criticisms have been raised from this decision. In contrast, facing both a rare disease and an orphan drug has been a justification for accepting less robust and reliable evidence.

Objectives: To discuss the methodological issues related to regulatory process of drugs for rare diseases, based on the Zolgensma case for treating SMA.

Methods: Critical appraisal study carried out at the Centre of Health Technology Assessment (HTA), Hospital Sírio-Libanês (São Paulo, Brazil).

Results: Open-label, unmasking and single-arm studies are far from the methodological rigor of a randomized double-blind controlled trial, considered the most appropriate primary study to assess the effects of healthcare interventions. Firstly, the presence of a similar comparator group is essential to estimate the real effect of the intervention and to assess if these effects are different from those observed with the use of best available option, placebo or natural course of the disease. Secondly, the similarity between the comparison groups helps to ensure that the effect observed can be exclusively attributed to the intervention, eliminating any other factor that differentiates the groups, such as disease severity, sex, age or comorbidities. The use of an adequate randomization method would certainly achieve this goal. Thirdly, the lack of allocation concealment can overestimate the size of the intervention effect by 37% to 41% (1,2). This means that, depending on the point estimate of the effect, an intervention that, in reality, has no benefits, may prove to be falsely beneficial. Fourthly, unmasking participants, personnel and outcome assessors lead to deviations in the process of conducting the study, such as adherence to treatment, reporting of adverse events, and also biased judgement of the outcomes. Here we pointed out four of the various biases (3) to which these studies may be exposed, impacting the direction and size of the findings, thus decreasing their reliability.

Conclusions: The case of Zolgensma raises a debate about the methodological criteria adopted by global regulatory agencies for approving the marketing of drugs, specifically in the rare disease scenario, and orphan drugs, at extremely high costs. Critical points related to the methods for planning and conducting the clinical studies that supported the regulatory process for drugs must be identified and revised.

Patient or healthcare consumer involvement: It is essential that patients, managers and health decision-makers understand the process of drug regulation in their country, as well as know the uncertainties related to the findings of pivotal studies used to support approvals.

References:

1. Moher, 1998
2. Schulz, 1995
3. Catalogue of Bias Collaboration. Catalogue of Bias. 2019. catalogofbias.org/biases/attrition-bias.

Rapid review report on early COVID-19 status in the Republic of Korea and global pandemic

Choi M[1], Shin HG[1], Jin Y[1]

[1] National Evidence-Based Healthcare Collaborating Agency (NECA), South Korea

Background: Since the first confirmed COVID-19 patient who traveled from Wuhan China was reported on 20 January 2020, COVID-19 infection has increased during February and March in the Republic of Korea.

Objectives: To provide information on the characteristics of the first two months of COVID-19 prevalence in Korea and to examine preliminary evidence from various sources in this unclear situation.

Methods: We used public data available from the Korea Center for Disease Control and Prevention (KCDC) and situation reports from World Health Organization (WHO) from February to 24 March. For additional information, we used health utilization data from OECD (Organization for Economic Co-operation and Development) for subgroup analysis. We performed a proportion meta-analysis. We searched PubMed, KoreaMed and CNKI (China National Knowledge Infrastructure) to identify the epidemiological characteristics of COVID-19 and treatment strategies. We also monitored the recommendations of domestic and global disease control institutions. We updated the search results and reports every two weeks.

Results: In Korea, the ratio of confirmed cases divided into two groups: before and after the occurrence of a large cluster infection explosion on 16 February from a religious group called Shincheonji Church. After WHO declared a global pandemic on 11 March 2020, the fatality rate of COVID-19 seems to be related to the number of hospitalizations per 1000 population. From the literature review, we identified a strong reproduction rate, asymptomatic period of infection, rate of exacerbated and current treatment.

Conclusions: The COVID-19 pandemic in Korea was inevitable. Following the early explosion of infection, the infection rate dropped thanks to rigorous tracing, widespread testing and acceptable health.

Patient or healthcare consumer involvement: None

Additional file: Figure

Summary of recommendations for breastfeeding of infants born to mothers with COVID-19

Yang N[1], Che S[2], Zhang J[3], Wang J[3], Ma Y[1], Wang M[4], Chen Y[1]

[1] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Cochrane China Network; Chinese GRADE Center, China; [2] Children's Hospital of Chongqing Medical University, China; [3] School of Public Health, Lanzhou University; Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Cochrane China Network; Chinese GRADE Center, China; [4] Department of Radiology, the First Hospital of Lanzhou University, China

Background: Breast milk is the best source of nutrition for infants, and numerous studies have shown that breastfeeding has multiple benefits. Cases of COVID-19 among pregnant and lactating women have also been confirmed. Physicians and lactating women need to make decisions whether or not to continue to breastfeed.

Objectives: To review existing recommendations form national or international authorities for breastfeeding of infants born to mothers with COVID-19.

Methods: Two review authors searched MEDLINE (via PubMed), Embase, Web of Science, China Biology Medicine disc (CBM), China National Knowledge Infrastructure (CNKI), and Wanfang Data. We also searched the following websites for relevant publications: World Health Organization (WHO), US Center for Disease Control (CDC), China National Health Commission (NHC), Google Scholar from 1 January 2020 until 4 April 2020. Two review authors independently screened the titles, abstracts and full-text articles; to identify and extract potentially relevant guidelines.

Results: Overall, 10 guidelines from five countries and WHO were included in the final review. China NHC guideline recommended that mothers with COVID-19 should be isolated, quarantined and suspend breastfeeding. Three guidelines from China CDC recommended that newborns be kept in isolation, pasteurized breastmilk from a milk bank can be used. One of two guideline from the US CDC and the UK Royal College of Obstetricians and Gynaecologists (RCOG) guideline indicated whether and how to start or continue breastfeeding should be determined by the mother and healthcare providers. Another US CDC guideline, the RCOG guideline, and the three guidelines from WHO, the Italian Society of Neonatology and German Nationale Stillkommission recommended continue to breastfeed with appropriate precautions (wash hands and wear mask). If their own health does not permit, mothers should be supported to express milk and allow healthy caregivers to breastfeed.

Conclusions: There is no consensus on whether mothers with COVID-19 should continue breastfeeding. High-quality evidence from systematic reviews of breastfeeding for COVID-19 mothers is needed to support recommendations.

What is going on in the future for evidence-informed health policy-making?

Doshmangir L[1], Yazdizadeh B[2], Sajadi HS[3], Mohtasham F[2], Majdzadeh R[2], Lavis J[4]

[1] Health Services Management Research Center, School of Management and Medical Informatics, Tabriz University of Medical Sciences, Tabriz, Iran; [2] Knowledge Utilization Research Centre, Tehran University of Medical Sciences, Tehran, Iran; [3] Knowledge Utilization Research Centre, University Research & Development Center, Tehran University of Medical Sciences, Tehran, Iran; [4] Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, Canada

Background: Various global and local trends are affecting health policy decisions and shaping the future of health systems. Identifying trends is imperative to predict emerging issues or challenges and plan to enhance the use of evidence in order to address them better.

Objectives: To determine the main trends that are likely to affect evidence-informed health policy-making (EIHP) in Iran and to identify the main challenges or opportunities for EIHP that Iran will face.

Methods: We conducted a content analysis of relevant documents to outline the list of trends that may affect the Iran health system. Then, we organized two focus group discussions and six face to face interviews with relevant national experts to discuss the trends affecting EIHP in Iran, the effect that they potentially may have, and the opportunities and threats emerging from these trends. We used framework analysis to analyze data. During the development of subthemes we coded data, analyzed memos, and generated subthemes based on the elements of the framework. We identified five themes and 15 subthemes.

Results: Five trends were likely to affect EIHP in Iran. Social, technological, and economic trends influence the EIHP more directly than political and environmental trends. The main challenges for EIHP, caused by these trends, are a growing need for more localized and high-quality research evidence in a reasonable time, more stakeholder participation and more capacity building for the production and use of relevant research evidence.

Conclusions: Each of the identified trends has various impacts on the use of research evidence in health policies. Social, technological, and economic trends have more effects and cause some challenges. More evidence will be needed to address public health problems. Furthermore, there will be more access to data and evidence. It shows the need to pay more attention to capacity building for EIHP.

Patient or healthcare consumer involvement: No patient or healthcare consumer was involved in this study.

HEALTH TECHNOLOGY ASSESSMENT

A living systematic review to support health technology assessment: The CADTH experience

Weeks L[1], Kim J[1], Lachance C[1], Carson E[1]

[1] Canadian Agency for Drugs and Technologies in Health, Canada

Background: Living systematic reviews (LSRs) – systematic reviews that are continuously updated as new evidence becomes available – are increasingly being used to support clinical practice and guideline development. LSRs are also well suited to support health technology assessment (HTA), although to date we are unaware of an LSR being conducted by an HTA organization.

Objectives: To describe the Canadian Agency for Drugs and Technologies in Health (CADTH)'s experience conducting their first LSR as part of an HTA on the topic of stereotactic ablative radiotherapy for the treatment of oligometastatic cancer. To discuss challenges and lessons learned.

Methods: Throughout the development of the HTA protocol, targeted conversations were had between the LSR clinical research team and individuals working on other aspects of the HTA, including health economists, qualitative researchers, medical librarians, project managers, knowledge mobilization officers, patient engagement officers, and members of the publishing team. The purpose was to explore process challenges and implications of the living model on other aspects of HTA conduct and reporting. We took detailed notes and documented decisions. Throughout the conduct of the LSR, through bi-weekly team meetings, we identified challenges and brainstormed solutions.

Results: We followed the methods outlined in the Guidance for the Production and Publication of Cochrane Living Systematic Reviews, with some adjustments to suit the CADTH and HTA context. For example, monthly search updates were not deemed possible due to competing demands of the medical librarians. Monthly alerts are instead being conducted during the conduct of the baseline review and, once the review transitions to living mode, updates will be conducted every three months for electronic databases and every six months for grey literature. Decision rules were established a priori, including both qualitative and quantitative signals, to guide decisions about when to incorporate new evidence into an analysis. The decision for the review to remain in living mode will be revisited annually and informed by the level of priority for decision-makers, the level of uncertainty that remains in the clinical evidence, and the existence of ongoing studies. Several process changes were required including to project schedule templates, engagement of clinical experts and peer reviewers, patient engagement, publishing, and knowledge mobilization.

Conclusions: With rapid health technology innovation and increasing attention to reassessment and disinvestment decisions, new models for evidence synthesis are needed. LSRs show promise to support HTA and will be used as appropriate in future CADTH HTAs, but processes and methods must also balance rigour and timeliness.

Patient or healthcare consumer involvement: Throughout the conduct of the LSR, CADTH is engaging people with oligometastatic cancer, guided by the belief that patients have knowledge, perspectives, and experiences that are unique and contribute essential evidence for HTA.

A standardized HTA report summary for rapidly presenting outside findings

Mitchell MD[1], Mull NK[1]

[1] University of Pennsylvania Health System, USA

Background: Health technology assessment (HTA) agencies have recognized past redundancies in evaluating new drugs, devices, and other technologies, and now understand that collaboration will increase the efficiency and sustainability of the HTA enterprise. This is called out specifically in the EUnetHTA Joint Action 3 objectives as "re-use" of published HTA results by other national and regional agencies. Our hospital-based evidence-based practice center is pursuing the same goals as the national and regional centers of EUnetHTA. We receive more requests for HTA reports than we are able to complete ourselves; but at times, we can meet a requestor's needs with a previously published HTA report. This can be challenging in several ways. First, the reports may include analysis based on product costs and health system implementation strategies in the country where the HTA report was produced and therefore may be only indirectly applicable to our end-users. Second, reports vary in how they describe the evidence and present it to end-users. Finally, those HTA reports may be too long and time-consuming for our clinician and administrative clients to read in full.

Objectives: To develop a standardized product for adaptation of existing HTA reports and their findings. Our aim was to provide HTA information to committee members and other decision-makers in a way that they could quickly understand both the conclusions of the reports and the strengths and weaknesses of the evidence base.

Methods: In an iterative process, we developed a new evidence report product. The Rapid Product Summary has three key features: a standard report format, a simplified ten-item instrument for reporting on the quantity and strength of evidence on the technology in question, and a rapid topic acquisition and report generation schedule that would fit within a typical 30-day committee meeting cycle. The structure of the report has a degree of flexibility to accommodate diverse types of requests.

Results: All Rapid Product Summaries completed to date have been progressed from acceptance of topic request to finished draft report in less than 30 days. Each of the Rapid Product Summary reports was three to four pages long, with all summary points included on the first page. Summarized HTA reports included ones from NICE, CADTH (Canadian Agency for Drugs and Technologies in Health), the ECRI Institute, and SBU (Swedish health technology assessment agency). We were unable to provide a Rapid Product Summary in response to about one-third of requests because no suitable source reports were identified.

Conclusions: By finding, adapting, and re-using existing published reports, local evidence-based practice centers can provide timely support for decisions about new and emerging technologies.

Additional file: [Rapid summary scale example](#)

An evaluation study of using semi-automation systematic tools for effective Health Technology Assessment in Korea

Choi M[1], Park J[1]

[1] National Evidence-Based Healthcare Collaborating Agency (NECA), South Korea

Background: Since systematic reviews need manual and time-consuming work, especially in study selection process and quality assessment, interest in semi-automation tools for systematic reviews has increased.

Objectives: We aimed to figure out the usefulness of semi-automation tools for systematic reviews in Health Technology Assessment (HTA) in Korea for effective systematic reviews.

Methods: We established an internal steering group including seven members who have experience on several HTA reports. We are investigating current HTA systematic review process to figure out the possible point of process for application of semi-automation tools. We will perform a pilot of systematic review to test various existing semi-automation tools. Finally, the internal-steering group will recommend a new protocol for the effective systematic review process for HTA in Korea.

Results: This study started in April 2020 and preliminary results will be available in early September.

Conclusions: We would like to recommend a process for conducting effective HTA systematic reviews using semi-automation tools in Korea.

Patient or healthcare consumer involvement: None.

Centre of Health Technology Assessment of Hospital Sírio-Libanês: using EBM to qualify and promote the sustainability of Brazilian public health system (SUS)

Riera R[1], Bagattini Â[2], Pachito D[2]

[1] Centre of Health Technology Assessment (HTA), Hospital Sírio-Libanês; Universidade Federal de São Paulo, Brazil; [2] Centre of Health Technology Assessment (HTA), Hospital Sírio-Libanês, Brazil

Background: PROADI-SUS is an innovative program, from the Brazilian Ministry of Health, through which five non-profit hospitals, including Hospital Sírio-Libanês (HSL), began to conduct projects using their skills to qualify and promote the improvement of the Brazilian public health system (SUS). In order to provide scientific and methodological support for these projects, based on the principles of evidence-based medicine (EBM), the Centre of Health Technology Assessment in HSL (C-HTA-HSL) was created.

Objectives: To describe the creation of the HTA Centre in Hospital Sírio-Libanês and to present the initial activities and the main EBM-related projects implemented so far.

Methods: Descriptive case study.

Results: The C-HTA-HSL was created in 2018 and it is currently a formal member of Brazilian HTA Network (REBRATS) and HTA Network of the Americas (RedETSA). Since its launch, the centre has supported the following projects of PROADI-SUS:

1) through a collaboration with Agência Nacional de Vigilância Sanitaria (Anvisa, the Brazilian national health surveillance and regulatory agency), we are mapping the evidence on the efficacy and safety of the use of platelet-rich plasma (PRP) for non-transfusion purposes in order to support the regulatory framework for the use of PRP in Brazil;
2) through a collaboration with the National Council of Justice, we are implementing a set of strategies to provide evidence-based support during the judicial process on health care in Brazil;
3) through a collaboration with Agência Nacional de Saúde Suplementar (ANS, the Brazilian regulatory agency for health insurance), we are providing evidence-based advice to the technical staff of the agency during the process of evaluating proposals for incorporating technologies into the list of products offered to users; in addition, we are reviewing the content of the scientific reports about evidences on clinical effects, cost-effectiveness and budgetary impact of 40 proposals; and
4) through collaboration with the local laboratory of tissue bioengineering, we are in charge of methodology for planning, conducting, analysing and reporting costs-effectiveness analyses of two randomized controlled trials: use of stem cells for children with cleft lip and palate and use of stem cells for children with craniofacial microsomia.

Conclusions: We believe that the creation of C-NATS-HSL can enhance the methodological rigor of the projects of PROADI-SUS, by the systematic adoption of EBM concepts, so that reliable research can support the decision-making process within a public, universal and equitable health system.

Patient or healthcare consumer involvement: We believe that the creation of C-NATS-HSL can enhance the methodological rigor of the projects of PROADI-SUS, by the systematic adoption of EBM concepts, so that reliable research can support the decision-making process within a public, universal and equitable health system.

Support: Brazilian Ministry of Health / PROADI – SUS, Hospital Sírio-Libanês.

Reporting quality of case series using computer tomography in patients with coronavirus disease 2019

Wang M[1], Luo X[2], Lv M[3], Lei J[1], Tian J[4], Chen Y[4]

[1] Department of Radiology, the First Hospital of Lanzhou University, Lanzhou, China; [2] School of Public Health of Lanzhou University, Lanzhou; Cochrane China Network for GRADE and Guideline Working Group, Lanzhou, China; [3] School of Public Health of Lanzhou University, Lanzhou, China; [4] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou; Cochrane China Network for GRADE and Guideline Working Group, Lanzhou, China

Background: In late 2019, a pneumonia caused by a novel coronavirus (SARS-CoV-2) emerged in Wuhan, China, and rapidly spread to 23 countries around the world. A large number of case series studies using computer tomography (CT) in patients with coronavirus disease 2019 (COVID-19) were published. However, the reporting quality of those studies is unknown.

Objectives: To evaluate the reporting quality of case series studies reporting widely used CT for COVID-19 patients.

Methods: We searched MEDLINE (via PubMed), Embase, the Cochrane library, Web of Science, Wanfang Data, CNKI (China National Knowledge Infrastructure), and CBM (Sinomed) between 1 January 2020 and 1 March 2020. We included case series studies focused on chest CT imaging for patients with COVID-19. We used the revised Preferred Reporting Of CasE Series in Surgery (PROCESS) guideline to assess the reporting quality of case series studies on COVID-19.

Results: We included 33 case series studies. Only 21.2% (7/33) of the studies used the term "case series" in the title, no article used "case series" as a keyword, and only one literature in the research method mentioned the type of study as a case series. A total of 60.6% (20/33) of the included studies were not comprehensively described the methods in the abstract, such as describe when it was done and by whom. No studies reported the trial registration. Almost half (48.5%, 16/33) of the included studies didn't perform ethical approval. Only 27.3% (9/33) studies reported the prospective or retrospective design, the single or multi-center, and consecutive or non-consecutive, however most of studies did not report whether the cases are consecutive or non-consecutive. There were 30.3% (10/33) of studies that did not report where the study was conducted, and the remaining only mentioned the name or geographical location of the hospital, but not the nature of the institution; 75.8% (25/33) did not report the exclusion criteria for participants or the epidemiological history/severity of the disease and other comorbid conditions; 57.6% (19/33) of image examination methods and parameters/contents and methods of image evaluation were not reported or incomplete; 21.2% (7/33) did not report the experience of image evaluators, and 42.4% (14/33) did not have quality control and consistency checks. Most of the literature (32/33) in the discussion section summarized the key results and 81.8% (27/33) of the articles compared with previous studies, but only 54.5% (18/33) articles discussed the advantages and limitations of research. 81.8% (27/33) studies did not declare conflicts of interest or funding sources.

Conclusions: The reporting quality of the case series studies in the field of CT diagnosis of COVID-19 is poor. Future studies should follow the reporting guideline of case series. In addition, the PROCESS guidelines are not applicable to the imaging field. A case series reporting guideline for imaging will be developed later.

Patient or healthcare consumer involvement: None.

Systematic review of endoscopic ultrasound guided radiofrequency ablation for pancreatic tumors

Shin H[1], Han H[1], Lee I[2], Jang J[3], Do J[4], Jung W[5], Choi J[2], Koh Y[6], Lee W[1], Shin C[1]

[1] National Evidence-Based Healthcare Collaborating Agency, Seoul, South Korea; [2] The Catholic University of Korea, St. Mary's Hospital, Seoul, South Korea; [3] Seoul National University Hospital, South Korea; [4] Chung-Ang University Hospital, Seoul, South Korea; [5] Samsung Medical Center, Seoul, South Korea; [6] National Police Hospital, Seoul, South Korea

Background: An insulinoma is the most common functional pancreatic neuroendocrine tumor (pNET). Increased fasting serum insulin and C-peptide assist in the diagnosis. Surgery is the standard care for pNET. However, it can lead to complications such as diabetes mellitus. Endoscopic ultrasound-guided radiofrequency ablation (EUS-RFA) is a promising therapy for the treatment of pancreatic neoplasms.

Objectives: To assess the safety and clinical effectiveness of EUS-RFA in pancreatic tumors.

Methods: We conducted a systematic review based on searching eight national databases including KoreaMed and three international databases including Ovid-MEDLINE, Ovid-Embase, and the Cochrane Library. After having excluded 1583 articles (including 386 duplicated articles), we reviewed 11 articles (7 case series and 4 case reports) in the final assessment. The article quality was evaluated using Scottish Intercollegiate Guidelines Network (SIGN), and the level of evidence and grade of recommendation were carefully determined and documented based on the quality.

Results: There was no procedure-related death. One of the studies reported severe complications (10%, acute pancreatitis with fever, perforation, a stenosis of the main pancreatic duct), but those that were appropriately treated and the tumors had completely disappeared at follow-up. After the therapy, clinical response with normalization of glucose levels was observed in all insulinoma cases. Also, radiological complete response was reported in most participants and some cases remained symptom free at 10 to 12 months of follow-up. As a result, EUS-RFA has an effectiveness to improve symptoms through ablating the tumors. We suggest that EUS-RFA can be used before the surgery. On the other hand, there was a lack of evidence to assess clinical effectiveness in pancreatic cancer and the studies reported a low complete response rate (25 ~ 64.7%) in non-functional pNET and cystic tumors. Therefore, in pancreatic adenocarcinoma, non-functional pNET and cystic neoplasms, further research is required in the form of long-term follow-up (more than five years) study design (Level of Evidence D).

Conclusions: EUS-RFA of insulinoma with symptom (≤ 2cm) is a minimally invasive, safe and clinically effective intervention. More data about clinical effectiveness and long-term results are needed in the other indications.

The effectiveness and safety of external trigeminal nerve stimulation (e-TNS) for preventive treatment of migraine

Park E[1], Jeon M[1]

[1] National Evidence-Based Healthcare Collaborating Agency, South Korea

Background: Migraine is a very common disease, approximately 15% of people are affected globally. Current available preventive antimigraine drugs have limited efficacy and have unpleasant adverse effect or contraindications. Due to these limitations, patients need safe and effective non-pharmacological therapy. The external trigeminal nerve stimulation (e-TNS) is reported effective for pain relief and patient satisfaction.

Objectives: This study aimed to evaluate the effectiveness and safety of e-TNS for preventive treatment of migraine.

Methods: We searched MEDLINE, Embase, the Cochrane Library and Korean databases, including KoreaMed, RISS, KISS, KISTI, KMBASE on 18 March 2019. The risk of bias was assessed by SIGN methodology checklist.

Results: Our literature search retrieved 1299 articles. The final nine articles were selected according to the selection and exclusion criteria. We assessed effectiveness for the parameters of migraine symptomatic improvement, including the changes in migraine duration, pain intensity, drug intake and frequency of migraine attack, based on eight articles. In a study comparing e-TNS and Flunarizine (a migraine preventive drug), changes in the number of migraine days and the intakes of acute migraine medications decreased similarly, but were not statistically significant, and the pain intensity measured with VAS-10 showed a significantly greater decrease in the Flunarizine group (P = 0.01). Compared with the sham control, the e-TNS group had a significant reduction in migraine days (P = 0.023), acute migraine medication intake (P = 0.0072) and frequency of migraine attacks (P = 0.044). In the single-arm studies, most studies reported a significant improvement in symptoms after using e-TNS. We assessed safety for procedure-related side effects based on five articles. The reported side effects included paresthesia, arousal and sleep disorder, skin rash or pressure feeling in the area of electrode attachment. The most frequently reported side effect was local paresthesia in five studies, and the rate was reported as 1.34% to 34.39%.

Conclusions: The effectiveness may be expected if e-TNS is used as a treatment alternative in patients with adverse effects or contraindications to migraine drugs. However, this doesn't have sufficient evidence to verify the effectiveness, more study results from RCTs compared with the sham control are needed. The safety is acceptable since the selected articles reported only mild side effects and a technology based on similar principles is already in use for other indications. The New Health Technology Assessment Committee deliberated that e-TNS is safe, but further research is required to verify the effectiveness.

Patient or healthcare consumer involvement: This study was performed with two researchers independently and subcommittee of seven members (neurologists, neurosurgeons, anesthesiologist, psychiatrist, and pediatrician). This conclusion will help migraine patients and clinicians to make clinical decisions.

Additional file: PRISMA flow diagram

The implementation of a learning system for early intervention services for psychosis in Quebec

Ferrari M[1], Iyer SN[1], Roy M[2], LeBlanc A[3], Abdel-Baki A[4]

[1] Department of Psychiatry, McGill University, Canada; [2] Department of Psychiatry and Neurosciences, Laval University, Canada; [3] Department of Family Medicine and Emergency Medicine, Faculty of Medicine, Laval University, Canada; [4] Department of Psychiatry, Université de Montréal, Canada

Background: Accessing evidence-based care is a fundamental principle of the Canadian universal healthcare system. Unfortunately, 26% of Canadians who experience mental illnesses have reported that they have unmet needs in mental health care. Early intervention services (EIS) for psychosis are widely recognized as being more effective for treating early psychosis than routine care. EIS attempt to decrease delays to treatment access, actively engage service users and families, offer appropriate patient-staff ratios, and integrate evidence-based psychosocial interventions. Although many provinces have implemented EIS, studies have revealed heterogeneity in the implementation of the essential components of EIS identified by international experts and guidelines.

Objectives: The Rapid Learning Health System (RLHS) for Quebec EIS project aims to address this gap by systematically collecting real-time data in eleven EIS programs in Quebec using an electronic data capturing platform and by building capacity to use this data for shaping clinical practices through targeted continuous education. Specifically, to 1) determine the feasibility of implementing a RLHS aimed at improving the quality of care of EIS, and 2) assess the impact of the RLHS on service compliance with standards of care across eleven Quebec EIS.

Methods: The RE-AIM model informs and guides data collection on the implementation process and impact of this RLHS. RedCap, a user-friendly electronic data capturing, repository and reporting system, will capture selected indicators of service quality. Simultaneously, Dialog+, an e-intervention, will provide feedback to and from service users and providers, to promote quality of life and care and to support shared decision making and measurement-based care.

Results: We will present learnings and preliminary descriptive data following the five RLHS phases: Scan (Performed knowledge synthesis activities and need assessment in services), Design and Implement (Stakeholders engagement, open communication to keep stakeholders informed and tailored capacity-building activities aimed at improving quality of services), and Evaluate and Adjust (Description of the data gathered, reporting system and feedback to services). We will focus on knowledge synthesis activities that informed this project; strategies to engage patients, families, and front-line staff; and IT technologies and data.

Conclusions: A RLHS can improve the uptake of clinical guidelines and evidence-based interventions in clinical settings and the translation of knowledge into practice. The RLHS approach has been shown to promote innovation in healthcare. This project will contribute to new evidence of its impact in the mental healthcare context. It will assess how real-time data and a learning community of EIS across Quebec can share best practices to improve clinical and yield province-wide outcomes.

Patient involvement: Patients, families, clinicians, and decision-makers have been involved to enhance the relevance, uptake, and sustainability of this project.

GUIDELINE DEVELOPMENT

A modeling approach to derive baseline risk estimates for GRADE recommendations: concepts, development, and results of its application on a guideline

Morgano GP[1], Wiercioch W[1], Anderson DR[2], Brożek JL[1], Santesso N[1], Xie F[1], Cuker A[3], Nieuwlaat R[1], Akl EA[4], Darzi A[1], Yepes-Nuñez JJ[5], Exteandia-Ikobaltzeta I[1], Rahman M[6], Rajasekhar A[7], Rogers F[8], Tikkinen KAO[9], Yates AJ[10], Dahm P[11], Schünemann HJ[12]

[1] Department of Health Research Methods, Evidence and Impact, McMaster University, Canada; [2] Department of Medicine, Dalhousie University, Canada; [3] Department of Medicine and Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, USA; [4] Department of Internal Medicine, American University of Beirut, Lebanon; [5] School of Medicine, Universidad de los Andes, Colombia; [6] Lillian S. Wells Department of Neurosurgery, University of Florida, USA; [7] Division of Hematology and Oncology, Department of Medicine, University of Florida, USA; [8] Trauma and Acute Care Surgery, Penn Medicine Lancaster General Health, USA; [9] University of Helsinki and Helsinki University Hospital, Finland; [10] Department of Orthopedic Surgery, University of Pittsburgh Medical Center, USA; [11] Department of Urology, University of Minnesota, USA; [12] Department of Medicine and Department of Health Research Methods, Evidence, and Impact, McMaster University, Canada

Background: Baseline risks are required to calculate absolute effect estimates, which are essential elements of evidence summaries produced for guideline panels. Systematic reviews of prognostic observational studies are scarce and the available estimates are often not directly applicable to patient-important outcomes. In some contexts, guideline panels revert to using surrogates to estimate baseline risks but this approach may introduce bias in the estimates of anticipated absolute effects.

Objectives: To develop an approach to model baseline risks for patient-important outcomes prioritized for recommendations when only baseline risks for surrogate outcomes are available.

Methods: This study was part of the American Society of Hematology (ASH) guidelines for the management of venous thromboembolism (VTE). The McMaster University GRADE Centre and the ASH guideline panel for the prevention of VTE in surgical patients developed a modeling approach based on explicit assumptions about the distribution of symptoms, anatomical location, and severity of VTE events.

Results: We applied the approach to derive modelled estimates of baseline risk. We used these estimates to calculate absolute measures of anticipated effects that informed the discussion of the evidence and the formulation of 30 recommendations. The approach increased transparency and reduced potential error in the decision-making process.

Conclusions: Our approach can assist guideline developers facing a lack of information about baseline risk estimates that directly apply to outcomes of interest. It also addresses potential bias of over- or underestimating absolute anticipated effects of interventions that can result from the use of surrogate data.

Patient or healthcare consumer involvement: Patients representatives were included in the guideline panel and contributed to the development of the model assumptions.

Additional file: visual abstract

A protocol for the practical application of human rights in World Health Organization guideline development

Thomas R[1], Thomas Bosco R[2], Takawira C[3], Peretz O[2], Garner P[4]

[1] Cochrane Infectious Diseases Group, UK; [2] WHO Guidelines Review Committee, Switzerland; [3] Liverpool School of Tropical Medicine, UK; [4] Cochrane Infectious Diseases Group, UK

Background: Equity, rights, gender and societal (ERGS) considerations are central to achieving global health and well being as well as more equitable distribution of good health. The World Health Organization (WHO) is mandated to ensure incorporation of ERGS into its guidelines. However, practical and effective implementation when developing global guidelines that cover a vast array of topics is challenging. In this collaborative project, we sought to evaluate how ERGS standards and principles have been used in WHO guideline development and what progress has been made since the adoption of formal guidance in the WHO Guideline Development Handbook (Chapter 5). We developed a protocol to guide this review that can be used to both evaluate guidelines, and as a prompt for guideline developers. The key innovation in our work is a protocol-driven set of ERGS expectations that can be tailored to different topics to drive guideline development.

Objectives: (1) To assess adherence to ERGS principles in 12 recent WHO guidelines, in terms of process and in the final product (content), and identify change over time. (2) To develop a protocol to effectively incorporate ERGS in: a) planning a guideline; b) conducting a guideline meeting; and c) evaluating guidelines.

Methods: Methods sought to address both procedural and content aspects of the guideline. Process evaluation: WHO has identified a set of process-based benchmarks to assess adherence to its guidance on human rights in the guideline development. These benchmarks will be used to evaluate the guideline development process including the developers composition. Content evaluation: We applied a modified version of domain 1 of the AGREE II tool combined with a set of prompts (Table 2) related to human rights.

Results: We will present results across the 12 guidelines evaluated detailing the full methods of the ERGS process evaluation and the content evaluation. We will discuss strengths and limitations of the method and suggest implications for a) planning a guideline; b) conducting a guideline meeting; and c) evaluating guidelines, and specifically how it can be used to inform recommendations.

Conclusions: Guideline adherence to ERGS principles varies greatly. We apply a novel evaluation methodology to illustrate how WHO incorporates these principles in guideline development, and propose new methods for meaningful inclusion in future guidelines.

Patient or healthcare consumer involvement: The proposed approach draws on professionals from within countries who experience health care as providers and as consumers. We intend to pilot consumer refereeing of the protocols in the development process.

Additional file: Table: content evaluation

Adaptation of a British Medical Journal (BMJ) Rapid Recommendation: challenges and perspectives

Ussing A[1], Andersen HK[2], Tarp S[1]

[1] Danish Health Authority and The Parker Institute, Denmark; [2] Danish Health Authority, Denmark

Background: The Danish Health Authorities have a desire to adapt high-quality and clinically important international guidelines into a Danish context to achieve national guidelines in an accelerated time course. The BMJ has developed the concept "Rapid Recommendations" where new and potentially practice-changing evidence is transformed into clinical guideline recommendations using the GRADE approach. The Rapid Recommendation team conducts systematic reviews to inform their guideline panel. We selected the BMJ guidelines and the underlying systematic reviews for adaption, since their methods fulfill our demands for transparency and quality.

Objectives: To develop a new clinical guideline based on adaption of a Rapid Recommendation guideline into a Danish context using the GRADE method without updating the underlying "body of evidence".

Methods: In a pilot project the Rapid Recommendation for oxygen therapy (1) was adapted to a Danish context. A guideline panel representing relevant medical specialties and methods specialists assessed the "body of evidence" from the systematic review underlying the rapid recommendation (2). The panel rated the importance of each outcome. The risk of bias assessments of the included trials and meta-analysis were reviewed, but only critical changes were allowed. Finally, the quality of evidence for each outcome were rated and the panel formulated their own recommendations.

Results: After public hearing and external peer review the new guideline was published, the process was completed within seven months. The adaption process revealed a need to clarify to the guideline panel, that the adaption was not a translation of recommendations, but a process of formulating new recommendations based upon their own assessments of the "body of evidence". The working group did not always agree with the risk of bias assessment or study inclusion in the underlying review, which led to rating down the quality of evidence.

Conclusion: Adaption of the BMJ rapid recommendation was feasible and enabled the development of a new guideline in an accelerated time course. Challenges in the adaption process were revealed but could be handled by information of the process, emphasizing that the adaption method implies independent assessment of the "body of evidence" and formulation of own recommendations.

Perspectives: New clinical recommendations and guidelines can be developed based on new high-quality systematic reviews and guidelines, leading to a time timesaving process.

Patient or healthcare consumer involvement: People with lived experience of hospitalization for acute medical conditions were members of the original guideline panel (1). Patient organizations were invited to the public hearing of the Danish guideline.

References:

1. Siemieniuk RAC, Chu DK, Kim LHY, Güell-Rous MR, Alhazzani W, Soccal PM, et al. Oxygen therapy for acutely ill medical patients: a clinical practice guideline. *BMJ* 2018;363:k4169.
2. Chu DK, Kim LH, Young PJ, Zamiri N., Almenawer SA, Jaeschke R., et al. Mortality and morbidity in acutely ill adults treated with liberal versus conservative oxygen therapy (IOTA): a systematic review and meta-analysis. *Lancet* 2018;391(10131):1693-1705.

An evaluation of methodological and reporting quality of guidelines in rare disease

Shi Q[1], Zhou Q[1], Wang Z[2], Zhao S[3], Xun Y[2], Lu S[4], Chen Y[5]

[1] The First School of Clinical Medicine, Lanzhou University, China; [2] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, China; [3] School of Public Health, Lanzhou University, China; [4] School of Medicine, University of Electronic Science and Technology of China, China; [5] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, Lanzhou, China

Background: Rare diseases present a global public health priority to healthcare practitioners and policy makers, causing significant morbidity and mortality and putting a great clinical, economic and social burden on the affected individuals, families and communities, especially for the low- and middle-income countries. Thus, standard effectiveness strategies and rigorous clinical practice guidelines that summarize the available evidence and provide recommendations are needed to improve the care of the millions of people worldwide who suffer from rare diseases. However, creating guidelines for rare diseases presents specific challenges and so far the quality of published guidelines on rare diseases is still uncertain.

Objectives: To assess the methodological and reporting quality of published clinical practice guidelines in rare disease.

Methods: We performed a systematic search of databases (PubMed, Embase and Orphanet), relevant guideline websites, and government health agency websites from their inception to April 2020 and selected clinical practice guidelines related to rare disease. No language restriction was applied. Four researchers independently evaluated the methodological quality of eligible guidelines using the AGREE-Ⅱ (Appraisal of Guidelines for Research and Evaluation Ⅱ) instrument, the degree of agreement was evaluated by intra-class correlation coefficient. Two researchers in pairs independently evaluated the reporting quality of eligible guidelines using the RIGHT (Reporting Items for practice Guidelines in Healthcare, for reporting quality) statement, disagreements were resolved through discussion or consulting a third researcher. We conducted statistical analyses using SPSS version 25.0.

Results: The results will be presented at the meeting.

Conclusions: The results will be presented at the meeting.

Patient or healthcare consumer involvement: None.

An overview of the evidence for the clinical practice guidelines in rare disease

Shi Q[1], Wang Z[2], Zhou Q[1], Zhao S[3], Xun Y[2], Chen Y[4]

[1] The First School of Clinical Medicine, Lanzhou University, China; [2] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, China; [3] School of Public Health, Lanzhou University, China; [4] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, Lanzhou, China

Background: Rare diseases represent a global public health priority to healthcare practitioners and policy makers. Clinical practice guidelines that summarize evidence and use clear criteria to generate recommendations are necessary to support decision-making in this field. However, there are numerous barriers to evidence synthesis and generation in rare diseases due to the lack of high certainty evidence. Strategies may include the use of indirect evidence, use of systematic observation forms, or the use of registry data, etc. Thus, understanding the evidence status used in the clinical practice guidelines in rare diseases is essential for follow-up research.

Objectives: To investigate the evidence overview of the clinical practice guidelines for rare diseases.

Methods: We conducted a scoping review on clinical practice guidelines for rare diseases. We performed a systematic search of databases (PubMed, Embase and Orphanet), relevant guideline websites, and government health agency websites from their inception to April 2020. We then selected clinical practice guidelines related to rare disease. No language restriction was applied. Two researchers independently screened the records and extracted data, disagreements were resolved through discussion or by referral to a third party. Data processing was performed by Stata statistical software 14 to present descriptive results.

Results: The results will be presented at the meeting.

Conclusions: The results will be presented at the meeting.

Patient or healthcare consumer involvement: None.

Analysis for the composition of the guideline-development group

Ma Y[1], Yang N[1], Wang J[2], Zhang J[2], Zhang H[1], Chen Y[1]

[1] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, China; [2] School of Public Health, Lanzhou University;Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, China

Background: In recent years, with the increase in the number of clinical practice guidelines (CPGs), their quality has received more and more attention. However, the institutions and handbooks for guideline development have not yet given unified requirements for who should participate in the development of CPGs and how to form a guideline development group (GDG) scientifically and reasonably, and there is also no clear standard for the research fields, scopes, scales and responsibilities of GDG.

Objectives: To analyse and compare the GDG formation from the existing CPGs.

Methods: We analyzed the GDG of 397 CPGs published in Chinese journals and 150 CPGs presented on Guideline International Network (randomly sampled 30 guidelines each year for a total of 150 guidelines) from 2013 to 2017, including the reporting situation, composition, scale and methodologist of GDGs.

Results: Of the 397 Chinese guidelines, 340 (85.6%) of them reported the composition of GDG consisting of secretary, writing group, consensus group, expert group, evidence group, steering group, advisory group and so on. There are many names for groups and the top three reported frequency groups are the writing group, the expert group, and the development group. In the guidelines that reported GDG, the number of groups ranged from 1 to 6, and the total number of people in GDG ranged from 3 to 137. Of the 150 G-I-N guidelines, 102 (68%) of them reported the composition of GDG and the top three reported frequency groups are the expert group, the development group, and the writing group. In the G-I-N guidelines that reported GDG, the number of groups ranged from 1 to 5, and the total number of people in GDG ranged from 1 to 60. Only 7 (1.8%) Chinese guidelines and 5 (3.3%) G-I-N guidelines clearly reported the involvement and the number of methodologists in the GDG, and the number of methodologists ranged from 1 to 4. More than 90% of the Chinese and G-I-N guidelines did not report the composition of the GDG or only reported the guidelines were developed and written by a clinician. Few guidelines reported the clear responsibilities of each group in GDG.

Conclusions: At present, the Chinese and English CPGs are not standardized enough on the reporting of the GDG, whether it is about the classification of GDG, the name of GDG, or the participation of the methodologist. In order to ensure the scientific quality of the guideline development methods, more guidance and handbooks on how to establish the GDG are needed in the future.

Patient or healthcare consumer involvement: None

Assessing the certainty across a body of evidence for comparative test accuracy

Yang B[1], Mustafa R[2], Bossuyt P[1], Brozek J[3], Hultcrantz M[4], Leeflang M[1], Schünemann H[2], Langendam M[1]

[1] Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Amsterdam UMC, University of Amsterdam, The Netherlands; [2] Michael G De Groote Cochrane Canada; McMaster GRADE Centres, Department of Health Research Methods, Evidence, and Impact, McMaster University, Canada; [3] Michael G De Groote Cochrane Canada and McMaster GRADE Centres, Department of Health Research Methods, Evidence, and Impact, McMaster University, Canada; [4] Swedish Agency for Health Technology Assessment and Assessment of Social Services (SBU), Sweden

Background: Development of recommendations for healthcare related tests and strategies may require evidence on the relative accuracy of competing testing strategies. Evidence synthesis methodology and the GRADE Working Group did, so far, not provide detailed guidance for rating certainty in comparative test accuracy.

Objectives: To develop guidance on how authors of evidence syntheses and health decision-makers, including guideline developers, can rate the certainty across a body of evidence for comparative test accuracy.

Methods: Through an iterative brainstorm-discussion-feedback process within the GRADE Working Group, we extended the existing GRADE guidance for assessing certainty of evidence for test accuracy [1] to scenarios in which two or more index tests are compared.

Results: Rating the certainty of evidence for comparative test accuracy shared many concepts and ideas with the existing GRADE guidance for test accuracy. The rating in comparisons of test accuracy required additional considerations, such as the selection of appropriate comparative study designs, additional criteria for judging risk of bias, and the consequences of using comparative measures of test accuracy. Distinct approaches to rating certainty were required for comparative test accuracy studies and between-study (indirect) comparisons.

Conclusions: This guidance will help provide a transparent assessment of the certainty of comparative test accuracy evidence, and facilitate the use of such evidence in decision-making regarding healthcare tests and strategies.

Patient or healthcare consumer involvement: Patients or healthcare consumers were not involved in the design or execution of this study.

References:

1. Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Hassan Murad M, et al. GRADE guidelines: 21 part 1. Study design, risk of bias and indirectness in rating the certainty across a body of evidence for test accuracy. *Journal of Clinical Epidemiology* 2020;122:P129-144. https://doi.org/10.1016/j.jclinepi.2019.12.020

Assessing the quality of clinical practice guidelines (CPGs) development in Taiwan using the AGREE II instrument

Chen KH[1], Chen C[2], Tam KW[3], Huang TW[4], Kuo KK[5]

[1] Cochrane Taiwan, Taipei Medical University; Post-Baccalaureate Program in Nursing, College of Nursing, Taipei Medical University; Department of Nursing, Wan Fang Hospital; Evidence-Based Knowledge Translation Center, Wan Fang Hospital, Taiwan; [2] Cochrane Taiwan, Taipei Medical University; Graduate Institute of Clinical Medicine, College of Medicine, Taipei Medical University; Department of Public Health, School of Medicine, College of Medicine, Taipei Medical University; Division of P, Taiwan; [3] Cochrane Taiwan, Taipei Medical University; Division of General Surgery, Department of Surgery, Shuang Ho Hospital, Taipei Medical University; Division of General Surgery, Department of Surgery, School of Medicine, College of Medicine, Taipei Med, Taiwan; [4] Cochrane Taiwan, Taipei Medical University; School of Nursing, College of Nursing, Taipei Medical University, Taiwan; [5] Cochrane Taiwan, Taipei Medical University, Taiwan

Background: Clinical practice guidelines (CPGs) are intended to improve quality of care based on the best available evidence. A credible CPG is an essential tool for knowledge translation. A quality CPG can facilitate clinical decision making. Therefore, it is imperative to evaluate the quality of CPGs for promoting healthcare.

Objectives: To systematically review the quality of evidence-based CPGs developed in Taiwan.

Methods: 1) Systematic search: Using "Guideline" [Publication Type], "Guidelines as Topic"[MeSH] as keywords, search for Cochrane, PubMed, Embase, CINAHL, NGC, SIGN, NICE, NZGG, G-I-N Library, JBI, National Digital Library of Theses and Dissertations in Taiwan, Airiti Library (Chinese database), Government Research Bulletin, and Google Scholar, limited to Chinese or English. 2) Quality assessment: Two researchers independently evaluated each CPG using the Appraisal of Guidelines for Research & Evaluation (AGREE II) Instrument. In case of ≧ 3 points difference in score between two review authors, a third expert was invited as tiebreaker.

Results: After screening, we included 96 CPGs published between 1999 and 2018. Most CPGs were published in Traditional Chinese (n = 78, 81.3%), the others in English (n = 18, 18.7%). Three-quarters (n = 72) of CPG development was based on systematic reviews, and a quarter (n = 24) by expert consensus. Four (4.2%) CPGs adapted the GRADE methodology. Only 13 (14%) CPGs provided an updated version. For funding, a total of 35 (36.5%) CPGs received grants from government, while others were developed independently by professional societies. The statements of conflicts of interest and related documents were rarely mentioned. The standardized score of AGREE II in the six domains were 81%, 55%, 47%, 63%, 31%, and 22%, and the overall quality was 4.5 points. The domains with lower scores were: "editorial independence", "application", and "rigour of development".

Conclusions: There is room for quality improvement in Taiwan's CPGs. For the future development, the use of a standardized international grading system, such as GRADE, is essential to ensure a high quality of methodology. In addition, the tools such as AGREE II for evaluation of the new CPGs will substantially improve the developed and updated CPGs through rigorous evaluation. It is difficult to obtain the necessary or complete information of developed CPGs in Taiwan without a credible agency; this hinders the research efforts and completeness of the data search. It would be most desirable to develop a CPG integration platform in Taiwan, either governmental or NGO-based, for better co-ordinating the development and quality assurance of CPGs and the indexing of the database for CPGs. In addition, the agency can play a role in the dissemination of CPG information for healthcare quality.

Patient or healthcare consumer involvement: Not applicable.

Australian living guidelines for management and care of patients with COVID-19 infection

Elliott J[1], Tendal B[1], Leder K[1], McDonald S[1], Murano M[1], Tate R[1], Norris S[2], Green S[1], Turner T[1], . .[3]

[1] School of Population Health and Preventive Medicine, Monash University, Australia; [2] Hereco, Australia; [3] On behalf of the National COVID-19 Clinical Evidence Taskforce, Australia

Background: COVID-19 is an infectious disease caused by a severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). After identification in December 2019, COVID-19 spread internationally and was declared a Public Health Emergency of International Concern by the World Health Organization in January 2020. Research evidence on COVID-19 was initially scant, but emerged quickly as the pandemic progressed. Australian healthcare workers needed rapid, evidence-based guidance on treatment of people with COVID-19.

Objectives: To rapidly develop an evidence taskforce and living guidelines for the management and care of people with COVID-19 infection in Australia, using collaborative living systematic review and guidelines methods.

Methods: The National COVID-19 Clinical Evidence Taskforce, convened by the Australian Living Evidence Consortium, and co-ordinated by Cochrane Australia, included more than 17 health professional organizations, representing many different groups providing care to people with COVID-19. Initial question development used three strategies: an online form distributed to the membership of the Taskforce, review of existing guidance, and discussions with clinical leaders. Preliminary recommendations were based on analysis of existing national and international guidelines on the treatment of adults with COVID-19. The second stage involves updating these recommendations using data from primary studies, using the Cochrane register of COVID-19 study reports, and systematic reviews conducted by others. Titles and abstracts are independently screened and data extracted by two members of the Cochrane Australia COVID-19 Living Guidelines team in Covidence. Where research is available to address our questions, we conduct living systematic reviews as the basis of evidence summaries, and draft evidence-based recommendations following GRADE methods, to be considered and revised by our guideline panels. Multidisciplinary guideline panels include clinicians with a range of clinical expertise and from a variety of clinical settings across Australia. Guideline panels consider, refine and agree new and revised existing recommendations at weekly meetings.

Results: The first guideline recommendations were published on 3 April 2020, two weeks after the Taskforce was initially formed. The site received more than 10,000 hits in the first 48 hours after publication. New and revised recommendations made by the guideline panels are published online each week and disseminated through traditional and social media channels. Conclusion: Living systematic review and living guideline methods; combined with national and international collaboration on evidence synthesis and clinical input enabled rapid development of an evidence taskforce and living guidelines for the management and care of people with suspected or confirmed COVID-19 infection in Australia.

Patient or healthcare consumer involvement: Consumers are represented on the Guidelines Leadership Group, and their input sought throughout the guideline process.

Avoiding informal recommendations: a framework for identifying and managing guideline statements

Hajizadeh A[1], Wiercioch WW[1], Neumann I[2], Leontiadis G[1], Akl E[3], Mbuagbaw L[1], Lotfi T[1], Schünemann H[1]

[1] McMaster University, Canada; [2] Pontifical Catholic University of Chile, Chile; [3] American University of Beirut, Lebanon

Background: Guidelines developed according to international standards, provide transparency in the development of the evidence-informed recommendations that result. Even in guidelines of high caliber such as those issued by the World Health Organization's Global TB Programme (WHO-GTB), additional actionable information convolutes the clarity of recommendations when these guideline statements are erroneously published outside of standardized methods, and/or are mistakenly classified. This paper proposes a conceptual framework for the identification and management of guideline statements; informal recommendations, good practice and implementation statements.

Objectives: The prime objective of this work is to highlight the existence of informal recommendations ubiquitous in WHO-GTB guidelines and related publications, and to propose a framework for the identification and subsequent classification of informal recommendations, implementation considerations and good practice statements. The secondary objective of this work is to prompt the appropriate development of informal recommendations to be presented as they inherently are; a WHO-GTB recommendation.

Methods: We took an iterative consensus approach to devise a conceptual framework for the identification and management of informal recommendations and other guideline statements. We invited 11 experts in health research and guideline development to participate. The conceptual framework was supported by examples classified in duplicate. We selected five PICO questions (Population, Intervention/Exposure, Comparison, Outcome) at random from five different WHO-GTB guidelines to guide the application and refinement of the conceptual framework. We performed identification, extraction, and classification in duplicate and subsequent results were verified by participants.

Results: Guideline statements (informal recommendations, good practice and implementation statements) are actionable statements that differ in terms of the PICO elements they share with the formal recommendation they accompany, their link to evidence, and their eligibility for formal development. All three statements are found to be pervasive among WHO-GTB publications.

Conclusions: WHO-GTB guidelines contain recommendations that are not always sufficient to answer the PICO question from which they arose. Additional guideline statements; informal recommendations, good practice, and implementation statements, provide additional actionable guidance. These statements should be systematically identified, and appropriately managed in a guideline's development and final presentation.

Patient or healthcare consumer involvement: This work has been conducted in collaboration with stakeholders involved in policy work in both public and clinical health.

Certainty of evidence and level of recommendation for complex versus simple interventions in Chilean clinical guidelines

Seron P[1], Oliveros M[1], Fuentes R[1]

[1] Universidad de La Frontera, Chile

Background: Complex interventions are characterized by the inclusion of a set of components that interact with each other in a complex and dynamic health system, with flexible designs that adapt to the implementation context and with outcomes of interest that are also diverse, multiple and often far from a direct causal pathway. In addition to being frequently implemented, they challenge the researchers who seek to evaluate them, both because of the difficulties in minimizing the drawbacks that impact the internal validity of the studies and their generalizability or external validity. This could influence the strength of the associated recommendations, among other things, by assessing the certainty of the evidence supporting them. The Chilean Ministry of Health is updating the clinical guidelines using the GRADE approach and the recommendations include different interventions.

Objectives: To compare the certainty of evidence and the strength of recommendation of complex versus simple interventions contained in Chilean clinical guidelines designed with the GRADE approach.

Methods: We included all Chilean guidelines that use the GRADE approach in their design. Then, the interventions derived from the questions formulated by the panel were classified as complex if they included at least four of the following complexity characteristics: population-level intervention, 'proactive' treatment, flexibility in the implementation, multiple health and social outcomes, longer causal pathway, and multiple implementation setting as the mandatory characteristic. Additionally, we extracted data about the evidence certainty and strength of the recommendation from guidelines. Two blinded researchers independently performed all processes and resolved disagreement by consensus. We performed a descriptive analysis of the data comparing both types of interventions.

Results: A total of 49 guidelines were included with 299 associated intervention recommendations. From these, 32 (10.7%) were complex compared with 267 (89.3%) simple. The certainty of evidence was **high** in 4% vs 4.3% (P > 0.05), **moderate** in 8% vs 19.8% (P > 0.05), **low** in 32% vs 26.1% (P > 0.05); or **very low** in 56% vs 49.8% (P > 0.05) of the complex compared with simple interventions respectively. The strength of recommendation was **strong** in 6.3% vs 13.5% (P > 0.05), **conditional** in 65.6% vs 82.3% (P < 0,05), **weak** in 6.2% vs 0.4% (P < 0.05), or **good clinical practice** in 21.9% vs 3.7% (P < 0.001) of the complex compared with simple interventions respectively.

Conclusions: There is no difference in the certainty of the evidence between complex and simple interventions, but more simple interventions are conditionally recommended, while more complex intervention are recommended weakly or as good clinical practice. There are considerations other than the certainty of the evidence that influence the strength of the recommendations in the Chilean clinical guidelines.

Patient or healthcare consumer involvement: None

Co-creation in developing international guidelines for the management of kidney disease

Tunnicliffe D[1], Howell M[1], Cheung M[2], Palmer S[3], Strippoli G[4], Tonelli M[5], Vandvik PO[6], Tong A[1], Craig J[7]

[1] The University of Sydney, Australia; [2] Kidney Disease: Improving Global Outcomes, Belgium; [3] University of Otago, New Zealand; [4] University of Bari, Italy; [5] University of Calgary, Canada ; [6] University of Oslo, Norway; [7] Flinders University, Australia

Background: The exponential increase in scientific literature has challenged traditional models of clinical practice guideline development in nephrology. Undertaking evidence synthesis that adheres to international standards for guideline development is resource-intensive and methodologically challenging. Improved collaboration between existing organizations with appropriate expertise is required for the development of guidelines on the diagnosis, monitoring, and treatment of kidney disease.

Objectives: To describe the collaboration between Cochrane Kidney and Transplant (CKT), Standardised Outcomes in Nephrology (SONG) Initiative, Kidney Disease: Improving Global Outcomes (KDIGO), and Making Grade the Irresistible Choice (MAGIC) to develop international guidelines for the management of kidney disease.

Methods: KDIGO has engaged CKT as the evidence review team for two guideline updates (KDIGO Glomerulonephritis Guideline; KDIGO Management of Blood Pressure in Chronic Kidney Disease Guideline) and one new guideline (KDIGO Management of Diabetes in Chronic Kidney Disease). CKT with KDIGO developed protocols that mapped PICO questions, relevant to the guideline topic to existing Cochrane Reviews. We identified SONG core outcomes as critical and important outcomes in guideline development. We searched the CKT Register of Studies to update reviews with input from review authors as necessary, and undertook new systematic reviews as required. We developed 'Summary of findings' tables and presented them in MAGICapp to the guideline Work Group, to inform the development of guideline recommendations.

Results: This collaboration between partners has led to 23 Cochrane Reviews being updated, two new Cochrane Reviews, and 27 systematic reviews of which there was no Cochrane Review available. The clinical practice guidelines are close to publication, with two of the three guidelines having undergone a public comment period. The partnership has reduced duplication of effort across organizations and ensured that guidelines incorporate core outcomes that matter to consumers in the evidence review process. The use of Cochrane systematic reviews and MAGICapp has improved the transparency in translation from evidence to recommendations and will allow for the rapid updating of guidelines in the future.

Conclusions: Developing collaborations in guideline development and using existing Cochrane Reviews has improved efficiency for all organizations and developed new opportunities for guidelines to adhere to international best-practice and include patient-important outcomes, with the potential for rapid updating.

Patient or healthcare consumer involvement: The inclusion of Standardised Outcomes in Nephrology (SONG) as critical and important outcomes for guideline development ensured that patient and caregiver priorities were being addressed. Additionally, in the diabetes guideline, patient and caregivers were involved in the guideline work group.

Developing a framework for building synergies between the selection of essential medicines and their use in practice guidelines

Piggott T[1], Lavis J[2], Moja L[3], Akl E[4], Schünemann H[1]

[1] Department of Health Research Methods, Evidence, and Impact, McMaster University, Canada; [2] McMaster Health Forum, Canada; [3] World Health Organization, Switzerland; [4] American University of Beirut, Lebanon

Background: The World Health Organization (WHO) has produced the Model List of Essential Medicines (EML) since 1977. Its purpose is to prioritize medicines through identifying the most effective therapeutic options in each disease area. Recommendations by health guideline groups on medications are important in defining appropriate uses and there is potential to build synergies between the processes of developing recommendations. In some countries the national EML and guidelines are aligned and complement each other.

Objectives: To explore decision criteria in EML recommendations and identify synergies between developing essential medicine lists and developing practice guidelines.

Methods: We are conducting semi-structured, open-ended qualitative interviews with purposefully sampled key informants with experience in the development of essential medicine lists and/or practice guidelines to ask about the decision process and link between the two. We are transcribing and analysing interviews in NVivo v12 using a closed-loop iterative coding and thematic analysis technique. Themes are extracted to support the development of a decision-framework and software solution to support cohesive approach and close alignment between EML and guidelines that may have greater impact than single component in isolation.

Results: The project interviews and analysis are in the process of completion, with final results including themes to be included for the abstract presentation.

Conclusions: The decision process for essential medicines has significant similarities with decisions made by health guideline groups. Building synergies these two processes can be reinforced by building a framework, and a software tool to support a united approach. Further evaluation is warranted to better understand the utility of this approach.

Consumer involvement: We worked with guideline and EML sponsor organizations to identify experts and end-users of guidelines and EMLs to invite to participate in this study. We will work with an advisory group through the process to involve consumers of guidelines and EMLs in the analysis and development of a framework and software solution.

Development of the European guidelines on breast cancer screening and diagnosis

Saz-Parkinson Z[1], Parmelli E[1], Janusch-Roi A[1]

[1] European Commission; Joint Research Centre, Italy

Background: The European guidelines on breast cancer screening and diagnosis are evidence-based guidelines developed within the European Commission Initiative on Breast Cancer (ECIBC). They have been developed by a multidisciplinary group (GDG) of experts in the field, as well as patients, using the GRADE approach and its evidence to decision (EtD) frameworks.

Objectives: To present the results of the development process of the European Breast Cancer Guidelines.

Methods: The European Commission selected, via an open call, a multidisciplinary GDG who voluntarily participate in the Guideline development process. The European Commission's Joint Research Centre (JRC) is responsible for the scientific co-ordination, overarching planning and budgeting of the Guidelines, ensuring conflict of interest management and transparently reporting the guideline process. Systematic reviews externally conducted by Cochrane Iberoamerica support each recommendation developed using GRADE's EtD frameworks, which are updated as new evidence becomes available according to our updating strategy (https://healthcare-quality.jrc.ec.europa.eu/discover-ecibc/methodologies/guidelines-updating).

Results: During the 4.5 years since start of ECIBC, 16 face-to-face meetings have taken place (39 total days) with approximately 27 GDG panel experts. Recommendations have gradually been published during this period, as they were finalized, on a dedicated webpage. The following recommendations have been developed: 21 recommendations on breast cancer screening, 19 on breast cancer diagnosis, 31 relating to communication issues around screening and diagnosis and 2 on training of professionals involved in breast cancer screening and diagnosis. Additionally, we have also developed three good practice statements on training and communication issues and 13 indicators have been developed to monitor breast cancer screening. In 2019, two recommendations were updated and in 2020, six were updated and a new one was developed as a result of the updating process. With regard to the strength of the recommendations in the European Breast Cancer Guidelines, 61 are conditional and 12 are strong recommendations (requiring at least 80% agreement among GDG members). There is low or very low certainty evidence for 48 of the recommendations and only high certainty evidence for two.

Conclusions: The development of the European Breast Cancer Guidelines has been a rapid learning process for all those involved, using the experience gained to improve and streamline processes in order to be able to achieve a development rate of approximately 1.5 recommendations per month. All lessons learned are helping design the new European Commission Initiative on Colorectal Cancer. Finally, transparent reporting of EtDs is allowing countries to rapidly adapt or adopt these recommendations to their specific healthcare contexts (Bulgaria, Italy, Spain, Tunisia, China and Chile, among others)

Patient or healthcare consumer involvement: Three patient members of GDG are fully involved in development and updating process.

Dietary recommendations for the obese and overweight: a systematic survey of global guidelines

Wang Q[1], Zhang Q[2], Pan B[3], Tian J[4], Ge L[5]

[1] School of Public Health, Lanzhou University, China; [2] School of Nursing, Lanzhou University, China; [3] Gansu Provincial Hospital, Lanzhou, China; [4] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, China; [5] Evidence-Based Social Science Research Center, School of Public Health, Lanzhou University, China

Background: WHO defined overweight and obesity as abnormal or excessive fat accumulation that presents a huge risk to health. Obesity and overweight can lead to a number of chronic diseases such as hypertension, type Ⅱ diabetes, sleep apnea, cardiovascular diseases, and even cancers. With the wildly increasing trend of obesity and overweight, it's urgent to offer reliable instructions for clinicians when making decisions. That's where guidelines are required. Nowadays, large quantities of guidelines have been published to offer professional guidance for concerning operators. However, recommendations on diet from different guidelines also differ from each other and the quality of these guidelines varies greatly, some of which are too low to be trusted.

Objectives: To systematically appraise the quality of obesity/overweight guidelines which make recommendations on diet and assess the consistency of their dietary recommendations.

Methods: We searched PubMed, Web of Science, Embase as well as main guideline websites including Guidelines International Network (GIN), National Institute for Health and Clinical Excellence (NICE), National Guideline Clearinghouse (NGC), Scottish Intercollegiate Guidelines Network (SIGN) for obesity/overweight guidelines and dietary recommendations. Four researchers independently assessed their methodological and reporting quality using the AGREE II instrument and RIGHT checklist and extracted dietary recommendations on managing obesity/overweight. We will summarize the general characteristics, evidence source information and details of the guideline recommendations using descriptive statistics. We calculate both the mean score and the standard deviation (SD) of all guidelines for each domain of AGREE II (Appraisal of Guidelines for Research and Evaluation) instrument. We also summarize the number of reported items and the reporting rate of each domain of the RIGHT (Essential Reporting Items for Practice Guidelines in Healthcare) checklist for the guidelines. The statistical analysis will be conducted using Excel 2013 (Microsoft Corp, Redmond, Washington, USA, www.microsoft.com).

Results: This study is ongoing and results will be presented at the Cochrane Colloquium as available.

Conclusions: This study is ongoing and results will be presented at the Cochrane Colloquium as available.

Patient or healthcare consumer involvement: No patients or healthcare consumers are involved in this study.

Drafting evidence-based recommendations in 10 days in the context of the COVID-19 pandemic: methods and pitfalls

Franco JVA[1], Escobar Liquitay CM[1], Peuchot V[2], Boietti B[2], Sanchez Thomas D[2], Torres Gomez F[2], Peroni L[2], Baez GN[3], Pisula P[3], Pollán J[2]

[1] Centro Cochrane, Instituto Universitario Hospital Italiano, Argentina; [2] Servicio de Clínica Médica, Hospital Italiano de Buenos Aires, Argentina; [3] Carrera de Medicina, Instituto Universitario Hospital Italiano, Argentina

Introduction: The arrival of the COVID-19 pandemic to our country introduced uncertainties regarding treatment options. Current recommendations were based on supportive care. The use of drugs capable of inhibiting viral replication such as hydroxychloroquine or lopinavir/ritonavir was suggested by other countries that faced the initial outbreaks of the pandemic.

Objectives: To generate rapid recommendations on available pharmacological interventions for the treatment COVID-19 and to identify the opportunities and pitfalls of this process in the context of ongoing pandemic disease.

Methods: We generated pragmatic searches in multiple databases searching direct and indirect evidence on the pharmacological treatments for COVID-19 (we excluded case reports, editorials and reviews). We created 'Summary of findings' tables and presented them to a multidisciplinary panel, including patients, to develop recommendations using explicit criteria under the Evidence to Decision Framework. The decision and the process were documented using a qualitative approach (notes and transcriptions of the working sessions).

Results: We screened 3191 references (21 March 2020) and included one randomized controlled trial for ritonavir/lopinavir and one observational study for hydroxychloroquine. Based on low to very-low quality of evidence these interventions caused little to no difference in the pre-defined critical outcomes identified by our community. Ten days after starting the project we had a virtual meeting with the panel and a decision was made to recommend the use of lopinavir/ritonavir for critically ill adults (weak recommendation) and if the patient had a contraindication to lopinavir/ritonavir, then use hydroxychloroquine (weak recommendation). Our main difficulties were: the emerging available evidence from primary studies, usually in forms of preprints, or available 'Summary of findings' tables from rapid reviews or health technology assessments (many of them with different outcomes, data or interpretation). Furthermore, scientific societies and local and foreign governments issued simultaneously conflicting recommendations, which caused some reluctance to the development of the first evidence-based recommendations of our institution.

Conclusions: We were able to generate rapid recommendations for COVID-19 in a short time span (10 days), however, we faced emerging methodological challenges in the evidence synthesis process and contextual challenges when formulating recommendations. This experience could help in the development of methodological guidance for recommendations in the context of an emergency such as pandemic disease.

Patient or healthcare consumer involvement: We conducted an open online poll to people in the community and healthcare professionals (> 400 responses) to identify critical outcomes for 'Summary of findings' tables. A patient advocate was incorporated into the decision panel.

Establishing thresholds for important benefits in screening guidelines

Helsingen LM[1], Zeng L[2], Siemieniuk RAC[3], Lytvyn L[3], Vandvik PO[4], Agoritsas T[5], Bretthauer M[1], Guyatt G[3]

[1] Oslo University Hospital and University of Oslo, Norway; [2] McMaster University and Sichuan University, Canada and China; [3] McMaster University, Canada; [4] University of Oslo and Lovisenberg Diakonale Hospital, Norway; [5] University Hospitals of Geneva and Faculty of Medicine of Geneva, and McMaster University, Switzerland and Canada

Background: International standards for clinical practice guidelines require explicit statements regarding how values and preferences influence recommendations. However, no cancer screening guideline has addressed the key question of what magnitude of benefit people would require to undergo screening, given its harms and burdens.

Objective: This article describes the development of a new method for guideline developers to define a threshold for the magnitude of benefit needed to undergo the burdens and harms of screening.

Summary of methods: The new method was developed and applied in the context of a recent BMJ Rapid Recommendation clinical practice guideline for colorectal cancer (CRC) screening. First, we presented the guideline panel with harms and burdens (derived from a systematic review) associated with the CRC screening tests under consideration. Second, each panel member completed surveys documenting their views of the expected benefits on CRC incidence and mortality that people would require to accept the harms and burdens of screening. Third, the panel discussed the results of the surveys and agreed on thresholds for benefits at which the majority of people would choose screening. During these three steps, the panel had no access to the actual benefits of the screening tests. In step four, the panel was presented with screening test benefits derived from a systematic review of clinical trials and microsimulation modeling. The thresholds derived through steps one to three were applied to these benefits, and directly informed the panel's recommendations.

Results: The panel inferred that at least half of people in the target population would require a colorectal cancer mortality reduction of 5 per 1000 over a 15-year period to undertake fecal immunochemical test (FIT) screening, and a reduction of 10 per 1000 for sigmoidoscopy or colonoscopy. They clarified that a difference in colorectal cancer mortality reduction of 10 or more per 1000 would prompt them to recommend colonoscopy or sigmoidoscopy over FIT, and a difference of 5 or more would prompt recommending FIT every year over FIT every two years and colonoscopy over sigmoidoscopy.

Conclusion: We present the development and application of a new, four-step method that enables incorporation of explicit and transparent judgments of values and preferences in a screening guideline. Guideline panels should establish their view regarding the magnitude of required benefit, given burdens and harms, before they review screening benefits, and make their recommendations accordingly. Making informed screening decisions requires transparency in values and preferences judgments that our new method greatly facilitates.

Patient or healthcare consumer involvement: Three people with colorectal cancer screening experience took part in the guideline development process, including defining thresholds of the required benefit.

Evidence and recommendations: a systematic analysis based on 77 COVID-19 guidelines and consensus statements

Luo X[1], Liu Y[2], Ren M[2], Wang L[2], Lv M[2], Zhang X[3], Chen Y[4], Wang M[5]

[1] School of Public Health of Lanzhou University, Lanzhou; Cochrane China Network for GRADE and Guideline Working Group, Lanzhou, China; [2] School of Public Health of Lanzhou University, Lanzhou, China; [3] The First School of Clinical Medicine, Lanzhou University, Lanzhou, China; [4] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou; Cochrane China Network for GRADE and Guideline Working Group, Lanzhou, China; [5] Department of Radiology, the First Hospital of Lanzhou University, Lanzhou, China

Background: Since the outbreak of COVID-19, a large number of guidelines and consensus statements have been developed, but the quality of the cited evidence and the recommendations are unclear.

Objectives: To systematically analyze the cited evidence of COVID-19 guidelines and consensus statements, compare the recommendations of existing evidence to provide guidance for clinicians.

Methods: We searched PubMed, Embase, CBM, CNKI and Wanfang Data from 1 December 2019 to 1 April 2020. In addition, we searched some official websites including US CDC, European CDC, National Health Commission of the People's Republic of China and some pre-print servers. Two trained researchers conducted the study selection and data extraction independently, then the third researcher checked the results. The information we extracted includes baseline characteristics, guideline and consensuses statements development information and recommendations. Finally, we systematically analyzed the extracted information.

Results: We included 77 records (24 guidelines and 53 consensus statements) from China (57, 74.0%), the US (10, 13.0%), WHO (6, 7.8%) and Europe (4, 5.2%). A total of 990 references were cited, with an average of 13 (0 to 90) references. Twenty-one (27.3%) guidelines and consensus statements did not cite references. Guidelines cited more systematic reviews (odds ratio (OR) 6.71, 95% confidence interval (CI) 3.26 to 13.83, P = 0.00), RCTs (OR 2.11, 95% CI 1.03 to 4.32, P = 0.04), and SARS/MERS studies (OR 3.31, 95% CI 2.39 to 4.60, P = 0.00) than consensus statements. More systematic reviews (OR 3.09, 95% CI 1.59 to 6.03, P = 0.00) and SARS/MERS studies (OR 2.63, 95% CI 1.78 to 3.88, P = 0.00) were cited in international guidelines and consensus statements. Guidelines and consensus statements published by associations cited more systematic reviews (OR 0.45, 95% CI 0.23 to 0.88, P = 0.02), RCTs (OR 0.42, 95% CI 0.19 to 0.92, P = 0.03) and SARS/MERS studies (OR 0.34, 95% CI 0.24 to 0.48, P = 0.00) than which published by non-associations. Thirty-one guidelines and consensus statements (40.26%) include the contains of diagnosis, of which, nucleic acid detection was the most recommended. Thirty-one guidelines and consensus statements (40.26%) include the contains of treatment, of which, 19 guidelines and consensus statements (24.68%) covered general treatment, 19 guidelines and consensus statements (24.68%) refereed to antiviral therapy, and 18 guidelines and consensus statements (23.38%) with oxygen therapy. Finally, main recommendations were extracted and summarized, and 23 recommendations were formed, including 3 for diagnosis, 12 for treatment, 3 for special population, and 5 for other aspects.

Conclusions: Current COVID-19 guidelines and consensus statements are based on less evidence and recommendations from different guidelines and consensus statements are inconsistent. More original research should be conducted, and the quality of guidelines and consensus statements should be strengthened.

Patient or healthcare consumer involvement: None.

How do cancer screening guidelines trade off benefits versus harms and burdens of screening: a systematic survey

Zeng L[1], Helsingen LM[2], Nampo FK[3], Wang Y[4], Siemieniuk RA[4], Yao L[4], Bretthauer M[4], Guyatt GH[4]

[1] West China Second University Hospital, Sichuan University; Department of Health Research Methods, Evidence and Impact, McMaster University, China; [2] Clinical Effectiveness Research Group, Institute of Health and Society, University of Oslo and Department of Transplantation Medicine, Oslo University Hospital, Norway; [3] Latin-American Institute of Life and Nature Sciences; Evidence-Based Public Health Research Group, Federal University of Latin-American Integration, Brazil; [4] Department of Health Research Methods, Evidence and Impact, McMaster University, Canada

Background: Cancer screening guidelines differ in their recommendations for or against screening. To be able to provide explicit recommendations, guidelines need to specify thresholds for the magnitude of benefits of screening, given its harms and burdens. Cancer screening guidelines have a potential advantage over guidelines in other areas of medicine because they often have one or two key benefit outcomes: reduction in cancer mortality or cancer incidence. Therefore, the central question guideline panels face may be framed as: Given the harms and burden of screening, what magnitude of its key benefits (effect on cancer incidence and/or mortality) would people require to undergo screening? In other words: what is the threshold above which people would undergo screening and below which they would not? By establishing such a threshold, a panel makes transparent, through a quantitative trade-off between benefits and harms or burdens, their assessment of the values and preferences of the target population.

Objectives: We evaluated how current cancer screening guidelines address the relative importance of benefits versus harms and burdens of screening.

Methods: We searched the Guidelines International Network, International Guideline Library, ECRI Institute, and MEDLINE. Two pairs of review authors independently performed guideline selection and data abstraction. We included all cancer screening guidelines published in English between January 2014 and April 2019. We established specific criteria to evaluate whether and how cancer screening guidelines defined a threshold for a key benefit outcome in the trade-off between benefits versus harms and burdens, and in cost-effectiveness evaluation. We also evaluated whether the guidelines qualitatively commented on the trade-off between benefits versus harms and burdens (e.g. a statement of benefits over harms and burdens, or a statement of the magnitude of net effect) or commented on the target population's values and preferences regarding the trade-off between benefits versus harms and burdens.

Results: Of 68 eligible guidelines, 25 included a statement regarding the trade-off between screening benefits versus harms and burdens (14 guidelines), or a statement of direction of the net effect (defined as benefits minus harms or burdens) (13 guidelines). None of these 25 guidelines defined how large a screening benefit should be to recommend screening, given its harms and burdens. Eleven guidelines performed an economic evaluation of screening. Of these, six identified a key benefit outcome; two specified a cost-effectiveness threshold for recommending a screening option. Eight guidelines commented on people's values and preferences regarding the trade-off between benefits versus harms and burdens.

Conclusions: Current cancer screening guidelines fail to specify the values and preferences underlying their recommendations. No guidelines provide a threshold at which they believe the benefits of screening outweigh its harms and burdens.

Patient or healthcare consumer involvement: Not involved

Additional files: Tables

Implementation of the transparent and rigorous evidence-based methodology of clinical practice guideline development in the Czech Republic

Klugarova J[1], Vrbova T[2], Pokorna A[1], Licenik R[3], Dusek L[4], Klugar M[5]

[1] Masaryk University, Faculty of Medicine, Czech National Centre for Evidence-Based Healthcare and Knowledge Translation (Cochrane Czech Rep./Czech JBI Centre of Excellence/MUNI GRADE Centre); Institute of Health Information and Statistics of the Czech Rep., Czech Republic; [2] Czech Health Research Council, Czech Republic; [3] Masaryk University, Faculty of Medicine, Czech National Centre for Evidence-Based Healthcare and Knowledge Translation (Cochrane Czech Rep./Czech JBI Centre of Excellence/MUNI GRADE Centre); Czech Health Research Council, Czech Republic; [4] Institute of Health Information and Statistics of the Czech Republic, Czech Republic; [5] Masaryk University, Faculty of Medicine, Czech National Centre for Evidence-Based Healthcare and Knowledge Translation (Cochrane Czech Rep/Czech JBI Centre of Excellence/MUNI GRADE Centre); Institute of Health Information and Statistics of the Czech Rep., Czech Republic

Background: Currently, in the Czech Republic, there is no such institution as the "National Centre for Clinical Practice Guidelines", which would systematically collaborate with all relevant stakeholders in the Czech Health System. There is high heterogeneity in level and also in a type of indicated care among healthcare facilities. In 2018, we established the Czech National Centre for Evidence-Based Healthcare and Knowledge Translation (CEBHC-KT) as the Czech methodological hub and umbrella for the Cochrane Czech Republic, Masaryk University Grade Centre and the Czech Republic Centre for Evidence-Based Healthcare: The Joanna Briggs Institute Centre of Excellence. Within the national guidelines project led by the Czech Health Research Council together with Ministry of Health of the Czech Republic and Institute of Health Information and Statistics of the Czech Republic, we developed the Czech National methodology of the Clinical Practice Guideline development based on the GRADE approach.

Objectives: To implement the transparent and rigorous evidence-based methodology of clinical practice guideline development into the Czech Health System.

Methods: The Czech National methodology of the Clinical Practice Guideline (CPGs) development is being implemented in the Czech Republic using the Joanna Briggs Institute model for implementation based on the clinical audit, GriP (Getting Research into Practice) strategy, follow-up audit and feedback. This implementation project is part of European ERASMUS+ project "Strategic Partnership in Innovation and Development of Evidence-Based Healthcare".

Results: Our baseline audit showed that in March 2017, there were 123 professional healthcare organizations which developed 1909 CPGs. However, the majority of these CPGs were "expert opinion" or "consensual" based guidelines, most of which lacked evidence-based medicine principles and methods. There were no CPGs based on the GRADE approach. The main obstacles to implementation of the Czech national CPG methodology were identified as: lack of knowledge of CPG development among all types of healthcare professionals (including the heads of the professional medical and allied healthcare organizations who are responsible for CPG development), and lack of motivation for high-quality CPG development and use. Therefore, we organized workshops focused on CPG development, including systematic review and GRADE methodology combined with individual co-operation with teams developing CPGs. Currently, we have developed 12 evidence-based trustworthy CPGs and we are working on a further 28.

Conclusions: Implementation is a crucial part of the evidence-based health care approach. However, it is also essential to choose a suitable implementation model and especially to engage key stakeholders regarding the topic of the change.

Patient or healthcare consumer involvement: The chair of the Patients Board is a member of the Guarantee Committee of the CPG Project, approving topics of new CPGs and also approving completed / final version of CPGs in the Czech Republic.

Methodology for developing the Public Health Agency of Canada sexually transmitted and blood-borne infections guidelines

Shanmugasegaram S[1], Gale-Rowe M[1]

[1] Public Health Agency of Canada, Canada

Background: The Public Health Agency of Canada's (PHAC) Centre for Communicable Diseases and Infection Control (CCDIC) develops the Sexually Transmitted and Blood-Borne Infections (STBBI) Guidelines (formerly the Canadian Guidelines on Sexually Transmitted Infections) in consultation with the National Advisory Committee on STBBI (NAC-STBBI). CCDIC is implementing a new methodology to balance the value of expert opinion with the increasing need to use a rigorous, systematic and transparent approach to formulate trustworthy guidance.

Objective: To provide an overview of the methodology for developing the PHAC STBBI Guidelines.

Methods: A review identified publications on best practice standards (e.g. GRADE approach) in guideline development and the methods manuals of major guideline developers. This informed the draft CCDIC methods manual, which underwent internal review before being implemented.

Results: The methodology is as follows: CCDIC conducts a topic selection and prioritization exercise, which includes completing an assessment tool to help the NAC-STBBI rank a list of topics. For each prioritized topic, CCDIC forms a Sub Working Group to engage experts and conducts a scoping exercise before an evidence review protocol is prepared. The GRADE approach is used, as appropriate, when conducting evidence reviews and developing recommendations. After NAC-STBBI voting, the recommendations are published in a statement.

Conclusion: By following this methodology, PHAC will produce trustworthy evidence-based STBBI recommendations. Updates to the manual will reflect lessons learned and, as appropriate, new developments in the guideline methodology field.

Patient or healthcare consumer involvement: CCDIC plans to identify and engage relevant stakeholders during various stages of the guideline development process.

Methods for the evaluation of the implementation of a rapid-recommendation on COVID-19 treatment

Peuchot V[1], Boietti B[1], Franco J[2], Escobar Liquitay C[3], Sanchez Thomas D[1], Torres Gomez F[1], Peroni L[1], Baez G[4], Pisula P[4], Pollan J[1]

[1] Servicio de Clínica Médica, Hospital Italiano de Buenos Aires, Argentina; [2] Departamento de Investigación, Instituto Universitario Hospital Italiano Servicio de Medicina Familiar y Comunitaria – Hospital Italiano de Buenos Aires, Argentina; [3] Biblioteca Central, Instituto Universitario Hospital Italiano, Argentina; [4] Alumno, Instituto Universitario Hospital Italiano, Argentina

Background: The arrival of the COVID-19 pandemic to Argentina introduced uncertainties regarding treatment options. Initial recommendations were based on supportive care. The use of drugs capable of inhibiting viral replication such as hydroxychloroquine or lopinavir/ritonavir was suggested by other countries that faced the first cases in the pandemic. We developed rapid recommendations for the treatment of COVID-19 following the GRADE framework.

Objectives: To evaluate compliance rates to quality of care indicators along with implementation of locally developed recommendations for the treatment of COVID-19.

Methods: We created a prospective registry of all COVID-19 cases admitted to our institution, including data of the critical outcomes that were considered when drafting these recommendations and indicators of the adequateness of the implementation of said recommendations. Considering that they were "weak" recommendations, we anticipate substantial variability. The indicators alongside the outcomes will feed to a dashboard that will be presented to the decision panel of the recommendation and the Crisis Committee of our institution in the following weeks. In addition, there will be a brief survey of the intended medical users of the recommendations to assess their views on their applicability, including barriers and enablers.

Results: We have drafted and pilot-tested the data collection forms and database. We expect to analyse the opportunities and challenges when implementing rapid recommendations in the context of a pandemic disease. We will present the results at the Colloquium.

Consumer involvement: We conducted an open online poll to people in the community and healthcare professionals (> 400 responses) to identify critical outcomes for SoF tables. A patient advocate was incorporated into the decision panel.

Methods of guideline adaptation: a review of guideline development guidance

El-Khoury R[1], Noureldine M[2], Abboud W[3], Khabsa J[1], Schünemann H[4], Akl E[5]

[1] Clinical Research Institute, American University of Beirut Medical Center, Beirut, Lebanon; [2] Faculty of Medicine, Lebanese American University, Byblos, Lebanon; [3] Department of Internal Medicine, American University of Beirut, Beirut, Lebanon; [4] Department of Health Research Methods, Evidence, and Impact (HEI), McMaster University, Ontario; Department of Medicine, McMaster University, Ontario, Canada; [5] Clinical Research Institute, American University of Beirut Medical Center, Department of Internal Medicine, American University of Beirut, Beirut; Department of Health Research Methods, Evidence, and Impact (HEI), McMaster University, Lebanon

Background: Guidelines are the cornerstone of high-quality evidence-based practice. Guidelines could be 1) developed from scratch, i.e. de novo; 2) adopted as is; or 3) adapted to the local setting taking into consideration contextual factors. Adapting guidelines accounts for contextual factors, while cutting on the required resources.

Objectives: To systematically review guideline adaptation processes as described in guidance documents.

Methods: We included any publicly available documents, and peer reviewed articles by guideline producing organizations mentioning adaptation. We abstracted in duplicate and independently information on the characteristics of the organizations and of the adaptation process. We synthesized data narratively.

Preliminary Results: Out of 137 guideline producing organizations considered, 36 provided guidance on guideline adaptation (total of 48 documents). The majority of the documents were in English (52%) and were produced by national organizations (67%). 19% of the documents were completely dedicated to adaptation and 50% referred to the ADAPTE-toolkit for adaptation. 27% of the documents addressed the consideration of equity and 29% assessed the baseline risk of the outcomes. Only 10% of the guidelines assessed the indirectness of the evidence.

Conclusions: The majority of the guideline adaption guidance documents do not assess the indirectness of the evidence, the baseline risk for the outcome or the equity while adapting a guidance document.

Patient or healthcare consumer involvement: Appropriate adaptation of guidelines should improve the quality of care, reduce variability in medical practices, and eventually improve patients' outcomes.

Moving from evidence to decisions when making recommendations in clinical practice guidelines guidance documents

Meneses-Echávez J[1], Bidonde J[2], Flottorp S[3], Sanabria AJ[4], Yepes-Nuñez JJ[5], Poklepović Peričić T[6], Puljak L[7], Bala MB[8], Storman D[9], Swierz M[8], Zając J[10], Alonso-Coello P[4]

[1] Norwegian Institute of Public Health; Facultad de Cultura Física, Deporte y Recreación, Universidad Santo Tomás, Bogotá, Colombia, Norway; [2] Norwegian Institute of Public Health, Oslo, Norway; School of Rehabilitation Science, College of Medicine, University of Saskatchewan, Canada, Norway; [3] Norwegian Institute of Public Health, Norway; [4] Iberoamerican Cochrane Centre; CIBERESP-IIB Sant Pau, Spain; [5] Universidad de los Andes, School of Medicine, Colombia; [6] Cochrane Croatia; University of Split School of Medicine, Croatia; [7] Center for Evidence-Based Medicine and Health Care; Catholic University of Croatia, Zagreb, Croatia; [8] Chair of Epidemiology and Preventive Medicine, Department of Hygiene and Dietetics, Systematic Reviews Unit; Jagiellonian University Medical College, Poland; [9] Chair of Epidemiology and Preventive Medicine, Department of Hygiene and Dietetics, Systematic Reviews Unit, Department of Adult Psychiatry; University Hospital Jagiellonian University Medical College, Poland; [10] Chair of Epidemiology and Preventive Medicine, Department of Hygiene and Dietetics, Systematic Reviews Unit, Jagiellonian University Medical College, Poland

Background: The process of moving from evidence to decisions (EtD) is an important step in guideline development. Ideally, a framework should guide this process. Various multicriteria EtD frameworks have been launched but their use has not yet been evaluated systematically.

Objectives: To identify and describe the processes and frameworks suggested for formulating recommendations in guidance documents for clinical practice guidelines development.

Methods: We searched the G-I-N library, MEDLINE, the Cochrane Methodology Register, and gray literature sources to identify guidance documents published since 2000, produced by worldwide institutions responsible for guideline development. We excluded those documents that only reported methodology for adapting, endorsing, or updating guidelines. Pairs of independent researchers selected and extracted data using a piloted case report form.

Results: This is a preliminary analysis of 37 guidance documents (mean publication year 2014, range 2004 to 2020). Most of the documents (22; 59%) were published by scientific societies, from the United States (14; 38%). Eleven guidance documents (30%) described a structured process for outcomes rating. Most of them (32; 86%) indicated a structured process for assessing the quality of the evidence, for which the GRADE approach was the most common (24; 65%), followed by SIGN (3; 8%). Only one document did not provide a system for grading the strength of recommendations, whereas 25 documents (68%) indicated a structured process for formulating recommendations. More than half (22; 59%) of the documents suggested the use of the GRADE-EtD frameworks for moving from evidence to recommendations; other frameworks included SIGN (2; 5%), Oxford (2; 5%), and NICE (1; 3%).

Conclusions: Considering guidance documents for producing clinical practice guidelines, GRADE was the most widely used approach for assessing the quality of the evidence, and the GRADE-EtD frameworks were the most common ones suggested by organizations devoted to guidelines development for the process of moving from evidence to recommendations. Not all organizations suggest structured frameworks for the EtD process.

Patient or healthcare consumer involvement: Even though almost all the documents (34; 92%) provided guidance about the composition of the guideline panel, only 23 of them (62%) suggested the involvement of patients or consumers representatives, and six (16%) encouraged the involvement of members of the public. A wider guidance about the involvement of patients or consumers' representatives and members of the public in clinical practice guidelines is highly encouraged.

Pharmacotherapy recommendations for obesity: a cross-sectional survey

Wang Y[1], Wang J[2], Pan B[2], Ge L[1], Chen Y[1]

[1] School of Public Health, Lanzhou University, Lanzhou, China; [2] Gansu Provincial Hospital, Lanzhou, China

Background: Obesity, a modern worldwide epidemic, not only has a great negative impact on people's health and quality of life but also causes a serious economic burden. The treatment of obesity involves several facets, including lifestyle changes, bariatric surgery, and pharmacotherapy. Due to the limitations of other treatments, pharmacotherapy is necessary as an adjuvant in the treatment of obesity. Five anti-obesity drugs have now been approved by US FDA: orlistat, phentermine plus topiramate, lorcaserin, naltrexone plus bupropion, and liraglutide. According to our knowledge, several clinical practice guidelines (CPGs) about pharmacotherapy for obesity have been published, yet the quality of the CPGs is mixed. So far, there are few studies systematically assessing the consistency of recommendations, methodological quality, and reporting quality about CPGs in pharmacotherapy for obesity. **Objective:** To systematically review the existing guidelines on pharmacotherapy for obesity and assess the consistency of their recommendations, methodological quality and reporting quality.

Methods: We searched the clinical practice guidelines (CPGs) in PubMed, Embase, Web of Science, CBM and four main guideline databases about pharmacotherapy for obesity. All searches were performed on 5 August 2019. Inclusion criteria are as follows:

1) published guidelines on overweight, obesity or weight loss;
2) CPGs reported recommendations on pharmacotherapy for obesity; and
3) The language of publication is English.

According to the eligibility criteria, two review authors screened titles and abstracts of all the retrieved bibliographic records independently, then we further evaluated the text in full for any potentially eligible studies. Due to the mature research in this field and the number of guidelines, only the guidelines are included, not consensus statements, and recommendations. We will use the AGREE II instrument to appraise methodological quality included in the guidelines, which comprises 23 key items organized within six domains and the RIGHT checklist will be used to evaluate the quality of reports included in the guidelines, which consists of 22 items by four review authors. Then, we will use descriptive statistics to summarize the general characteristics, evidence source information, and details of the recommendations of the guidelines. We will use the interclass correlation coefficient (ICC) with 95% CIs to measure the agreement among the four review authors. We will rank the ICC scores into five groups: minor (0.01 to 0.20), fair (0.21 to 0.40), moderate (0.41 to 0.60), substantial (0.61 to 0.80), very good (0.81 to 1.00). We will conduct all the analyses using SPSS version19.0.

Results: This study is ongoing and results will be presented at the Cochrane Colloquium as available.

Conclusions: This study is ongoing and results will be presented at Cochrane Colloquium as available.

Patient or healthcare consumer involvement: Not applicable.

Physical activity recommendations for obesity: a cross-sectional survey

Wang Y[1], Wang J[2], Pan B[2], Ge L[1], Chen Y[1]

[1] School of Public Health, Lanzhou University, China; [2] Gansu Provincial Hospital, Lanzhou, China

Background: Obesity has become the fifth leading cause of human death worldwide. It can cause numerous negative health effects, such as increased risk of diabetes, cardiovascular disease, hyperlipidemia, arteriosclerosis, and cancer. The World Health Organization (WHO) has defined overweight and obesity as a chronic disease. Physical activity to achieve the purpose of weight loss by consuming excess fat has always been considered one of the most important and effective ways for obesity prevention and weight control. According to our knowledge, there are so many clinical practice guidelines (CPGs) on physical activity for obesity, however, there are few studies systematically assessing the consistency of recommendations, methodological quality, and reporting quality about CPGs in physical activity for obesity.

Objectives: To systematically review the existing guidelines on physical activity for obesity and assess the consistency of their recommendations, methodological quality, and reporting quality.

Methods: We searched for CPGs in PubMed, Embase, Web of Science, CBM and four main guideline databases to identify existing physical activity for obesity guidelines. All searches were performed on 5 August 2019. The included research needs to meet the following criteria: 1) published guidelines on overweight, obesity or weight loss, 2) CPGs reported recommendations on physical activity for obesity, 3) the language of publication is English. We conducted pilot-literature selection to ensure high inter-rater reliability of reviewers. Firstly, and according to the established eligibility criteria, two review authors independently screened the titles and abstracts of the retrieved records. For records that met the criteria for inclusion, we screened the full-text. Due to the mature research in this field and the number of guidelines, only the guidelines are included, not consensus statements, and recommendations. We will use the Appraisal of Guidelines for Research and Evaluation, 2nd edition (AGREE II) instrument to appraise the methodological quality included in the guidelines, which includes 23 items in six domains. Then, we will use the Reporting Items for Practice Guidelines in Healthcare (RIGHT) tool to evaluate the quality of reports included in the guidelines, which consist of 22 items in 7 domains. We will also use Excel 2013 to sort out the scores in each domain of AGREE II, expressed as main scores and standard deviation (SD). We will use SPSS 21.0 to calculate the Interclass Correlation Coefficient (ICC) and 95% Confidence Interval (95% CI) among reviewers. The ICC scores will be ranked into five groups: minor (0.01 to 0.20), fair (0.21 to 0.40), moderate (0.41 to 0.60), substantial (0.61 to 0.80), very good (0.81 to 1.00). We will conduct all the analyses in SPSS version 21.0.

Results: This study is ongoing and results will be presented at the Cochrane Colloquium as available.

Conclusions: This study is ongoing and results will be presented at the Cochrane Colloquium as available.

Patient or healthcare consumer involvement: No applicable.

Prioritizing topics for and adopting/adapting WHO guidance: a critical interpretive synthesis

Piggott T[1], Alvarez E[1], Lavis J[2]

[1] Department of Health Research Methods, Evidence, and Impact, McMaster University, Canada; [2] McMaster Health Forum, Canada

Background: Health guidance development is a key normative function of the World Health Organization (WHO). Within the WHO system, while guidance is primarily developed by headquarters or regional offices, it could be developed and shared for adaptation/adoption at many levels (global, regional and country office). Priorities too could be generated at different levels.

Objectives: To identify the elements of a framework that would assist decision-makers in prioritizing topics where guidance is needed and in adopting/adapting guidance to their own context; paying particular attention to differences by level of decision-making (global/regional/national/sub-national) and type of health guidance (clinical/public health/health system).

Methods: We conducted a critical interpretive synthesis (CIS) by using formal literature searches and an inductive process to interpret meaning from the evidence reviewed. We connected with key content experts for input at multiple stages throughout the CIS process. We conformed to the PRISMA guidelines and registered this study in PROSPERO. We searched 13 academic and grey literature databases. We conducted title and abstract screening and full-text screening using pre-defined criteria and extracted included studies using a pre-defined data-extraction form. We coded included studies in NVIVO v12 to support thematic analysis and inform a best-fit framework.

Results: We identified 2763 unique records in our search, and included 42 studies and grey literature documents for extraction, coding, and thematic analysis. Included documents addressed adaptation (51%), adoption (42%), contextualization (33%), prioritization (28%), and implementation (60%) of health guidance. The themes included:

1) transparency in guideline prioritization and recommendation development is important for adoption, adaptation and implementation;
2) WHO headquarters should generate guidance that is more implementable;
3) contextualization of health guidance from global to local should be a continuous process to facilitate implementation;
4) GRADE ADOLOPMENT and the tool in GRADEpro could support contextualization/implementation of guidance; and
5) alternate pathways exist for health guideline contextualization/implementation.
6) We present a framework for conceptualizing different types of guidance across different levels in the formal WHO system, including the use of methodologies such as formal prioritization tools and GRADE adolopment to facilitate movement of guidance and eventual implementation.

Discussion: This CIS presents a conceptual framework for health guidance prioritization and adoption/adaptation in the WHO system. We identify that the GRADE adolopment methodology could serve to support the framework we present. This framework has relevance to the WHO system to support effective collaboration on health guidance prioritization and adoption/adaptation.

Consumer involvement: We worked with guidelines and WHO.

Process of identifying national health research priority areas in the Ministry of Health, Malaysia through evidence-based practice

Muhamad NA[1], Bakhtiar MF[2], Too CL[2], Abu Sapian R[3], Musa NSE[3], Johari MZ[4], Lai WH[5], Omar MI[3], Aris T[6]

[1] Sector for Evidence-Based Healthcare, National Institutes of Health, Ministry of Health, Malaysia; [2] Institute for Medical Research, National Institutes of Health, Malaysia; [3] National Institutes of Health, Malaysia; [4] Institutes for Behavioural Health research, National Institutes of Health, Malaysia; [5] Institute for Clinical Research, National Institutes of Health, Malaysia; [6] Institute for Medical Research, National Institutes of Health, Ministry of Health, Malaysia

Background: Policymakers in the Ministry of Health (MoH), Malaysia have been working to identify health research priority setting to improve the use of health research resources in terms of funding, personnel, and equipment in the most effective way under the 12th Malaysia Plan 2021-2025. To make evidence-informed decisions regarding the health research priority areas in the medical landscapes, the research core team under the 12th Malaysia Plan needs to provide evidence to justify reasons on the selected areas identified as health priorities. The process for health research priority-setting adapts the use of evidence-based practice in decision making.

Objective: To describe the steps of health research priority-setting in Malaysia and improve the use of resources for optimum uptake of research by using evidence-based practice and improve the health and well-being of all Malaysians.

Methods: Step 1: Engagement with decision-makers and stakeholders. We identified all decision-makers and stakeholders through their area of expertise. The list included public and private medical universities, research providers, research funders, government agencies and non-governmental organizations. We asked them about the 1) specific gaps in the area for health research, 2) skills and knowledge of relevance to the areas for health research, and 3) health research objectives and outcomes. Step 2: Providing support, information sharing and co-ordination related to setting health priorities The core team for health research priority-setting developed a survey form for both academics and stakeholders/policymakers which is called an advisory committee, in order to identify key health research areas to be prioritized. Step 3: Health research priority identification phase. Participants were split into a few working groups based on subject expertise to identify general priority research areas within their domain. Each group began by reviewing a list of potential health research topics and the knowledge gaps around that topic. Step 4: Draft ranking, advocate and support for all types of systematic reviews in identifying gaps for research priority Topics were placed in high priority, moderate priority, or low priority categories. These first draft will be presented to the whole group in plenary and an extensive discussion was held to refine the list. Lists were then revised into a second draft form. Step 5: Product document for health research priority-setting. The final draft which will be developed by the working group will be shared with all policymakers, stakeholders and researchers of the MoH, Malaysia and public and private medical universities for further comments.

Conclusions: Organizing this systematic priority setting process lays the groundwork for future priority setting workshops to be conducted using internationally recognized best practices.

Patient or healthcare consumer involvement: We invited NGOs and public at large to determine their area of interest during the preparation of health research priority-setting.

Quality appraisal of Chinese rapid guidelines

Yang N[1], Zhang J[2], Wang J[2], Ma Y[1], Wang M[3], Chen Y[1]

[1] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Cochrane China Network; Chinese GRADE Center, China; [2] School of Public Health, Lanzhou University; Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Cochrane China Network; Chinese GRADE Center, China; [3] Department of Radiology, the First Hospital of Lanzhou University, China

Background: In response to a public health emergency, such as the novel coronavirus (2019-nCov) outbreak, the rapid guidelines should be developed in a shorter time frame. However, a challenge with developing rapid guidelines is maintaining reporting and methodological rigor with in condensed timeframe.

Objectives: To evaluate the reporting and methodological quality of Chinese rapid guidelines with the RIGHT checklist (Essential Reporting Items for Practice Guidelines in Healthcare) and AGREE Ⅱ (Appraisal of Guidelines for Research and Evaluation).

Methods: We systematically searched MEDLINE (via PubMed), Chinse electronic databases including Wanfang Data, China National Knowledge Internet and Chinese Biology Medicine Library from the inception to February, 2020. We also searched the National Health Commission of PRC, the Chinese Medical Journal Network and Baidu to identify additional potential guidelines. Two independent review authors used The RIGHT (Essential Reporting Items for Practice Guidelines in Healthcare) instrument to perform a systematic appraisal with 22 items for included Chinese rapid guidelines. Each item was rated as "Yes" for total compliance, "Unclear" for partial compliance or "No" for non-compliance, respectively. We also calculated the number and proportion of reported items for each item. Statistical analyses were produced using SPSS version 16.0 for Windows. We appraised the methodological quality of the included Chinese rapid guidelines using the Appraisal of Guidelines for Research and Evaluation II (AGREE II) instrument. We assessed agreement among review authors by intra-class correlation coefficient.

Results and conclusions: This study is ongoing and results will be presented at the Cochrane Colloquium as available.

Reasons for randomized controlled trial retraction and their impact on clinical practice guidelines: a cross-sectional study

Shi Q[1], Wang Z[2], Zhou Q[1], Ma Y[2], Zhao S[3], Lu S[4], Xun Y[2], Chen Y[5]

[1] The First School of Clinical Medicine, Lanzhou University, China; [2] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, China; [3] School of Public Health, Lanzhou University, China; [4] School of Medicine, University of Electronic Science and Technology of China, China; [5] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, Lanzhou, China

Background: High-quality randomized controlled trials (RCTs) occupy an extremely important position on the evidence pyramid, either as standalone studies or as part of clinical practice guidelines. However, retracted RCTs may affect the development of clinical practice guidelines and hence recommendations to stakeholders. Understanding the reasons for the withdrawal of RCTs and their impact on clinical practice guidelines is essential for authors to circumvent common problems, as well as for the guideline developers and other evidence users to judge the reliability of RCT results.

Objectives: To investigate the reasons for RCT retraction and the impact of retracted RCTs on clinical practice guidelines.

Methods: Electronic databases (including PubMed and Embase) and Retraction Watch were retrieved from their inception until 27 March 2020. We searched "Retracted, Retraction, Withdrawal, Withdrawn, Randomized Controlled Trial, placebo" words in the form of keywords combined with free words. Two researchers independently screened the records and extracted data. We excluded non-medical RCTs and resolved disagreements through discussion or by consulting a third researcher. Data processing was performed by Stata statistical software 14 to present descriptive results.

Results: We retrieved 3471 studies and included 445 retracted RCTs. The reasons for reaction and the impact on clinical practice guidelines will be presented at the meeting.

Conclusions: The results will be presented at the meeting.

Patient or healthcare consumer involvement: None.

Survey of the needs, guideline development processes, and collaborative efforts of guideline-producing organizations

Siedler M[1], Sultan S[2], Ogunremi T[3], Morgan RL[4], Getchius T[5], Dahm P[2], Ginex P[6], Falck-Ytter Y[7], Fatheree L[8], Murad MH[9], Temple Smolkin R[10], Jakhmola P[11], McFarlane E[12], Mustafa R[13]

[1] University of South Florida, USA; [2] University of Minnesota, USA; [3] Public Health Agency of Canada, Canada; [4] McMaster University, Canada; [5] American Heart Association, USA; [6] Oncology Nursing Society, USA; [7] Case Western Reserve University, USA; [8] College of American Pathologists, USA; [9] Mayo Clinic, USA; [10] Association for Molecular Pathology, USA; [11] US Centers for Disease Control and Prevention, USA; [12] National Institute for Health and Care Excellence, UK; [13] University of Kansas Medical Center, USA

Background: Guidelines are developed by organizations of varying size, structure, and resources worldwide. However, little is currently known about the needs and challenges faced by guideline producers.

Objectives: To assess the perceived needs of guideline-producing organizations worldwide, with a specific focus on identifying challenges related to collaboration between organizations.

Methods: The survey targeted organizations and individuals involved in guideline development. Survey questions were developed by the U.S. GRADE Network and Guidelines International Network using an iterative approach. The survey was pilot-tested among attendees of a guideline development workshop, and included free-response, multiple-choice, and Likert-scale questions. We used electronic mailing lists, social media, and word-of-mouth to disseminate the survey using convenience and snowball sampling methods from November 2019 to April 2020.

Results: A total of 171 responses were included in the analysis, representing 30 countries and 116 unique organizations, which included professional societies, academic institutions, government agencies, industry, patient advocacy, and other stakeholders with membership sizes ranging from below 1000 to more than 30,000. The number of staff dedicated to guideline development ranged from less than 3 to more than 15. Nearly half (48.5%) of respondents indicated that they collaborated with other organizations to produce guidelines. A lack of adequate resources to develop rigorous guidelines rapidly was the most notable need indicated in the survey. Difficulties in reconciling differences in methodology and the time required to establish collaborative agreements were the most relevant barriers to collaboration.

Conclusions: Efforts to improve guideline development should focus on providing resources for rapid guideline development, capacity building, and facilitation of collaborative agreements.

Patient or healthcare consumer involvement: None.

The classification and management of conflicts of interest in Chinese clinical guidelines

Xun Y[1], Zhou Q[2], Shi Q[2], Wang Z[3], Yang N[1], Chen Y[1]

[1] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, Lanzhou, China; [2] The First School of Clinical Medicine, Lanzhou University, Lanzhou, China; [3] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; Cochrane China Network; Guideline International Network Asia; Chinese GRADE Center, China

Background: Conflicts of interest can affect the credibility, independence, and quality of clinical practice guidelines. Guideline developers, users, and researchers are trying to manage and reduce conflicts of interest worldwide. The evaluation results of the report quality of the Chinese guidelines show that the "Funding and Interest" report rate is only 10% to 15%. The classification and management of these conflicts of interest are unclear.

Objectives: To analyze the classification and management strategies of conflicts of interest of the Chinese guideline developers, and give suggestions on how to address conflicts of interest in clinical practice guidelines in China.

Methods: We will include in this review all Chinese guidelines available through MEDLINE, Web of Science, Embase, China National Knowledge Infrastructure, Wanfang, China Biology Medicine disc, Medlive (http://www.medlive.cn/) since inception until the present. These guidelines may be from any healthcare specialty or field for any disease or disease group. Two review authors will independently assess all guidelines and extract data on guideline specialty, guideline disease area, year of publication, developing body, whether reported the conflicts of interest or not, the classification and management of these conflicts of interest, etc.

Results and conclusions: This study is ongoing and results will be presented at the Evidence Summit as available.

Patient or healthcare consumer involvement: None.

The majority of reporting guidelines development lack conducting the Delphi method: a systematic search of reporting guidelines

Banno M[1], Tsujimoto Y[2], Kataoka Y[3]

[1] Department of Psychiatry, Seichiryo Hospital, Japan; [2] Department of Healthcare Epidemiology, Graduate School of Medicine and Public Health, Kyoto University, Japan; [3] Hospital Care Research Unit, Hyogo Prefectural Amagasaki General Medical Center, Japan

Background: Previous guidance of reporting guidelines recommends incorporating the Delphi method, a method to integrate the opinions of a group of experts, for consensus in the development of reporting guidelines.

Objectives: To clarify whether reporting guidelines generally perform the Delphi method, what factors may be associated with the employment of the Delphi method, and the reporting quality of the Delphi method in reporting guidelines.

Methods: We included all reporting guidelines (n = 244) in Enhancing the QUAlity and Transparency Of health Research (EQUATOR) Network published after 1 January 2011. We investigated the trends and factors associated with performing the Delphi method, using a Cochrane-Armitage test and multivariable logistic regression, respectively. Moreover, we assessed the reporting quality of the Delphi method among current reporting guidelines.

Results: Of 244 guidelines, 62 (25 %) performed the Delphi method for consensus development. The proportion of reporting guidelines that conducted the Delphi method was less than 10 % in 2011 and 2012, and 29 % in 2019. Year of publication (adjusted odds ratio (AOR) 1.35, (95% confidence interval (CI) 1.15 to 1.58)), number of authors (1 to 19: AOR 3.35 (95 % CI 1.48 to 7.59), 20 or more: 3.96 (95 % CI 1.70 to 9.19)), compared with 1 to 9 and multiple and simultaneous publications (AOR 2.54 (95 % CI 1.01 to 6.37)) were associated with the use of the Delphi method. The reporting quality of the Delphi method was moderate in most reporting guidelines using the Delphi method.

Conclusions: The use of the Delphi method in reporting guidelines is insufficient. Users and review authors should carefully appraise the consensus development in the reporting guidelines.

Patient or healthcare consumer involvement: None.

The use of Cochrane Reviews to inform clinical recommendations in the Danish National Clinical Guideline development

Korfitsen CB[1], Kirkegaard Mikkelsen M[1], Händel MN[1], Ussing A[1], Walker C[1], Rohde JF[1], Andersen HK[2], Tarp S[1]

[1] The Parker Institute, Bispebjerg and Frederiksberg Hospital and the Danish Health Authority, Denmark; [2] The Danish Health Authority, Denmark

Background: The purpose of the guidelines from the Danish Health Authority is to support evidence-based clinical decision-making. This is done by using the GRADE approach with a maximum of 10 outcomes per question (between one and three critical outcomes). The GRADE approach recommends using systematic reviews as the base for the development of healthcare recommendations. One of Cochrane's main aims is to inform clinical decision-making by developing independent and updated systematic reviews of high quality and relevance. Use of Cochrane Reviews in guideline development depends on discrepancies between the aim of the review and the guideline, especially in relation to the selection of outcomes.

Objectives: To explore the use of Cochrane Reviews as a source of evidence to inform clinical recommendations in guideline development.

Methods: We reviewed guidelines published since 2014, available at the Danish Health Authority website. Firstly, we noted recommendations based on evidence derived from Cochrane Reviews and topic (using the Cochrane topic taxonomy). Secondly, we categorised these review-informed recommendations according to the extent to which the pre-specified clinical question outcomes were covered: 1) all, 2) all critical, 3) some critical, or 4) any (non-critical).

Results: A total of 340 evidence-based recommendations and 211 expert consensus recommendations (no randomized controlled trials available) were published. Of the 340 recommendations, 104 (31%) included evidence partly or fully derived from Cochrane Reviews. Among these, 27 (26%) had all critical and non-critical outcomes, 34 (33%) had all critical outcomes, 34 (33%) had some critical outcomes and nine (9%) had only some non-critical outcomes estimated with evidence from Cochrane Reviews (fig. 1). Recommendations with evidence from Cochrane Reviews varied from 0% to 75% across topics, where most topics had coverage of 20% to 40%. Pregnancy and childbirth had the highest coverage of 75% of four recommendations whereas skin disorders had the lowest with 0% of 10 recommendations (fig. 2).

Conclusions: One-third of the evidence-based recommendations used Cochrane Reviews to inform the clinical recommendations and the coverage was similar across most topics. In 59% of these recommendations Cochrane Reviews were used for all critical outcomes. However, in 74% additional data extraction from primary studies was necessary because not all outcomes were informed by the Cochrane Review. This evaluation should inform future Cochrane Reviews to emphasize all outcomes important for clinical decision-making.

Patient or healthcare consumer involvement: Developing Danish clinical guidelines includes the involvement of patient organizations as part of the guideline-working group, patients reviewing clinical questions prior to searching the evidence base and patients are represented in the reference group reviewing and commenting on the final guidelines before publication.

The Use of GRADE methods to support the recommendations for health interventions in the World Health Organization public health guidelines

Chen N[1], Chen XL[1], Yang K[1]

[1] Evidence-Based Social Science Center, Lanzhou University, Lanzhou, China

Background: Interventions that seek to identify effectiveness require an assessment of the degree of certainty that estimates of outcomes reported in a group of studies are high enough to support decisions or recommendations.

Objectives: To characterize the distribution of recommendations strength and confidence in estimates of effect in World Health Organization (WHO) guidelines that have used the GRADE approach and graded strength of recommendations and confidence in effect estimates.

Methods: We reviewed the guidelines listed in the WHO guidelines database as of October 2019. We identified those that use GRADE and, in these guidelines, examined the distributions of strong and weak and associated confidence in estimates of effect (high, moderate, low, and very low).

Results: We identified 218 WHO public health guidelines; 164 (75.2%) related to GRADE methods. The remaining data extraction is in progress.

Conclusions: The current data indicate that strong recommendations based on low or very low confidence in effect estimates are very frequently made in WHO guidelines. Further study to determine the reasons for such recommendations is warranted.

Patient or healthcare consumer involvement: None

Use of GRADE in Chinese clinical practice guidelines

Xun Y[1], Zhou Q[2], Yang N[1], Wang Z[1], Shi Q[2], Ma Y[1], Luo X[3], Chen Y[1]

[1] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, Lanzhou, China; [2] The First School of Clinical Medicine, Lanzhou University, Lanzhou, China; [3] School of Public Health, Lanzhou University, Lanzhou, China

Background: Evaluating the quality of the evidence on which guidelines are based is an essential step in the development of appropriate guideline recommendations for practice, the use of which should be beneficial for patients.

Objectives: To determine whether and to what extent GRADE methodology has been and is currently being used in Chinese clinical practice guidelines.

Methods: We will include in this review all Chinese guidelines available through MEDLINE, Web of Science, Embase, China National Knowledge Infrastructure, Wanfang, China Biology Medicine disc, Medlive (http://www.medlive.cn/) since 2004 (which is when GRADE was mentioned) until the present. These guidelines may be from any healthcare specialty or field for any disease or disease group. Two review authors will independently assess all guidelines and extract data. We will assess guidelines to determine whether they have used the GRADE approach. Guideline specialty, guideline disease area, year of publication, developing body, whether used GRADE approach or not? etc. Appraise these guidelines using the AGREE II to find the difference of methodological quality.

Results and conclusions: This study is ongoing and results will be presented at the Evidence summit as available.

Patient or healthcare consumer involvement: None.

Use of systematic reviews and rapid evidence synthesis in setting up National Health Research Priorities in Malaysia

Muhamad NA[1], Abu Sapian R[2], Musa NSE[3], Bakhtiar MF[4], Too CL[4], Ab Ghani RM[1], Johari MZ[5], Lai WH[6], Aris T[4]

[1] Sector for Evidence-Based Healthcare, National Institutes of Health, Ministry of Health, Malaysia; [2] Research Management Unit, National Institutes of Health, Ministry of Health, Malaysia; [3] Collaboration and Innovation, National Institutes of Health, Ministry of Health, Malaysia; [4] Institute for Medical Research, National Institutes of Health, Ministry of Health, Malaysia; [5] Institute for Behavioral Health Research, National Institutes of Health, Ministry of Health, Malaysia; [6] Institute for Clinical Research, National Institutes of Health, Ministry of Health, Malaysia

Background: The Ministry of Health (MoH), Malaysia is committed to providing health services for all citizens and is interested in strengthening the country's health research capacity. Health research priority setting should be as evidence-based as possible, while also incorporating the views of a wide range of stakeholders. The use of systematic reviews to identify research gap in setting health research priorities ensures that research funded has the greatest potential public health benefit.

Objectives: To identify the range of health research activities in Malaysia for health research priority-setting.

Methods: The Research Policy and Planning Division under National Institutes of Health (NIH), MoH created a series of workshop since September 2019, with representation from the MoH, public universities, and external facilitators who were public health specialist and academician. We gathered information through rapid evidence synthesis to provide a body of knowledge to assist in the development of health research priority-setting for Malaysia. The MoH priority setting process was adapted from the Combined Approach Matrix and the Essential National Health Research methods. We identified areas through evidence mapping of systematic reviews and impact evaluations. The stakeholders focused on the quality of existing evidence for policy-makers and identified the research gaps. All the gaps were highlighted and synthesized.

Results: These workshops used a combination of two different priority setting models. The ENHR and CAM models were modified by the workshop facilitators, refined further by the workshop planning team, and adapted based on feedback during the workshop. Because of this combination, the workshop could be moulded to participants' needs and methods could be flexible in how they were used.

Conclusions: Organizing this systematic health research priority-setting process lays the groundwork for future health research priority-setting workshops to be conducted in using this combination of internationally recognized best practices. Therefore, it has been important to evaluate the process of health research setting in order to utilize the best method in the future. This exercise was conducted in a limited resource setting but was found to be feasible and recommendable. The input of constituencies that were not included in this workshop is important for improving the equity of a priority setting processes in the future.

Patient or healthcare consumer involvement: Adoption of policies by involving the public to identify values of equity and accountability in the healthcare systems. Identifying research gap with limited resources.

RESEARCH WASTE

A meta-epidemiological study: could prospective publication of protocols promote the overall reporting and methodological quality of non-Cochrane systematic reviews and meta-analysis?

Lu T[1], Wang C[1], Lu C[1], Ke L[2], Ge L[1], Yang K[1]

[1] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China; [2] Department of Liver Surgery, The First Affiliated Hospital, Sun Yat-sen University, China

Background: The association between prospective registration and the overall reporting and methodological quality of systematic reviews (SRs) has been investigated. The impact of prospective publication of protocols on publicly reviewed journals and prospective registration only on the quality of SRs and meta-analysis (MAs) may be different. However, whether prospective publication of protocols could improve overall reporting and methodological quality of SRs and MAs still remains unknown.

Objectives: The primary purpose of this study is to investigate whether the overall reporting and methodological quality of SRs and MAs with published protocols are superior to those of MAs only registered.

Methods: We will search the PubMed database to identify (paired) non-Cochrane SRs and MAs of randomized controlled trials (RCTs) published from 2015 to 2019 in English. The main search terms include: "meta-analysis", "meta-analyses", "systematic review", "systematic reviews" and "random*". We will divide the rest of relevant reviews into two groups: non-Cochrane SRs and MAs with published protocols and those only registered when we finish titles and abstracts screening. For each group, the first 60 eligible studies from each group will be randomly selected. We will separately extract and cross-check key information (such as first author, number of authors, year of publication, journal, types of disease, details of intervention and control, sample size, number of included RCTs, funding, etc.) of SRs and MAs. Two independent review authors will assess the reporting and methodological quality of included studies using PRISMA and R-AMSTAR (A MeaSurement Tool to Assess systematic Reviews). We will resolve any disagreements through discussion. The assessment method of reporting quality based on each PRISMA item will be: "1" point for each "yes", "0.5" for each "partial", and "0" point for any other responses ("no" and "cannot answer"). The assessment method of methodological quality based on each R-AMSTAR item will be: "1" point for zero or one criteria met, "4" points for all criteria met and other situations for "2" or "3" points. For each PRISMA item, we will calculate odds ratio (OR) with 95% confidence interval (CI) to compare the complete compliance of two groups. We will calculate the mean difference (MD) and 95% CI for each item of R-AMSTAR to compare the methodological quality of two groups. We will also use univariable and multivariable linear regression analysis to explore basic characteristics (such as number of authors, year of publication, sample size, number of included RCTs, funding, etc.) whether they are associated with the overall reporting quality and met.

Results: This study is ongoing, and the result of meta-epidemiological study will be submitted to a peer-reviewed journal for publication.

Conclusions: This study will provide comprehensive information on whether the overall reporting and methodological quality of SRs and MAs with published protocols are better than those of MAs only registered.

Patient or healthcare consumer involvement: Not applicable.

Challenges and opportunities of evidence-based research in rapid learning health systems

Klugar M[1], Lund H[2], Brunnhuber K[3], Valimaki M[4], Yost J[5], Nußbaumer-Streit B[6], Puljak L[7], McCurtin A[8], Bala M[9], Klugarová J[1]

[1] Czech National Centre for Evidence-Based Healthcare and Knowledge Translation (Cochrane Czech Republic, Czech CEBHC: JBI Centre of Excellence, Masaryk GRADE Centre), Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University, Czech Republic; [2] Centre for Evidence-Based Practice; Western Norway University of Applied Sciences, Norway; [3] Elsevier, Digital Content Services, Data Platform Operations, UK; [4] University of Turku, Faculty of Medicine, Department of Nursing Science; Xiangya Nursing School, Central South University, Finland and China; [5] M. Louise Fitzpatrick College of Nursing, Villanova University, USA; [6] Cochrane Austria, Department for Evidence-Based Medicine and Evaluation, Danube-University Krems, Austria; [7] Center for Evidence-Based Medicine, Catholic University of Croatia, Croatia; [8] School of Allied Health, University of Limerick, Ireland; [9] Department of Epidemiology and Preventive Medicine, Medical College, Jagiellonian University, Poland

Background: In our era of information, "oversaturation" is a rapid essential keyword, especially in the context of health systems and health policymakers. On the other hand, health systems need high quality unbiased, not confounded evidence which will be implemented in a timely way, according to the approach known as evidence-based health care. The fundamental question is whether the researchers who are developing primary or secondary research are also following the evidence-based approach, which we can call evidence-based research (EBR). European COST Action EVBRES (Evidence-Based RESearch) defined EBR as "the use of prior research in a systematic and transparent way to inform new study so that it answers the questions that matter in a valid, efficient and accessible manner." The EBR approach is needed to prevent research waste. It includes the whole research process from study concept and design through to placing new results in the context of earlier similar evidence.

Objectives: The objective of this special session is to explore the challenges and opportunities of evidence-based research in rapid learning health systems. 1) Raising awareness and accepting the EBR challenge by all relevant stakeholders 2) Preparation of educational materials for teaching the EBR 3) Improvement of co-operation among relevant stakeholders to increase efficiency, update and development of systematic reviews.

Methods: EVBRES consists of four working groups using a variety of quantitative, qualitative, and mixed methods (e.g. scoping reviews, systematic reviews, Delphi studies, qualitative studies) to design and implement interventions to promote the EBR approach among various stakeholders.

Results: Working group one will describe key stakeholders' role, such as ethic committees, funding agencies, journals and patient groups, in solidifying the EBR approach. Working group two will introduce piloted designed training school of EBR; approach to best design and pedagogical methods enabling permanently available online training school of EBR; and barriers and facilitators to become evidence-based in research. Working group three will show improvements in the efficiency of production and updating of systematic reviews. Working Group four will explore methods to detect redundant research as well as measurable outcomes of implementing evidence-based research approaches that are relevant to researchers and key stakeholders.

Conclusions: EVBRES currently has participants from 39 European COST Action member countries. In addition, more than 10 international partner countries are also involved. As a result of presenting this special session for Cochrane community EVBRES looks forward to the valuable discussions of evidence-based healthcare experts to forward our thinking on how EBR fits in rapid learning health systems.

Patient or healthcare consumer involvement: Patients and consumers who are part of the EBR approach are very important stakeholders in EVBRES.

Don't throw the baby out with the bathwater: are publishers' requirement for searches to be 'up-to-date' penalizing more complex and rigorous reviews?

Stokes G[1], Sutcliffe K[1]

[1] EPPI-Centre, University College London, UK

Background: Journal editors and reviewers are often concerned that results of systematic reviews may be out-of-date, and thereby provide misleading information to decision-makers. Searches conducted a year or more previously can result in a journal rejecting a review, or in a requirement that searches are updated. However, as review methods become increasingly complex, synthesis requires additional time placing many reviews at risk of rejection by publishers.

Objective: To consider whether the focus on rapid publication might hinder production of the most rigorous or ambitious reviews.

Method: We focused on recent developments in review designs and methods to consider whether current journal guidelines may be at odds with newer review approaches intended to increase review quality.

Results: A variety of factors may improve the quality of systematic reviews whilst also increasing the time needed to undertake them. Several recommended review processes increase the number of steps required in a review. These include: methods to enhance rigor and transparency in presentation of review findings such as GRADE and CerQUAL; methods to enhance transparency relating to intervention components such as the TiDIER checklist; and efforts to enhance patient and public involvement (PPI) in the post-search phases of reviews. Several factors increase the volume of literature being synthesized. As the extent of available research literature continues to grow exponentially this can result in larger numbers of studies to screen for inclusion in each review. Similarly, the scope of many reviews is broader than the PICO approach in which a single intervention, outcome and population are specified. EPPI-Centre's work with policy-makers has shown that policy concerns often require reviews that cover a range of potentially useful interventions to enable consideration of where best to invest resources. Other factors relate to the extent of analysis required to ensure utility for decision-makers. Review approaches designed to understand the implications of intervention and contextual complexity, such as mixed-method research synthesis or realist synthesis, often involve multiple synthesis stages. A related point is that these reviews often require a high level of conceptual development and interpretation which also increases the time required for synthesis. Hi-tech solutions for searching and screening may reduce the time burden for parts of the review process in the near future, but this is unlikely to impact significantly on reviews that require iterative searching or high levels of interpretation.

Conclusion: Publishers' focus on timely publication of reviews poses a risk that research teams will favour questions that are easier to review quickly and avoid questions that require a broader scope or deeper analysis, ultimately leading to production of reviews that are less useful to decision-makers.

Patient or healthcare consumer involvement: PPI is highly relevant. To avoid tokenism ample time is needed to engage in PPI at various stages of a review.

Identifying resource intensive areas of systematic review production and updating: a scoping review

Nussbaumer-Streit B[1], Ellen M[2], Klerings I[1], Baladia E[3], Spijker R[4], Working Group 3 AG8 EVBRES COST Action CA17117

[1] Cochrane Austria, Austria; [2] Department of Health Systems Management, Guilford Glazer Faculty of Business and Management and Faculty of Health Sciences, Ben-Gurion University of the Negev, Israel; [3] Red de Nutrición Basada en la Evidencia (RED-NuBE), Academia Española de Nutrición y Dietética, Pamplona, Spain; [4] Cochrane Netherlands, The Netherlands

Background: Ideally, primary researchers base new studies on prior studies to avoid research waste. In case no up-to-date systematic review (SR) on a topic is available, primary researchers would need to conduct or update one themselves. However, the resource intensity associated with this process often keeps primary researchers from doing so. To facilitate the uptake of evidence-based research, the production and update of SR need to be accelerated.

Aim: To identify the most resource intensive areas when conducting or updating a SR in order to find out where the largest gain in improving efficiency whilst sustaining quality might be possible. In addition, we would like to know why these areas are resource intensive in order to identify suitable methods to address them.

Methods: We will conduct a scoping review (protocol available: osf.io/fby54). We will include SR and empirical and simulation studies that assess resource use in systematic reviews of health interventions, diagnostic, or prognostic studies, without any limits on languages or publication status. We will also include qualitative studies that assessed reasons why diverse steps are resource intensive. We will limit our search to studies published from 2009 onwards to get an overview of resource use of current SR processes. An information specialist will perform the search, following the three-step process recommended by the Joanna Briggs institute. In the first step, we will conduct a limited, focused, search of Ovid MEDLINE. We will perform a second, comprehensive, search using all identified keywords and index terms across the following databases: Ovid MEDLINE, Scopus (Elsevier), Science Citation Index Expanded, Social Sciences Citation Index, and Current Contents Connect (via Web of Science). In the third step, we will screen the reference list of identified reports and articles for additional studies. We will also conduct reference chasing, search for dissertations and theses via ProQuest, and conduct a handsearch of relevant conference proceedings (e.g. the Cochrane Colloquium).

Results: We plan to complete the scoping review in August and be able to present results in October at the Cochrane Colloquium. As a result of this research project we aim to give an overview of the most resource intensive steps of the systematic review process, as well as factors influencing this resource intensity. Results of this scoping review will be complemented by a qualitative study on reasons for resource intensity of SR steps and will feed into a Delphi study that aims to prioritize areas in the systematic review process and methods that are most relevant and promising for speeding up the review process. This should guide future methods improvement and validity studies in this area and ultimately help to accelerate the systematic review production without compromising quality.

Patient or healthcare consumer involvement: Increasing evidence-based research ultimately benefits users of healthcare: research becomes more efficient and results can improve practice faster.

Identifying the evidence gaps in Cochrane fertility reviews

Showell M[1], Stedman K[1], Farquhar C[1], Jordan V[1]

[1] University of Auckland, New Zealand

Background: Cochrane aims to produce the 'gold standard of evidence'. The Cochrane Gynaecology and Fertility Group (CGF) produce systematic reviews (SRs), within this subject area, using trials from a database of over 19,500 randomized controlled trials (RCTs). We believe many of these RCTs are not currently adding to the evidence base as there are gaps in the topics covered by the existing SRs.

Objectives: The main aim of this project is to identify gaps in evidence by identifying the RCTs published in the area of gynaecology that have not been used in Cochrane SRs.

Methods: We conducted an audit of all fertility trials in the CGF specialized register, published from 2010 to 2011. The search function in CRS web could quickly identify which trials had been 'included', 'excluded' or simply 'not used' in SRs across the Collaboration. We then classified the 'not used trials' into their population/condition and intervention. From this point we looked at the existing SRs in CGF to assess whether they in fact could be incorporated. The remaining 'not used trials' were grouped into potential review topics.

Results: Our database search found 564 trials from 2010 to 2011. Of these, we excluded 59 as they did not meet inclusion criteria, 318 were already used in Cochrane SRs, of the unused trials, 108 could fit into a review if the review was updated (19%), and 79 trials (14%) had no existing SR topic and were classified as 'unused'. The population groups were categorized as 'women', 'men', 'oocytes' and 'sperm', the intervention categories varied in accordance to the different treatments for each population. The largest population group with unused trials was 'Women' (52%) and within this group, the most common unused intervention was 'Chinese herbal medicines' (17%). In the second largest population group 'Oocytes' (23%), 'Preservation' was the largest intervention group (28%). In order to develop new review titles, we looked at those topics that had at least three unused trials and developed eight proposed titles. The new title with the most unused trials was 'Chinese medicine for women undergoing assisted reproductive technologies' for which we found six unused trials.

Conclusions: We were satisfied that Cochrane SRs were covering most topics in fertility, however there are improvements to be made, both in developing new titles and in updating existing reviews. By identifying the populations and interventions not currently systematically reviewed by Cochrane, we can now develop priority topics and thus provide better healthcare evidence and reduce research waste.

Patient or healthcare consumer involvement: Subfertile couples will be asked to prioritise the proposed new review titles, in terms of importance, and we will only progress the most relevant of these titles into Cochrane fertility SRs.

Impact of reporting quality on risk-of-bias assessment in occupational health and safety trials

Tikka C[1], Verbeek J[2], de Boer AGEM[3], Hoving J[2], Boschmann JS[2], Hulshof CTJ[3]

[1] Finnish Institute of Occupational Health, Finland; [2] Cochrane Work Review Group, Academic Medical Centre, University of Amsterdam, The Netherlands; [3] Coronel Institute of Occupational Health, Academic Medical Centre, University of Amsterdam, The Netherlands

Background: It is our impression that trials used in Cochrane Work Reviews are not well reported and are often considered at high risk of bias. We hypothesized that if trials were better reported this would facilitate the risk-of-bias assessment and result in a lower risk of bias. We focus on the most important aspect of randomized controlled trials (RCTs): randomization and allocation concealment.

Objectives: We want to assess how well authors of occupational health and safety (OHS) trials report the process of randomization and allocation concealment. In addition, we want to find out if better reporting quality is associated with a lower risk-of-bias assessment.

Methods: We conducted a systematic literature search in PubMed for RCTs published after 2010 in 18 OHS journals, with workers or workplaces as participants, irrespective of intervention and comparison. We excluded cross-over trials, protocols, pilot studies and studies that only report secondary outcomes, cost-effectiveness analyses, and exposure studies. We rated the quality of reporting as high, moderate, low, or very low for each article based on the number of CONSORT 2010 checklist items complied with (100%, < 100% to 75%, < 75% to 50%, < 50%). We restricted the assessment to items relevant for random sequence generation, allocation concealment mechanism, and randomization implementation. We applied the Cochrane risk-of-bias 2.0 tool to assess the risk of bias of the random sequence generation and allocation concealment for one primary outcome for each trial based on data available from publications without contacting the authors for additional information.

Results: We included 136 articles reporting on 130 randomized and cluster randomized trials (cRCTs). The reporting quality was high in 13 articles, moderate in 16, low in 28, and very low in 79 articles. Quality was slightly better for articles of RCTs (n = 91) than cRCTs (n = 39) with 26% and 13% high to moderate quality articles. CONSORT interpretation varied between assessors and needed considerable discussion before consensus was reached. The same held for the risk-of-bias 2.0 tool. We will report on the association of a low risk of bias assessment in studies to a high, moderate and low or very low reporting quality.

Conclusions: We found that reporting quality was low or very low for most articles reporting on randomized and cluster randomized trials in OHS journals, but this is not different from general medical journals. Clearer guidance from journals, CONSORT and Cochrane might help trial authors better implement, report and judge the techniques used to reduce the risk of selection bias in randomized controlled trials, such as through simplification of language, providing templates and examples in guidance documents.

Patient or healthcare consumer involvement: Better reporting of trials may improve the quality of research used to inform healthcare decisions and help consumers critically appraise and interpret trial results.

Impact of retracted randomized controlled trials on systematic reviews and meta-analysis: a cross-sectional study

Zhou Q[1], Shi Q[1], Wang Z[2], Lu S[3], Chen Y[4], Yang K[4]

[1] The First Clinical Medical School, Lanzhou University, China; [2] Evidence-Based Medicine Center, Lanzhou University, China; [3] School of Medicine, University of Electronic Science and Technology of China, China; [4] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, Lanzhou, China

Background: Randomized controlled trials (RCTs) are usually incorporated into systematic reviews and meta-analyses to form the best decision evidence. In recent years, due to academic misconduct and other reasons, many RCTs have been retracted. It would be useful to investigate the number of retracted RCTs that are included in systematic reviews and meta-analyses, the impact of this on their results, and how such research waste can be avoided.

Objectives: To investigate the impact of retracted randomized controlled trials on systematic reviews and meta-analysis.

Methods: We searched electronic databases (including PubMed and Embase) and Retraction Watch from inception to 27 March 2020. We searched "Retracted, Retraction, Withdrawal, Withdrawn, Randomized Controlled Trial, placebo" words. We conducted citation searching of the Web of Science database to find systematic reviews and meta-analyses that cited retracted randomized controlled trials. We will conduct a sensitivity analysis of systematic reviews and meta-analyses to explore the impact of retracted randomized controlled trials.

Results: The results will be presented at the meeting.

Conclusions: The conclusions will be presented at the meeting.

Patient or healthcare consumer involvement: None.

Publicly funded research in New Zealand: is the money being well spent?

Jordan V[1], Greenwood G[2], Glen J[3], Farquhar C[4], Showell M[4]

[1] University of Auckland, New Zealand; [2] University of Otago, New Zealand; [3] Health Research Council of New Zealand, New Zealand; [4] University of Auckland, New Zealand

Background: In order to fulfill the implied contract between researchers and the participants of research, trial results need to be disseminated. Patients and carers have identified publication and dissemination as one of the key areas of research wastage that they are concerned about. Australian research has reported the time lag between when funding was allocated by the National Health and Medical Research Council (NHMRC) and when results were published. Of trials that were funded between 2008 and 2011 only 50% had been published. The median time was 7.1 years after funding had been received. This research wastage was estimated to be AUD $30 million over a three-year period. The Health Research Council (HRC) in New Zealand is the "crown agency which is responsible managing the Government's investment in health research". The HRC funds research programmes (up to $5 million of funding over five years) and projects (up to $1.2 million over three years). These programmes and projects have been selected through a rigorous peer review system.

Objectives: To determine the length of time to publication for publicly funded research in New Zealand.

Methods: We investigated projects and programmes funded by the HRC from 2006 to 2014. The HRC supplied us with the collated programme and project information publicly available on their website. In order to determine if researchers had disseminated their research findings we undertook a search of the literature databases using investigators names and subject area. To find clinical trial registrations we searched the World Health Organization International Clinical Trials Registry Platform and Clinical trials.gov. For conference publications and journal articles we searched MEDLINE, Embase and Google Scholar. Searches were conducted from December 2019 to February 2020.

Results: A total of 374 programmes and projects were funded over this time. We were able to identify published findings for 191 of these (51%) which means 183 (49%) of the programmes and projects that were publicly funded had not published results that were able to be found in peer-reviewed journals. This unpublished research had received in total $258,988,275 NZD. Data will also be presented on time to publication for those published programmes and projects and a more in-depth look by discipline and funding level will be discussed.

Conclusions: There is serious concern about the level of dissemination of publicly funded research. Funding research granting bodies need additional resources in order to actively monitor funded projects. This could ensure that dissemination of results are achieved. Research funding organizations should also consider levers to encourage publication of complete and usable research results such as making public the outputs of their own monitoring processes (as per the WHO joint statement on public disclosure of results from clinical trials).

Patient or healthcare consumer involvement: Failure to disseminate clinical research is a betrayal of the relationship between researchers and consumers.

Quality assessment of evidence-based checklists identifying predatory journals and comparison of checklist items against a consensus definition

Ng JY[1], Haynes RB[1]

[1] Department of Health Research Methods, Evidence and Impact, Faculty of Health Sciences, McMaster University, Hamilton, Canada

Background: Researchers, clinicians and policy makers rely on the validity of published scientific research to make well-informed decisions that can have significant and life-changing impacts. Predatory journals (PJs) currently pose a threat to the quality and integrity of scientific publishing.

Objectives: We conducted a systematic review to identify evidence-based checklists (EBCs) to detect PJs, and compared checklist items against a consensus definition.

Methods: We updated a recently published systematic review of EBCs to detect PJs. All identified checklists were assessed using Cukier et al.'s (2019) five-item risk of bias assessment; checklists scoring "yes" for three or more items were deemed to be an EBC. Boateng et al. (2018)'s nine steps of scale development was used to identify how many steps were completed in the generation of each EBC. Items from each of the EBCs were then compared against the five components of Grudniewicz et al.'s (2019) consensus definition of predatory journals and publishers (PJPs), a standard that did not exist when the checklists were published. Table 1 contains methodological studies cited in this abstract.

Results: We identified four EBCs. None of them had completed step 1 of Boateng et al. (2018)'s 9 steps of scale development and validation, indicating their early stages of development. We assessed 47 items from the four EBCs against Grudniewicz et al. (2019)'s definition of PJPs, of which 28 items met the definition. Eighteen of these items were the same or very similar to at least one other item, while 10 items were unique.

Conclusions: Checklists to detect PJs are at an early stage of development and lack assessment of reliability and validity. To a varying degree, the EBCs contain items that match the new consensus definition of PJs, but also lack some of its features. The most obvious reason for this is that there had not been consensus on a definition for PJs at the time that the scales were published. With the recent consensus definition, checklists now have a clearer target, can make adjustments, and hopefully proceed to address methods standards for measurement scales.

Patient or healthcare consumer involvement: None directly as yet – consumers and authors of the medical literature will be involved in steps leading to a consensus scale for evaluating open access journals for predatory features.

Table 1. Methodological studies informing the current study

Citation	Relevant Methodology
Cukier S, Helal L, Rice DB, Pupkaite J, Ahmadzai N, Wilson M, Skidmore B, Lalu M, Moher D. Checklists to detect potential predatory biomedical journals: A systematic review. medRxiv. 2019 Jan 1:19005728.	5-item risk of bias tool for identifying evidence-based checklist
Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: a primer. Frontiers in Public Health. 2018 Jun 11;6:149.	9 steps of scale development and validation
Grudniewicz A, Moher D, Cobey KD, Bryson GL, Cukier S, Allen K, Ardern C, Balcom L, Barros T, Berger M, Ciro JB. Predatory journals: no definition, no defence. Nature. 2019 Dec.	Consensus definition of "predatory journals and publishers" based on a three-round modified Delphi survey

Quality of reporting for randomized controlled trials published in Latin American and Spanish journals: a survey of three clinical specialities

Bachelet VC[1], Navarrete MS[1], Barrera-Riquelme C[1], Carrasco VA[1], Díaz RA[1], Ibarra ÁA[1], Lizana FJ[1], Meza-Ducaud N[1], Pardo-Hernandez H[2], Rivera-Bacilio J[1], Saavedra MG[1], Tapia-Davegno C[1], Villanueva J[3]

[1] Escuela de Medicina, Facultad de Ciencias Médicas, Universidad de Santiago de Chile, Chile; [2] Iberoamerican Cochrane Center, Biomedical Research Institute Sant Pau (IIB Sant Pau) – CIBER Epidemiología y Salud Pública (CIBERESP), Spain; [3] Facultad de Odontología, Universidad de Chile, Chile

Background: The Consolidated Standards of Reporting Trials (CONSORT) reporting guideline was developed to improve the accuracy, completeness, and transparency of the reporting of randomized controlled trials (RCTs). Gaps and imprecisions of reporting hamper the assessment of the methodological quality and internal and external validity. Few studies have been conducted on the impact of CONSORT on RCTs published in Latin American and Spanish journals.

Objectives: To assess the reporting quality of RCTs of three clinical specialities published in Spanish and Latin American journals, as well as to evaluate changes over time and associations of quality with journal and country indicators.

Methods: We conducted a systematic survey of all RCTs published in Spanish-language journals in dentistry, neurology, and geriatrics. Our data source was the BADERI database that inputs into the Cochrane Central Register of Controlled Trials (CENTRAL). BADERI is a repository of journals that have been handsearched and RCTs that have been identified. We included randomized or quasi-randomized controlled trials with a recoverable full text published between 1990 and 2018. We excluded articles not addressing the clinical specialities of interest, conference proceedings, pilots or feasibility studies, secondary analysis on RCTs, translations of RCTs published elsewhere, letters, and editorials. After updating and completing BADERI to more recent years, we studied the whole population of RCTs for each clinical field, so no sample size calculation was done. We extracted data on 23 CONSORT items, plus four additional items (Table 1). The primary outcome was the total score of the predefined 23 CONSORT 2010 items for each RCT. The secondary outcome was the overall score for each CONSORT 2010 item predefined for this study.

Results: We report the interim analysis on 165 included dental RCTs from a total of 369 eligible for selection. We report the number of RCTs published by country and by time periods (Table 2). Seven countries of the region have published RCTs with a mean score of 13.2 (score range 0 to 34). The highest mean score is Colombia (15.0), and the lowest is Peru (7.7). The mean score in the 1990 to 2006 period was 10.9, which increased to 16.1 for the 2016 to 2018 period (Table 2).

Conclusions: Despite the widespread adoption of the CONSORT reporting guideline, there is still a significant gap in reporting standards in the Spanish-language journals that may hamper the inclusion of RCTs from Latin America and Spain in systematic reviews. This study provides evidence of this gap and should raise awareness in the region for the need to strengthen the methodological competencies of local clinical investigators and editors.

Patient involvement: None.

Additional files: Table 1; Table 2

Quality of reporting in randomized controlled trials of therapeutic cardiovascular medical devices

Zhang L[1], Chen W[1], Gong X[1], Zhang M[1]

[1] West China Hospital of Sichuan University, China

Background: Therapeutic medical devices play an important role in the treatment of cardiovascular diseases. The reliability of the randomized controlled trial, which is the best design for assessing treatment effects, largely depends on the information found in published reports. Information regarding the quality of reporting in trials of therapeutic medical devices is limited.

Objectives: To explore the quality of reporting (compliance with the Consolidated Standards of Reporting Trials (CONSORT) statement) in randomized controlled trials of therapeutic cardiovascular medical devices.

Methods: We conducted a cross-sectional study to assess the reporting quality of randomized controlled trials that tested the effects of therapeutic cardiovascular medical devices. The quality of reporting was assessed against a modified CONSORT checklist, including 47 items from the CONSORT statement and CONSORT extension. We also examined the specific items regarding medical devices. We undertook univariable and multivariable linear regressions to explore potential factors associated with CONSORT scores.

Results: We identified 115 randomized controlled trials. The mean (standard deviation) CONSORT score was 20.5 (5.0). The extent of compliance with the CONSORT reporting guideline differed substantially across items: 5 of the 47 items were reported adequately across trials (more than 90%), and 10 were reported adequately in less than 5% of trials. Less than 50% of the trials reported additional items related to the medical device. Multivariable regression analysis showed that trials published in general journals (coefficient 7.44, 95% confidence interval (CI) 5.50 to 9.38), with larger sample sizes (coefficient 2.30, 95% CI 0.76 to 3.83), and multi-center studies (coefficient 3.14, 95% CI 1.27 to 5.01) were associated with a higher quality of reporting.

Conclusions: The overall reporting quality in randomized controlled trials of therapeutic medical devices is suboptimal, particularly in terms of items regarding surgeons and hospitals. We suggest that the existing CONSORT statement and extension should be modified to be more applicable to therapeutic medical devices.

Patient or healthcare consumer involvement: Further guidelines or checklists for reporting should be developed, including more important information about device-related items, especially about device information and surgeons.

Redundant clinical trials challenging research ethics and hurting patients

Jia Y[1], Wen J[1], Qureshi R[1], Ehrhardt S[1], Celentano D[1], Robinson K[2]

[1] Bloomberg School of Public Health, the Johns Hopkins University, USA; [2] School of Medicine, the Johns Hopkins University, USA

Background: Recently the proliferation of scientific publications has led China to become the largest producer of scientific and clinical evidence. However, there are concerns over the redundancy of the research produced by Chinese scientists. A study is needed to evaluate the redundancy of clinical trials to save resource and protect patients.

Objectives: To estimate the number of extra major adverse cardiac events (MACEs) that occurred among patients with coronary artery disease (CAD) who were deprived of statins in redundant clinical trials conducted in Mainland China.

Methods: We searched PubMed, Embase, the Cochrane Central Register of Controlled Trials (CENTRAL), and SinoMed until March 2020 for randomized or quasi-randomized trials conducted in Mainland China comparing statins with placebo or no treatment among patients with any subtype of CAD. In the primary analysis, we defined trials as clinical practice guidelines (CPGs)-based redundant trials as those conducted after two CPGs were released explicitly recommending statins to patients with stable angina pectoris and acute coronary syndrome. The primary outcome was the number of extra clinical events, including MACEs, that were attributable to deprivation of statins in CPG-based redundant trials. The extra clinical events were the difference between the actual clinical events that occurred in the control groups and the expected clinical events if patients in the control group had received treatment as in the statins group. In the sensitivity analysis, we defined trials as cumulative meta-analyses (CMAs)-based redundant trials as those conducted after CMAs illustrated that sufficient evidence had accrued to confirm the benefits of statins.

Results: After the CPGs were released, 1864 and 91 redundant trials were initiated or continued recruiting, respectively. In total the CPG-based redundant trials recruited 197,296 patients, of which 96,481 were allocated to the control group and deprived of statins for 26,295 person-years. More than 4484 extra clinical events were reported in 401 trials, including 3367 MACEs and 1117 other or unspecified events. The 3367 MACEs consisted of 623 deaths, 939 cases of myocardial infarction, 201 cases of stroke, 113 cases of revascularization, 397 cases of heart failure, and 1097 cases of relapsed or deteriorated angina pectoris. The 2302 CMA-based redundant trials reported more than 6289 additional clinical events, including 4745 MACEs and 1.544 other or unspecified events. The 4745 MACEs consisted of 834 deaths, 1355 cases of myocardial infarction, 256 cases of stroke, 208 cases of revascularization, 508 cases of heart failure, and 1586 cases of relapsed or deteriorated angina pectoris.

Conclusions: Redundant and possibly unethical clinical trials conducted in Mainland China for statins among patients with CAD have resulted in numerous avoidable deaths and major cardiac events.

Patient or healthcare consumer involvement: None.

Reporting of essential items in comparative accuracy studies: a literature review

Vali Y[1], Yang B[1], Olson M[1], Leeflang MMG[1], Bossuyt PMM[1]

[1] Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands

Background: About half of the diagnostic test accuracy (DTA) reviews in the Cochrane Library are comparative DTA reviews. The ideal design of primary studies in these reviews are comparative accuracy studies, which compare the relative performance of two or more diagnostic tests. As any other form of research, such studies should be reported in an informative manner to allow replication and to be useful for decision-making.

Objectives: To assess whether and how essential components of test comparisons were reported in comparative accuracy studies.

Methods: We evaluated 100 comparative accuracy studies, published in 2015, 2016 or 2017, randomly sampled from 238 comparative DTA systematic reviews. We extracted information on 20 reporting items, pertaining to the identification of the comparison, its validity, and the actual results of the comparison.

Results: Most comparative accuracy studies (n = 73) failed to identify the comparative nature of the study in their title. A majority did not report which index tests were compared in their methods section (n = 67). About a third (n = 36) did not report the comparison as a study objective or a hypothesis. Although most studies (n = 86) reported how participants had been assigned to index tests, we could often not evaluate whether index test interpreters had been blinded to the results of other index tests (n = 66), nor could we identify the sequence of index tests (n = 51) or the methods for comparing measures of diagnostic accuracy (n = 59). Joint contingency table (2x2x2 tables) data were only reported by nine from 89 paired comparative studies. More than half of studies (n = 64) did not provide measures of precision (such as P values or confidence intervals) for the reported comparative accuracy estimates.

Conclusions: Essential components of test comparisons are frequently missed or incompletely reported by comparative accuracy studies in systematic reviews. This could consequently impede identification of these studies and interpretation of their findings. Explicit guidance for reporting comparative accuracy studies may facilitate the production of full and informative study reports.

Patient or healthcare consumer involvement: Well-conducted comparative accuracy studies, have the potential to yield high-certainty evidence for informing clinical decision-making regarding tests. Considering their importance, they should be reported meticulously to allow evidence users to appreciate their findings and to consider their applicability to different patient groups and healthcare settings.

The reality of unclearly reported RCTs during the selection period of systematic review: a telephone interview survey

Lu CL[1], Jiang Y[1], Zhang X[1], Xiong M[2], Jin XY[3], Li X[1], Liu JP[1]

[1] Centre for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, China; [2] Dongzhimen Hospital, Beijing University of Chinese Medicine, China; [3] Third Affiliated hospital, Beijing University of Chinese Medicine, China

Background: With the rapidly increasing number of randomized controlled trials (RCTs) conducted and published, systematic review (SR) searches find a lot of studies. It is often hard to deal with meta-analysis and other syntheses. Some trials did not report completed information on methodology, such as just reporting the "random" without specific information. If review authors cannot contact the trials authors, they have to identify trials as unclear risk of bias and downgrade certainty.

Objectives: To confirm the reality of unclearly reported trials and the eligibility to be included in SRs, and to provide some suggestions for rapid review and the avoidance of research waste.

Methods: Based on the update of a SR of RCTs from 2013 to 2019, we conducted telephone interviews with all authors of 283 trials that only reported "random". We asked about the following details in the following order: specific randomization methods, concealment methods, diagnostic examination methods, dropout and losing visiting participants details and reasons, protocol registry number or link, funding details, sample size calculation method, and intervention details. We searched the contact number based on the authors' affiliation and email address from the Internet. If the interviewees answered the wrong randomization methods, we terminated the interview.

Results: Using the author affiliation information from the articles, we failed to find all authors of 147 (51.94%) trials, and the reception desk of the hospital or the department said there was no such person. We successfully contacted all the authors of 122 (43.11%) trials but they rejected to answer any questions about the article. The authors of 11 (3.89%) trials replied that their studies were semi-randomized or a controlled study without randomization. Of these 11 trials, nine were published by a single author. Another one (0.35%) trial published by a single author answered that she could not remember the randomization method and rejected to answer other questions. As for another one (0.35%) trial published by a single author, we contacted the reception desk of department and the staff said the author was going abroad for advanced studies and no one could contact her. Only one (0.35%) trial published by three authors answered that they used the random table but rejected to answer other questions.

Conclusions: Among 283 trials, we could include only one trial for specific randomization method. Rapid review is generated by high-quality evidence, the certainty of which is assessed based on reporting information. We suggest that journal editors follow the Consolidated Standards of Reporting Trials (CONSORT) to report the quality of RCTs. A rigorous RCT should include not just a single person but a collaboration of investigators, a clinical expert, a methodologist, and a statistician. To avoid research waste and useless publications, we should improve the selection criteria, such as by excluding RCTs with a single author.

Patient or healthcare consumer involvement: High quality research can help patient and healthcare consumers to make better decisions.

The structure of research questions in rehabilitation randomized-controlled clinical trials: a methodological study

Arienti C[1], Lazzarini SG[1], Patrini M[1], Pulijak L[2], Pollock A[3], Negrini S[4]

[1] IRCCS Fondazione Don Carlo Gnocchi, Milan, Italy; [2] Center for Evidence-Based Medicine and Health Care, Catholic University of Croatia, Zagreb, Croatia; [3] Nursing, Midwifery and Allied Health Professions Research Unit, Glasgow Caledonian University, Glasgow, UK; [4] Department of Biomedical, Surgical and Dental Sciences, University of Milan "La Statale", Milan, Italy; IRCCS Istituto Ortopedico Galeazzi, Milan, Italy

Background: Randomized controlled clinical trials (RCTs) are considered the gold standard study design to evaluate the effectiveness of interventions in biomedical research. The choice of appropriate study design is informed by a clear research question (RQ). The RQ represents the starting point for research studies to evaluate effectiveness of interventions because it guides the definition of population, interventions, comparators and outcomes (PICO), and this consequently informs the development of the optimal study design to answer the question of interest. It is reported that RQs of rehabilitation RCTs lack clarity and are rarely defined using the PICO format.

Objectives: To assess whether and how the PICO format is described to frame RQs in RCTs focussed on the effectiveness of rehabilitation interventions.

Methods: We conducted a methodological study, systematically exploring the RQs within RCTs of rehabilitation interventions. We searched all the best journals according to the European Society of Physical Rehabilitation and Medicine criteria. Eligibility criteria were: RCTs published between 1 January and 31 December 2019, and addressing a RQ relating to the effectiveness of an intervention in any clinical rehabilitation setting. Two review authors extracted information relating to the RQ, objective or hypothesis, and assessed whether each of the four PICO elements were presented. We used the Cochrane Risk of Bias assessment tool to assess risk of bias in the included studies. We calculated the percentage of trials that clearly stated each PICO element and associated 95% confidence interval (CI). Variables were considered to be statistically significant at alpha = 0.05.

Results: After removal of duplicates, 247 records have been screened; of these, 97 RCTs met the inclusion criteria and were included in the analysis. Preliminary results show that 48% of the studies use "objective" form for the aim and 33% do not use the PICO format. The analysis is still ongoing and the final results will be shown during the Cochrane Colloquium.

Conclusions: RCTs of rehabilitation interventions published in 2019 often fail to adequately report all four PICO elements. In order to support evidence-based rehabilitation, it is important that steps are taken to improve the definition and reporting of PICO elements. To address this, a new reporting checklist for rehabilitation RCTs (RCTRACK checklist) is under development to incorporate an item relating to adequate reporting of RQs.

Patient or healthcare consumer involvement: Not applicable.

The use of systematic reviews and meta-analyses when designing and reporting surgical randomized controlled trials: a cross-sectional analysis

Lu C[1], Ke L[2], Lu T[1], Lan T[3], Guo Y[4], Ge L[1], Shen R[2], Hua Y[2], Yang K[1]

[1] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, China; [2] Department of Liver Surgery, The First Affiliated Hospital, Sun Yat-Sen University, China; [3] Department of Pancreatobiliary Surgery, The First Affiliated Hospital of Sun Yat-Sen University, China; [4] Organ Transplant Center, The First Affiliated Hospital, Sun Yat-Sen University, China

Background: Well-designed and conducted surgical randomized controlled trials (RCTs) are essential to surgery clinical practice, and the prerequisite of planning RCTs should be based on an extensive literature review, especially systematic reviews (SRs) and meta-analysis (MAs). However, knowledge integration before conducting RCTs is not uniformly applied.

Objectives: This cross-sectional study is to evaluate the use of SRs and MAs in conducting surgical RCTs using two high-impact factor surgery journals as data sources, including *Annals of Surgery* (Ann Surg) and the *British Journal of Surgery* (Br J Surg).

Methods: We searched studies published in Ann Surg and Br J Surg journals on 15 January 2020 through the PubMed database from 1 January 2015 to 31 December 2019. The search terms were as follows: "Ann Surg [Journal]" OR "Br J Surg [Journal]". All records were downloaded from PubMed and imported into Endnote X9 software, then two independent authors read each title, abstract, full-text, and selected RCTs. We will extract and cross-check data separately using a standard form: name of first author, year of publication, journal, funding, type of trial, type of intervention, sample size and significance of the study (P values was classified as positive, < 0.05, or negative, ≥ 0.05). For each RCT, we will document the number of SRs cited in the introduction, methods, and discussions by reviewing each publication's reference list for the following key words: meta-analysis, systematic review, and Cochrane Review. Each full-text will be reviewed to determine if the SR can be explicitly cited as justification for conducting the RCT, whether the cited SR can be inferred as forming the basis for initiating the RCT or just be cited in other ways unrelated to the RCT development. All discrepancies will be resolved by consensus. Descriptive data will be presented as frequencies and percentages along with 95% confidence intervals (95% CI). All proportions will be calculated alongside their 95% CIs using the Wilson binomial proportion method. All data analyses will be performed using Excel 2019 (Microsoft, WA, USA) and Stata version 13.0 (StataCorp, College Station, TX, USA).

Results and conclusions: We retrieved 4404 records from the PubMed database and identified 371 RCTs after full-text screening. The data extraction and data analysis are ongoing. Eventually, the results of this study will be submitted to a peer-reviewed journal for publication.

Patient or healthcare consumer involvement: Not applicable.

To share or not to share data: how valid are copious randomized controlled trials?

Bordewijk E[1], Wang R[2], van Wely M[1], Costello M[3], Norman R[4], Teede H[2], Gurrin L[5], Mol B[2], Li W[2]

[1] Universitair Medisch Centrum (UMC), Amsterdam, The Netherlands; [2] Monash University, Australia; [3] The University of New South Wales, Australia; [4] The University of Adelaide, Australia; [5] The University of Melbourne, Australia

Background: Increasingly individual participant data (IPD) is being shared and integrated from randomized controlled trials (RCTs) for systematic reviews and other righteous purposes. Granting open access of data has implications for the promotion of fair and transparent conduct of RCTs, which is crucial when arguing for reproducibility in research. It is, however, still common for authors to choose to withhold IPD, limiting the impact of and confidence in the results of RCTs and systematic reviews based on aggregate data. In our recent IPD meta-analysis evaluating the effectiveness of first-line ovulation induction for polycystic ovary syndrome (PCOS), IPD was only available from 20 RCTs whereas IPD from 34 RCTs was not available. We found that the summary effect sizes of meta-analyses of RCTs not providing IPD were different from those of RCTs that provided IPD. Several aggregate data meta-analyses have been performed on this topic.

Objectives: To understand if RCTs that did not share IPD have lower quality and more methodological issues than those that shared IPD in an IPD meta-analysis evaluating first-line ovulation induction for PCOS.

Methods: We assessed and compared the shared and non-shared IPD RCTs on the following criteria: risk of bias, GRADE approach, adequacy of trial registration; statistical issues (description of statistical methods and reproducibility of univariable statistical analysis); excessive similarity or difference in baseline characteristics that is not compatible with chance (Monte Carlo simulations and Kolmogorov-Smirnov test); and miscellaneous methodological issues.

Results: Overall, the non-shared RCTs had worse performance regarding the assessment of the risk of bias and the GRADE approach when compared to the shared RCTs. Adequate trial registration was found in 33% of the shared IPD RCTs versus 0% in the non-shared RCTs (P = 0.012). In total, 7/17 (41%) shared RCTs and 19/28 (68%) non-shared RCTs had issues with the statistical methods described (P = 0.079). The median (range) of inconsistency rate of univariable statistical results for the outcome(s) was 0 (0 to 0.63) (14 RCTs applicable) in the shared group and 0.44 (0 to 1) (24 RCTs applicable) in the non-shared group (P = 0.0033). The distribution of simulation generated P values from all baseline continuous variables did not significantly violate the expected uniform distribution in the shared group (P = 0.1626), suggesting that these baseline characteristics are likely to be the results of proper randomization. However, it was significantly violated in the non-shared group (P = 4.535*10^-8) (Figure 1).

Conclusions: The IPD meta-analysis on evaluating first-line ovulation induction for PCOS preserves better validity than meta-analyses using aggregate data. The availability of IPD might be a good indicator of the quality and methodological soundness of RCTs when performing systematic reviews.

Patient or healthcare consumer involvement: None.

Figure 1.
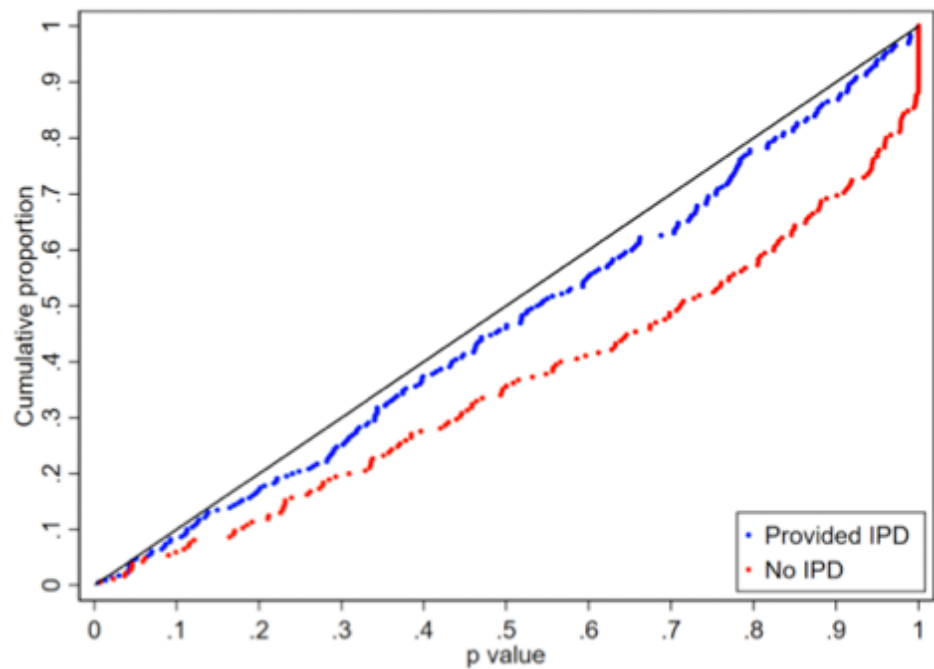
Updating guidelines for reporting systematic reviews and meta-analyses: development of the PRISMA 2020 statement

Page MJ[1], McKenzie JE[1], Bossuyt PM[2], Boutron I[3], Hoffmann T[4], Mulrow CD[5], Shamseer L[6], Moher D[7]

[1] School of Public Health and Preventive Medicine, Monash University, Australia; [2] Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Amsterdam University Medical Centres, University of Amsterdam, The Netherlands; [3] Paris Descartes University, France; [4] Institute for Evidence-Based Healthcare, Faculty of Health Sciences and Medicine, Bond University, Australia; [5] University of Texas Health Science Center at San Antonio, USA; [6] School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Canada; [7] Centre for Journalology and Canadian EQUATOR Centre, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Canada

Background: The PRISMA Statement was published in 2009 and designed to help authors of systematic reviews (SRs) prepare a transparent report of what they did and what they found. Authors and journals have widely adopted PRISMA, as evidenced by its 60,000 citations and 200 journals and SR organizations endorsing its use. A decade old and in need of incorporating advances in SR methodology and terminology and to reflect changes to the publishing landscape, we set out to update the original PRISMA reporting guideline.

Objectives: To describe the processes used to update the PRISMA 2009 statement for reporting SRs and meta-analyses, and summarise changes made to the guideline.

Methods: We conducted a selective review of documents providing reporting guidance for SRs, to generate ideas for how to modify the PRISMA 2009 statement. We invited 220 SR methodologists, authors and journal editors to provide feedback via an online survey on suggested modifications to PRISMA that arose from the selective review. The results of these projects were discussed at a 21-member in-person meeting in September 2018. Following the meeting, we drafted the PRISMA 2020 statement and refined it based on feedback from co-authors and sought examples that best illustrated the items.

Results: The review of 60 documents providing reporting guidance for SRs resulted in a bank of 221 unique reporting items and revealed that all sections of PRISMA 2009 could be modified or supplemented with additional guidance. Of the 110 respondents to the survey, more than 66% recommended keeping six of the PRISMA 2009 checklist items as they were and modifying 15 of the checklist items using the wording suggested by us. Attendees of the in-person meeting supported the revised wording for several items but suggested rewording for others, and further refinements were made over four drafts of the checklist. The PRISMA 2020 statement now consists of updated guidance intended to facilitate transparent reporting of SRs, with or without meta-analysis. Nearly all PRISMA 2009 checklist items were modified in some way (e.g. disaggregated into multiple sub-items to facilitate clarity). The statement also includes new reporting guidance to reflect advances in methods to identify, select, appraise and synthesise studies, and to enhance the reproducibility of SRs.

Conclusions: We anticipate that the PRISMA 2020 statement will benefit researchers, editors, peer reviewers, guideline developers and policy makers involved in conducting or using SRs and meta-analyses. Ultimately, we hope that uptake of the guideline will lead to more transparent reporting of SRs, so facilitating decision-making in the development of evidence-based guidance for clinical practice and health policy.

Patient or healthcare consumer involvement: No patients or healthcare consumers were involved in this research.

Willingness to share individual participant data, and barriers and facilitators to data sharing: a retrospective cohort study and cross-sectional survey

Tan AC[1], Askie LM[1], Hunter KE[1], Barba A[1], Seidler AL[1]

[1] NHMRC Clinical Trials Centre, University of Sydney, Australia

Background: Individual participant data (IPD) meta-analysis is often regarded as the gold standard for evidence synthesis. An IPD approach can substantially improve data quality by offsetting inadequate reporting and allowing more flexible and detailed analyses such as participant-level subgroup analysis (Tierney, Stewart, and Clarke, 2019). The International Committee of Medical Journal Editors declared data sharing an ethical obligation and since 2019 requires all clinical trial registrations to contain a data sharing plan (Taichman et al., 2017). Yet, IPD meta-analyses are often limited by the reluctance of many investigators to share their data. Data sharing statements submitted to clinical trial registries offer a unique resource to explore barriers and facilitators to data sharing (Tierney et al., 2019).

Objectives: Our aim was to determine data sharing willingness reported in recently registered clinical trial records, how this relates to clinical trial characteristics, and principal investigators' attitudes, motivations and barriers to data sharing.

Methods: We conducted a retrospective cohort study and an online cross-sectional survey of all interventional trials registered on the Australian New Zealand Clinical Trial Registry (ANZCTR) from 1 December 2018 to 30 November 2019.

Results: In the retrospective cohort study (n = 1517), data sharing willingness was low (22%) and, as shown in the Figure, was significantly lower for trials with industry involvement (odds ratio (OR) 0.52, 95% CI 0.37 to 0.72) and in non-randomized (OR 0.69, 95% CI 0.52 to 0.91) or phase 1 trials (OR 0.38, 95% CI 0.21 to 0.65, Chi^2 (5) = 19, P = 0.002). Data sharing willingness was not associated with health condition studied, presence of a control group, trial purpose, sample size or primary sponsor country. In the online cross-sectional survey (n = 281, 23% response rate), investigators showed strong support for the concept of data sharing (77%) but a substantially lower intention to share data for their own trials (40%). Major concerns included inappropriate data use by other investigators (97%), obtaining study participant informed consent to share (97%), preventing misinterpretation or misleading secondary analyses (97%) and spending undue time or effort preparing the data for sharing (96%).

Conclusions: There is insufficient willingness to share data, particularly amongst trials with industry involvement. Addressing the identified barriers would support the reuse and replication of research, strengthen the transparency and reliability of trials, and facilitate statistical techniques such as IPD meta-analyses.

Patient or healthcare consumer involvement: Healthcare consumers will be invited to comment on this research project.

Additional file: Trial characteristics

CONFLICT OF INTEREST AND RESEARCH INTEGRITY

Conflict of interest disclosure in systematic reviews of surgical interventions and devices: a cross-sectional survey

Yu J[1], Su G[2], Hirst A[3], Li Y[1]

[1] Chinese Evidence-Based Medicine Center, West China Hospital, Sichuan University, China; [2] School of Preclinical and Forensic Medicine, Sichuan University, China; [3] Nuffield Department of Surgical Sciences, University of Oxford, UK

Background: A conflict of interest (COI) is acknowledged as an important source of bias in the design, conduct and reporting of studies. Surgery, with its wide use of medical devices and surgeons' preference, is a field with high potential for COI.

Objectives: To investigate the frequency and types of conflicts of interest of systematic reviews including RCTs assessing surgical interventions.

Methods: We conducted a methodological survey by searching for systematic reviews (SRs) involving randomized controlled trials of surgical interventions published in 2017 via PubMed. We gathered information regarding general characteristics, source of funding, COI disclosure and the type of COI. We conducted a multiple logistic regression analysis to explore factors associated with the reporting of COI disclosure in SRs.

Results: We identified 163 systematic reviews in 2017, almost half of SRs (55.8%) reported their funding resource and 90.1% reported authors' COI disclosure. Among 147 systematic reviews that provided COI disclosures, 33 (20.2%) declaimed at least one author reporting any type of COI. Cochrane SRs were more likely to report funding source (81.4% versus 51%, P = 0.003) and at least one type of COI than non-Cochrane SRs. SRs that reported funding and significance of primary outcome were associated with better reporting of COI (adjusted odds ratio (OR) 12.6, 95% confidence interval (CI) 1.5 to 102.3; adjusted OR 3.79, 95% CI 1.29 to 11.09).

Conclusion: Although most SRs stated their COI disclosure, the majority of them were no COI disclosure and we found less information regarding non-financial COI. Disclosure should be required before the publication of a systematic review and be better reported regarding the nature of interests. Additionally, more efforts from editors and peer-reviews in medical journals should be made to verify the accuracy and integrity of conflict of interests.

Ethics approval and consent to participate: not applicable.

Conflict of interest policies in practice guideline development: a systematic survey of published guidance

Khabsa J[1], Itani Z[2], Noureldine HA[3], Nonino F[4], Khamis MM[5], Bejjani J[6], Schünemann HJ[7], Akl EA[8]

[1] Clinical Research Institute, American University of Beirut Medical Center, Lebanon; [2] Department of Epidemiology and Population Health, Faculty of Health Sciences, American University of Beirut, Lebanon; [3] Faculty of Medicine, Lebanese American University, Lebanon; [4] IRCCS Institute of Neurological Sciences of Bologna, Italy; [5] Faculty of Medicine, Fayoum University, Egypt; [6] Faculty of Medicine, American University of Beirut, Lebanon; [7] Department of Health Research Methods, Evidence, and Impact (HEI), McMaster University; Department of Medicine, McMaster University, Canada; [8] Department of Internal Medicine, American University of Beirut, Lebanon

Background: Conflict of interest (COI) can influence different steps of the guideline development process, from topic selection to that of guideline dissemination. In 2011, the Institute of Medicine (IOM) published standards for trustworthy clinical practice guidelines, focusing on the importance of COI recognition and management in the guideline development process.

Objectives: To survey guideline development guidance regarding their COI policies, including disclosure, management and reporting.

Methods: We included guidance documents and COI policies related to guideline development. We compiled a comprehensive list of guideline-producing organizations using multiple sources. We searched for the organizations' guidance documents and policies on their websites, PubMed, using Google, and reviewing the references lists of relevant articles. Two review authors assessed the organizations and documents for eligibility, and abstracted data on organizations' characteristics and COI policies, including disclosure of relationships of interest, verification of disclosures, assessment of the risk of COI, management of those conflicts, and their reporting.

Results: We identified a total of 137 guideline-producing organizations, of which 96 had a COI policy, either in the form of a non-brief section in their guideline development guidance (64%), or as separate dedicated document (36%). The median date of last update of the COI policies was 2017 (interquartile range 2015 to 2018). A disclosure form was made publicly available by 50% of organizations. The vast majority of organizations required panelists to disclose their relationships of interest (90%), followed by executive committee, oversight committee, or board (38%), peer reviewers (40%), and outside experts, consultants and presenters (26%). Only 20% of organizations mentioned disclosure by systematic review authors and/or methodologists. Assessment of risk of COI was mentioned by 50% of organizations. However, only 25% of the total sample provided details on the criteria considered and/or the method of doing so. Most common management strategies were exclusion from the group (63%), from discussion (40%), and from voting (40%). Very few organizations addressed COI management issues related to the systematic review (SR) process, which included exclusion of conflicted individuals from the SR process, restrictions during participation in the SR process (e.g. exclusion from performing evidence selection, verification by another group member, recusal from assessment of the quality of the evidence). The majority of organizations mention publicly reporting on their COI (73%). Very few explicitly mention reporting on amount of COI (3%), and 25% report on the method of evaluation and management of COI.

Conclusions: There is large variability in how guideline-producing organizations handle COI. While most address requirements related to panelists and management issues related to developing recommendations, very few specifically address the SR team and the SR process.

Patient or healthcare consumer involvement: None.

Data integrity of 35 randomized controlled trials in women's health

Bordewijk EM[1], Wang R[2], Askie LM[3], Gurrin LC[4], Thornton JG[5], van Wely M[6], Li W[2], Mol BW[2]

[1] Centre for Reproductive Medicine, Amsterdam UMC, Netherlands; [2] Department of Obstetrics and Gynecology, Monash University, Australia; [3] National Health and Medical Research Centre (NHMRC) Clinical Trials Centre, University of Sydney, Australia; [4] Centre for Epidemiology and Biostatistics, School of Population and Global Health, The University of Melbourne, Australia; [5] Faculty of Medicine and Health Sciences, University of Nottingham, UK; [6] Centre for Reproductive Medicine, Amsterdam UMC, The Netherlands

Background: While updating a systematic review for the Cochrane Database for systematic reviews on the topic of ovulation induction, we observed unusual similarities in a number of randomized controlled trials (RCTs) published by two authors from the same institute in the same disease spectrum in a short period of time.

Objectives: We undertook a focused analysis of the data integrity of all RCTs published by the two authors.

Methods: We searched the PubMed database for authors 'Badawy' or 'Abu Hashim' using the affiliation 'Mansoura' restricting to RCTs. We made pairwise comparisons to find identical or similar values in baseline characteristics and outcome tables between trials. We also assessed whether baseline characteristics were compatible with chance, using Monte Carlo simulations and Kolmogorov-Smirnov test.

Results: For 35 trials published between September 2006 and January 2016, we found a large number of similarities in both the baseline characteristics and outcomes of 26. Analysis of the baseline characteristics of the trials indicated that their distribution was unlikely to be the result of proper randomization.

Conclusions: Our analyses suggest serious data integrity issues in published RCTs from these authors. They have been cited 13 times in meta-analyses within Cochrane Reviews. The procedures demonstrated in this paper may help to assess data integrity in future attempts to verify the authenticity of published RCTs.

Patient or healthcare consumer involvement: Not applicable.

How to deal with conflict of interest in guideline development: a systematic review of guidance documents

Wang Z[1], Shi Q[2], Zhou Q[2], Xun Y[3], Zhao S[4], Chen Y[5]

[1] China; [2] The First School of Clinical Medicine, Lanzhou University, Lanzhou, China; [3] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China; [4] School of Public Health, Lanzhou University, Lanzhou, China; [5] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, Lanzhou, China

Background: The development of clinical practice guidelines (CPGs) must be fair and transparent. Conflicts of interest (COI) are one of the most important factors affecting the independence and reliability of guidelines. The existence of COI may lead to the exaggeration of ineffective interventions and the concealment of effective interventions, which is an important source of potential bias in the development of guidelines. Many guidance documents on the development of CPGs have referred to how to declare and manage COI. But there has been no detailed description of issues such as whether the definition or sort of COIs are different and whether the method of interest management is same. This point is directly related to the quality, credibility and authority of the guidelines developed by referring to different guidance document makers.

Objectives: To assess how guidance documents for developing CPGs address COI during the development of clinical guidelines.

Methods: We collected the guidance documents for developing clinical practice guidelines by searching PubMed and Google search engine (Alphabet). We also searched the reference lists of all eligible documents and relevant literatures for additional materials not captured by the aforementioned searches. We included documents that provided guidance on the entire development process of practice guidelines and mentioned COI. We excluded documents that were written by individuals, were outdated versions that had been subsequently updated, or were focused on specific aspects of guideline development (such as updating; systematic reviews, or the GRADE process). Two researchers independently screened the records and extracted data. We extracted the title of guidance documents, publication date, development organization, etc. as basic information. For the information on COI, the primary framework was based on the form of the WHO Handbook for Guideline Development (2nd edition), and the information outside the primary framework was supplemented in an iterative way. The information related to COI in the current guidance documents were finally comprehensively summarized and graded according to the frequency of the report.

Results: We retrieved 89 guidance documents. Of these, 66 reported COI. The detailed information of COIs will be presented at the meeting.

Conclusions: The results will be presented at the meeting.

Patient or healthcare consumer involvement: None

Industry funding of patient and health consumer groups: systematic review with meta-analysis

Fabbri A[1], Parker L[1], Colombo C[2], Mosconi P[2], Barbara G[3], Frattaruolo MP[3], Lau E[1], Kroeger CM[1], Lunny C[4], Salzwedel DM[4], Mintzes B[1]

[1] Charles Perkins Centre and School of Pharmacy, Faculty of Medicine and Health, The University of Sydney, Camperdown, Sydney, Australia; [2] Mario Negri Institute for Pharmacological Research IRCCS, Italy; [3] Gynaecology Unit, IRCCS Ca' Granda, Ospedale Maggiore Policlinico, Milan, Italy; [4] Cochrane Hypertension Review Group, Therapeutics Initiative, Department of Anesthesiology, Pharmacology and Therapeutics, Faculty of Medicine, University of British Columbia, Vancouver, Canada

Background: Patient and health consumer groups often rely on the financial support of the pharmaceutical and medical device industries. Considering their important role in health policy, concerns have been raised because of conflicts of interest and potential limitations to their independence.

Objectives: To synthesise studies on prevalence of industry funding, transparency, and positions of patient and health consumer groups.

Methods: We searched MEDLINE, Embase, Web of Science, Scopus and Google Scholar (from inception to January 2018). We included observational studies reporting at least one of the outcomes: prevalence: percentage of groups that accept industry funding, percentage of groups' funding from industry, and number of funders per group; transparency: proportion of industry funded groups reporting the source of funding; positions: association between industry funding and organizational positions on health and policy issues. We carried out duplicate independent data extraction and assessed study quality using an amended version of the Checklist for Prevalence Studies developed by the Joanna Briggs Institute. For meta-analyses of prevalence, we used a DerSimonian-Laird estimate of single proportions with Freeman-Tukey arcsine transformation. We used GRADE to assess the quality of the evidence per outcome.

Results: We included 26 cross-sectional studies. Most were conducted in the United States and Europe. We meta-analysed prevalence of industry funding, proportion of industry funded groups that disclosed funding, and prevalence of group policies on sponsorship. We cannot present summary estimates for industry funding prevalence and organizational policies because of the high level of unexplained heterogeneity. Seventeen (65%) studies were of high quality and 9 (35%) of low quality. The prevalence of industry funding ranged from 20% (12/61) to 83% (86/104) (15 studies). Among groups having received industry funding, 27% (95% confidence interval (CI) 24% to 31%) disclosed the information on their website. In submissions to governmental consultations, disclosure varied from 0% to 91% (two studies). From 2% to 64% of groups had policies on sponsorship. Industry funded groups tend to hold positions in line with sponsors' interests (four studies).

Conclusions: Patient groups often receive pharmaceutical and medical device industry funding; few disclose them on their websites. Shifts in disclosure may have occurred, as the relevant studies were published between 2008 and 2012. There was also little research in middle- or low-income countries. Few groups had formal sponsorship policies. Our findings of prevalent funding but limited governance raise concerns about risks of commercial biases in the representation of patients' interests. To limit these risks, mechanisms for increased public financing could be explored.

Patient or healthcare consumer involvement: Two authors have been involved for years with women's health and consumer groups. One representative of a Canadian patient group was involved in commenting on the findings.

Reporting of drug trial funding sources and author financial conflicts of interest in Cochrane and non-Cochrane meta-analyses

Turner K[1], Carboni-Jiménez A[2], Benea C[3], Elder K[3], Levis B[4], Boruff J[2], Roseman M[5], Bero L[6], Lexchin J[7], Turner E[8], Benedetti A[2], Thombs B[9]

[1] University of Ottawa, Canada; [2] McGill University, Canada; [3] Lady Davis Institute for Medical Research, Canada; [4] McGill University and Keele University, Canada and UK; [5] Kingsway Medical Centre Family Health Organization, Canada; [6] University of Sydney, Australia; [7] York University, Canada; [8] Oregon Health and Science University, USA; [9] Lady Davis Institute for Medical Research, Jewish General Hospital and McGill University, Canada

Background: A previous study found that only 30% of Cochrane Reviews of drug trials published in 2010 reported the funding source of some or all included randomized controlled trials (RCTs), 7% reported trial author-industry financial ties, and 7% reported trial author-industry employment. It is not known if reporting has improved since Cochrane implemented a policy to require reporting in 2012 or how Cochrane meta-analyses compare to non-Cochrane meta-analyses.

Objectives: 1) To investigate the extent to which recently published meta-analyses report trial funding, author-industry financial ties, and author-industry employment from included RCTs, comparing Cochrane and non-Cochrane meta-analyses. 2) To examine characteristics of meta-analyses independently associated with reporting funding sources of included RCTs. 3) To compare reporting among recently published Cochrane meta-analyses to Cochrane Reviews published in 2010.

Methods: We searched PubMed on 19 October 2018 and selected the 250 most recent meta-analyses listed in PubMed that included a documented search of at least one database, statistically combined results from ≥ 2 RCTs, and evaluated the effects of a drug or class of drugs.

Results: Ninety of 107 (84%) Cochrane meta-analyses reported funding sources for some or all included trials compared with 21 of 143 (15%) non-Cochrane meta-analyses, a difference of 69% (95% confidence interval (CI) 59% to 77%). Percent reporting was also higher for Cochrane meta-analyses compared with non-Cochrane meta-analyses for trial author-industry financial ties (44% versus 1%; 95% CI for difference, 33% to 52%) and employment (17% versus 1%; 95% CI for difference, 9% to 24%). In multivariable analysis, compared with Cochrane meta-analyses, the odds ratio for reporting trial funding was ≤ 0.11 for all other journal category and impact factor combinations. Compared with Cochrane Reviews from 2010, reporting of funding sources of included RCTs among recently published Cochrane meta-analyses improved by 54% (95% CI 42% to 63%), and reporting of trial author-industry financial ties and employment improved by 37% (95% CI 26% to 47%) and 10% (95% CI 2% to 19%).

Conclusions: Reporting of trial funding sources, trial author-industry financial ties, and trial author-industry employment in Cochrane meta-analyses has improved substantially since 2010 and is much higher than in non-Cochrane meta-analyses. Reporting requirements similar to those of Cochrane should be implemented and enforced by other journals.

Additional file: [Figure](#)

A Cochrane-driven scientific initiation

Fontes LE[1], Nogueira C[1], Pessata M[1], Altoé K[1], Franklin Leitão E[1], Riera R[1]

[1] Cochrane Brazil Rio de Janeiro, Brazil

Background: To encourage and support new incomers is one of Cochrane's priorities according to its strategic plan. The Cochrane Training website provides several learning opportunities, with emphasis (but not exclusively) on reading, preparing, conducting, and publishing systematic reviews. There are tools and online training guidance that might build capacity amongst healthcare professionals, but it is also an excellent source of learning for undergraduate students. Scientific Initiation (SI) is a program directed to undergraduate students, in order to connect them to study groups and lines of research. It intends to stimulate students, tutored by an experienced professor, to learn techniques and scientific methods, as well as to encourage scientific thinking and creativity. Petropolis' medical school (FMP/FASE), a private institution in the state of Rio de Janeiro, Brazil, doesn't have a formal discipline in Evidence-Based Healthcare. In spite of that, there are some initiatives being conducted successfully to approximate students and high-quality research, such as the launch of Cochrane Brasil Rio de Janeiro (CBRJ) Affiliate Center, part of Brazilian Cochrane Network, hosted by FMP/FASE, and a partnership with the Centre For Evidence-Based Medicine from the University of Oxford.

Objectives: To describe the experience of a Cochrane-driven scientific initiation to undergraduates at FMP/FASE.

Methods: Descriptive study.

Results: CBRJ conducts a one-year program, with face-to-face and virtual learning activities. The undergraduates are encouraged to apply for four positions available. The selection process involves curriculum analysis and a face-to-face interview. The SI engages the students to become Cochrane members and to get into action in different fields such as knowledge translation, research, organizational skills, and education. Some activities carried out during the program are:

1) training sessions in evidence-based principles using free materials from the Cochrane Training website;
2) attendance in training workshops and webinars to build capacity as potential new review authors;
3) translation of Cochrane materials into Portuguese, contributing to the spreading of relevant evidence and educational materials;
4) active participation in CBRJ organizational tasks, promoting a sense of community and involving them to the Cochrane structure; and
5) promotion of Cochrane evidence among other students and healthcare professionals locally.

Some of the Cochrane resources used are Cochrane Evidence Essentials, Cochrane Journal Club, and Students for Best Evidence blogs.

Conclusions: A scientific initiation underpinned in free Cochrane materials is feasible and effective in delivering high-quality training to undergraduates, encompassing different fields such as knowledge translation, education, and organizational skills.

Patient or healthcare consumer involvement: We believe that offering a proper introduction to health evidence, and how to use it to make informed health choices, would improve patient outcomes in the future.

EDUCATION AND TRAINING

A measurement properties systematic review of instruments for assessing undergraduate nursing students' evidence-based practice knowledge, attitudes, and skills

Cardoso D[1], Riitano D[2], Stephenson M[2], Santos E[1], Cardoso ML[3], Rodrigues M[1], Apóstolo J[1]

[1] Portugal Centre for Evidence-Based Practice: A JBI Centre of Excellence, Health Sciences Research Unit: Nursing, Nursing School of Coimbra, Portugal; [2] Joanna Briggs Institute, School of Translational Health Science, University of Adelaide, Australia; [3] Health Sciences Research Unit: Nursing, Nursing School of Coimbra, Portugal

Background: Using evidence-based practice (EBP) in clinical practice can have multiple benefits, such as improvement of patient outcomes, reduced healthcare costs, and increased quality of care. As education is a strategy to promote the EBP use in clinical practice, undergraduate nursing curricula should be based on EBP principles to train future nurses on EBP use. Therefore, good quality instruments are needed to measure undergraduate nursing students' EBP attitudes, knowledge, and skills and, consequently, the impact of the EBP educational programs in those outcomes.

Objectives: To assess the measurement properties of the available instruments for measuring undergraduate nursing students' EBP knowledge, attitudes, and skills.

Methods: We conducted this systematic review in accordance with an a priori published protocol and registered it in PROSPERO (CRD42017074920). We performed a search strategy in three steps from 1996 to July 2018 to find studies in Portuguese, English and Spanish. Two independent review authors analyzed the title/abstract and the full-text against the inclusion criteria. Using the COSMIN Checklist, two independent review authors performed the critical appraisal and a third review author analyzed the disagreements. We validated details of general characteristics of the instruments; characteristics of the study populations in which the measurement properties were validated; and extracted the results of the measurement properties from each included paper. We created overview tables to synthesize the data.

Results: We found 1942 records. From these, 11 papers included reporting data on the following five instruments: Evidence Based Practice Questionnaire; Student Evidence-based Practice Questionnaire; Evidence-based Practice Knowledge Assessment in Nursing; Evidence Based Practice Evaluation Competence Questionnaire; and Evidence-based practice profile questionnaire. No study assessed measurement error and criterion validity. Only the internal consistency was assessed by all studies with very good methodological quality. Responsiveness was assessed only by one study but with inadequate methodological quality. The methodological quality of the structural validity varied across the eight studies from inadequate to very good. The methodological quality of the remaining measurement properties assessed in the included studies varied from inadequate to adequate.

Conclusions: We found five instruments to measure undergraduate nursing students' EBP knowledge, attitudes, and skills. However, only two measured the three constructs of interest. Due to the low number of studies per instrument version (e.g. language and context), it was not possible to perform a best-evidence synthesis.

Patient or healthcare consumer involvement: We sought the input of nurses and nursing educators to guide the design of the systematic review.

A systematic review of quality improvement education for medical students: Kirkpatrick's learning model

Brannan G[1], Rao A[2], Brooks E[2], Mills L[3], Harnden A[4], Alderson T[5], McQuiston J[4]

[1] Department of Graduate Medical Education, McLaren Macomb Hospital, Mount Clemens, Michigan; GDB Research and Statistical Consulting, Athens, Ohio, USA; [2] Department of Internal Medicine, McLaren Macomb Hospital, Mount Clemens, Michigan, USA; [3] Department of Graduate Medical Education, McLaren Macomb Hospital, Mount Clemens, Michigan, USA; [4] Department of General Surgery, McLaren Macomb Hospital, Mount Clemens, Michigan, USA; [5] Department of Obstetrics and Gynecology, McLaren Macomb Hospital, Mount Clemens, Michigan, USA

Background: The 2001 Institute of Medicine's report, Crossing the Quality Chasm, revealed a fractured American Healthcare System and a critical need for quality improvement (QI) training. Since that time, medical education accrediting bodies such as AAMC and ACGME have incorporated the findings and implemented QI into the training curriculum at both the level of medical student and residents, respectively. In addition, residents are required to engage in quality improvement activities during their training. Earlier studies indicated a general recognition of the need and benefits of QI training. Beyond meeting accreditation requirements, it is also critical to determine the impact of a curriculum. The Kirkpatrick model has been used in medical education and other fields to assess the learner and the impact of a curriculum. It is based on outcomes in the context of four areas: participation, modification, behavioral change, and benefit to organization or patients.

Objective: To determine the QI curriculum learning impact using the Kirkpatrick model. This inquiry was part of a comprehensive systematic review to determine current training and curriculum opportunities and challenges in teaching quality improvement to medical students.

Methods: We identified publications in PubMed, Embase, and SCOPUS from 1 January 2009 to 31 December 2018 using a structured search strategy. We followed the PRISMA guideline. For comparison, we identified fundamental components in each article including study population, intervention performed, educational QI component, major findings, and learning outcomes. We used the Kirkpatrick model to determine trainee learning outcomes: impact on learners' satisfaction, changes in attitudes, knowledge and skills, changes in learners' behavior, organizational changes, and patient benefits. We generated simple descriptive statistics such as frequency and percentage to summarize the results.

Results: A total of 29 studies were identified from 3889 peer-reviewed abstracts. Results indicated that Kirkpatrick's Model level 2b: modification of knowledge/skill (65.5%) and level 4a: change in organizational practice (69%) were found to be most prevalent across the studies. Compared to previous systematic reviews on QI training for medical students, there was a significant increase in studies demonstrating Kirkpatrick level 4b: patient benefits (27.6%).

Conclusions: The trends are consistent with medical school QI training becoming more effective and shifting from an individualized effect to a larger organizational effect. Patient benefits were also observed in a significant number of studies.

Patient or healthcare consumer involvement: The patients' needs and interest were represented in the studies reviewed in different ways. Trainees worked on problems critical to patients, used real patient data for audit, and learned to work in inter-professional teams to effectively improve patient care.

AMSTAR2 or ROBIS to teach evidence synthesis?

Lee S[1]

[1] School of Pharmacy, Monash University Malaysia, Malaysia

Background: Systematic reviews (SRs) are an important source of evidence that are used to inform decision making. As such, it is important that healthcare students are taught how to interpret, and judge the methodological quality of systematic reviews. Several tools currently exist to appraise the quality and risk in SRs, but their evaluation and ease of use by students have never been examined.

Objectives: To describe the experience of using both AMSTAR2 and ROBIS tool for teaching of undergraduate pharmacy students

Methods: Second year pharmacy students were randomly allocated two SRs for evaluation. Students were required to evaluate one review using AMSTAR2 while the second review was evaluated using ROBIS. We recorded the time to complete each tool and also students' preference and perception of the ease of use for each tool.

Results: A total of 100 students completed and rated five SRs using AMSTAR2 and ROBIS. Students took a mean of 36.2 + 13.9 minutes to complete the AMSTAR2 compared to 23.5 + 13.6 minutes using the ROBIS tool. All of the students reportedly preferred to use AMSTAR2 compared to ROBIS in grading SRs. Students reported that they felt the ROBIS tool was very subjective especially since it was their first experience in evidence synthesis. As such, students preferred AMSTAR2 due to its simple and easy to administer with clear guiding questions for each domain.

Conclusions: The AMSTAR2 tool appears to be a valid and simple tool for teaching pharmacy students to grade SRs. Some of the signalling questions in ROBIS were judged to be very difficult to assess and understand

Patient or healthcare consumer involvement: No patient or healthcare consumer involvement was planned due to the methodological character of the study.

An education and skills-building program to deliver competent, confident, and efficient evidence-users

O'Mathúna D[1], Gallagher-Ford L[1]

[1] Helene Fuld Health Trust National Institute for Evidence-Based Practice in Nursing and Healthcare, The Ohio State University, USA

Background: Evidence-based decision-making/practice (EBDM/P) is a systematic problem-solving approach in healthcare that: improves quality and population health outcomes by approximately 30%, reduces costs, and empowers clinicians to fully engage in their roles, all of which align precisely with the concept of rapid-learning health systems. EBDM/P requires clinicians who can use evidence like Cochrane Reviews effectively and regularly. Researchers at our Institute and elsewhere have found discrepancies between clinicians' beliefs about evidence-based practice (EBP) and their EBP knowledge and skills. The situation falls far short of the 2003 Institute of Medicine goal that 90% of healthcare decisions be evidence-based by 2020. To overcome barriers to EBP, our Institute developed a 5-day training program called the EBP Immersion which we will describe in this presentation. Since 2012 this program has been delivered to nearly 3000 clinicians, and its educational outcomes have been evaluated. We will present the results here also.

Objectives: To describe the 5-day intensive EBP education and skills-building training program. To provide evidence on how our EBP Immersion develops competent, confident, and efficient evidence-users.

Description: This oral presentation will begin with an overview of the EBDM/P education and skills-building program developed at our Institute: the EBP Immersion. This has been delivered up to 20 times per year to cohorts of about 25 people, mostly healthcare clinicians. Immersions have been conducted around the US and overseas to nearly 3000 attendees. The presentation will provide an overview of the Immersion content, demonstrate the tools and templates used, and explain the teaching/learning strategies employed. Our program includes real-time expert mentoring, and strategies to support follow-up EBP projects at participants home institutions. Research accompanies the program to evaluate its effectiveness in producing competent, efficient, enthusiastic, and empowered evidence-users across disciplines and the care continuum. The workshop will present research findings demonstrating the program's short-term and sustained effectiveness on EBP attributes (beliefs, knowledge, competence, and implementation). We will present exemplars of how this program has underpinned significant practice transformations for individuals and healthcare organizations.

Conclusion: Participants will come away from this presentation with a clear understanding of the varied, innovative, and effective components of this unique program and how it is helping to make a difference across the healthcare continuum.

Patient or healthcare consumer involvement: Patients and consumers have not participated in this program to date, but efforts to involve them have already been implemented. Indirectly, they should benefit from having healthcare practitioners who are competent, confident, and efficient evidence-users.

Capacity building for evidence-based child health through digital platforms

Singh M[1], Singh M[1], Vorukolu S[1], Pradhan P[1], Singh P[1], Jaiswal N[1], Chauhan A[1]

[1] Postgraduate Institute of Medical Education and Research, Chandigarh, India

Background: Indian healthcare is experiencing a paradigm change and there is an emphasis on evidence-based medicine. Increased efforts are being made to build capacity amongst healthcare professionals and various modalities have been explored to reach healthcare professionals in remote areas.

Objectives: To build capacity in evidence-based child through the use of ECHO model and online courses.

Methods: The initiation of online sessions on evidence-based medicine conducted by partnering with Extension of Community Healthcare Outcomes to connect with medical colleges and medical research units in the North eastern region of the country. Sessions on evidence-based medicine were delivered via Zoom software under the project ECHO. These sessions were to develop the knowledge and expertise of the clinicians in the remote areas of north eastern region in the field of evidence-based medicine. The participants were subject to pre- and post- questionnaires to assess their knowledge, awareness and practices (KAP) in evidence-based medicine.

Results: The current online course had 49 participants enrolled for a period of one year. The scientists involved in the delivery of online lectures were trained through 2-day workshops in the use of ECHO and troubleshooting of any problems. During one course, 20% of the participants wrote protocols and submitted them for Prospero registration. Two systematic reviews have been completed and submitted for publication. There was significant increase in the KAP as depicted by the analysis of the pre- (mean score 11.71 ± 0.62) and post- questionnaires (mean score 15.64 ± 0.47).

Conclusions: This model has potential to be replicated all over the country to build capacity and improve medical practice.

Capacity building in evidence-based child health: six years' experience

Singh M[1], Jaiswal N[1], Chauhan A[1], Pradhan P[1], Singh M[1], Agarwal A[1], Mathew J[1], Malhotra S[1], Shafiq N[1], Dutta S[1]

[1] Postgraduate Institute of Medical Education and Research, India

Background: India has seen a paradigm shift in the process of guidelines and policy formulations. More guidelines are now making use of the evidence than relying on expert opinions and consensus alone. While the change was happening, the Indian Council of Medical Research, New Delhi commissioned a Centre for Advanced Research in Evidence-Based Child Health. The centre was based at Department of Pediatrics, Postgraduate Institute of Medical Education and Research, Chandigarh.

Objectives: To build capacity, conduct SRs and promote the practice and training of evidence-based health care in children

Methods: Short courses on 'How to practice and How to teach evidence-based child health' and workshops on "Protocol development and Review completion" were conducted at various sites in North and North-eastern parts of the country. The short courses and workshops followed a standard pre-designed module that included pre-tests and post-tests with similar questions during each course to evaluate the knowledge of evidence-based child health. Short courses included lectures followed by small group interactive sessions on critical appraisal of randomized controlled trials, diagnostic test accuracy and conduct of SRs. The short courses and workshops were free for the participants

Results: The phase 1 of the Centre for Advanced Research was from March 2012 to March 2018. During the six-year period 50 workshops and short courses were conducted at various medical colleges in North India and North East India. Over 1500 healthcare professionals including clinicians, researchers, paramedics and nurses were trained and sensitized to the principles of evidence-based medicine and conduct of systematic reviews. We compared the pre-test and post-test scores and post-test scores for the participants were always high showing significant gain in knowledge. The participants trained through these short courses and workshops have been a part of 40 systematic reviews, of which 21 got published in peer reviewed national and international journals. Course participants have also been part of over 24 high-quality systematic reviews including Cochrane Reviews, of which 11 have been published in peer reviewed national and international journals. The systematic reviews were also acknowledged by National Technical Advisory Group for Immunization (NTAGI) and were discussed in the policy meetings.

Conclusions: The six years of the Centre for Advanced Research has provided significant contribution towards capacity building of healthcare professionals in evidence-based child health and also led to the inclination towards conducting and publishing systematic reviews. The Centre for Advanced Research in Evidence-Based Child Health has now started its second phase with additional responsibilities.

Cochrane Knowledge Translation virtual mentoring pilot scheme: what we learnt and what next?

Head K[1], de Silva D[2], Watts C[3]

[1] Cochrane Central Executive Team, France; [2] The Evidence Centre, UK; [3] Cochrane Central Executive Team, UK

Background: Cochrane is testing mentoring to help build skills and capacity for knowledge translation. Mentoring is defined as "a professional relationship in which an experienced person (the mentor) assists another (the mentee) in developing specific skills and knowledge". Such programmes are common in academic institutions and have been linked to enhanced productivity and self-confidence, but very little is known about whether mentoring works across a global community to support knowledge translation.

Objectives: Cochrane set up a 'virtual' knowledge translation mentoring pilot to test whether mentoring was feasible across the global Cochrane community and whether it would help participants develop confidence in knowledge translation.

Methods: The Cochrane Knowledge Translation team defined eligibility criteria for mentors and mentees, advertised and recruited participants, matched pairs of mentors and mentees and provided documentation and training. Over the six-month period between September 2019 and March 2020, 12 people undertaking knowledge translation projects in Cochrane (mentees) were mentored by people who had experience of knowledge translation (mentors). It was planned that the 'pairs' would meet online for about 60 minutes per month for six months. The pilot was evaluated independently using a before and after design with surveys and interviews.

Results: This session will present the results for the implementation and outcomes from the pilot. In terms of the implementation, our initial findings included that there was high demand for the scheme and that applying online worked well. We have learnt that good matching and preparatory training is important and requires resources to implement. Data on mentor and mentee outcomes is still being analysed. Early results indicate that those who took part reported increased knowledge and confidence. We are aware of some knowledge translation projects where mentoring has helped to bring tangible improvements. We will present how the results from the pilot scheme will support Cochrane's future mentoring schemes both within knowledge translation and in other areas of the organization.

Patient or healthcare consumer involvement: Although there was no direct patient and healthcare involvement in the mentoring scheme, the projects on which mentor and mentee pairs worked were knowledge translation projects. These projects aim to support the use of Cochrane evidence in our target audiences which include patients and consumers. Future mentoring schemes may use the information gathered from this pilot to support mentoring in different audiences, which could include patients and healthcare consumers, potentially as both mentors and mentees.

Cochrane UK & Ireland Trainees Advisory Group journal club – using Twitter to engage trainees in evidence-based medicine

Gould R[1], Carter E[1]

[1] Cochrane UK, UK

Background: Founded in 2016, and supported by Cochrane UK, the Cochrane UK and Ireland Trainees Advisory Group (CUKI-TAG) aims to improve the engagement of medical and dentistry trainees with evidence-based medicine, in particular the work of Cochrane. Since 2018, CUKI-TAG has used Twitter to run a number of online journal clubs on a variety of topics. The journal club date and research paper to be discussed are advertised in advance on social media. Questions to be used during the journal club, to help stimulate discussion on the paper, are prepared prior to the event and reviewed by the Cochrane UK team. Journal club participants join the event by using #cochranetrainees, or following @CochraneUK, on Twitter.

Objectives: To evaluate the impact of CUKI-TAG Twitter journal clubs. To discuss the challenges and future opportunities for the CUKI-TAG Twitter journal club.

Methods: We used Twitter Analytics to evaluate the number of participants, number of impressions (number of people who could potentially see a tweet) and engagements (number of people who interact with a tweet e.g. reply, click on a link within tweet). We sought feedback from CUKI-TAG committee members via an online survey tool and during committee meetings. We also received ad-hoc qualitative feedback from participants.

Results: Six Twitter journal clubs have taken place since 2018. The mean number of participants was 31, with a mean of 125 tweets and 1.66 million impressions for each journal club. Feedback from the CUKI-TAG committee identified paper selection, finding topics applicable to a range of specialities, making trainees aware of the journal club and running the journal club at a manageable pace as the main challenges.

Conclusions: The CUKI-TAG Twitter journal club appears to be an effective way to reach a large number of participants. Going forwards we plan to run journal clubs more regularly (monthly/ bi-monthly) and will try to broaden engagement by selecting topics relevant across multiple specialties and advertising the journal club through organizations that come into regular contact with trainees such as training bodies. Additionally, we would like to help trainees develop skills in other areas, such as reflective practice and knowledge implementation, by generating discussions, for example, on how trainees could use evidence-based medicine to change their own practice and the practice of the team or department they work within. Additionally, the Twitter journal club is low cost and offers the opportunity to engage trainees rapidly and globally. This may be of particular value in low income countries where there may not be ready access to evidence-based medicine training.

Patient or healthcare consumer involvement: No patient or healthcare consumer involvement.

Creating a community of practice for dissemination in Cochrane

Lagosky S[1]

[1] Cochrane Central Executive Team, Germany

Background: For Cochrane evidence to be used to inform healthcare decision-making, it is important that we share the findings of the reviews with our target audiences in formats and languages that they can understand and use. We do this by creating dissemination products (such as blogshots, podcasts and Plain Language Summaries) which provide a tailored summary of the review findings, appropriate to the needs of the target audience. In 2019, as part of Cochrane's Knowledge Translation strategy, we developed a 'Dissemination Checklist' which aims to help producers of dissemination products improve the quality of the products. To implement the checklist we ran a training workshop in January 2020 for key people in Cochrane who regularly disseminate Cochrane Reviews. The training was the first opportunity to bring together a small group of people from within Cochrane who are committed to delivering high quality dissemination. We invited all those involved to join a community of practice to try to promote the benefits of peer-to-peer learning and support.

Objective: The aim of the community of practice is to create a space and convene conversations where people interested in (and currently practicing) dissemination can connect to:

1) share information on their work and experiences doing dissemination;
2) learn ideas and solutions others doing similar work through peer-to-peer learning; and
3) increase collaboration on projects and reduce duplication of effort across a wide geographic spread.

This presentation will share our experiences in developing and running the community of practice. We will present the challenges, learning and successes from the project.

Project Activities: The community of practice interacts largely via Cochrane's Slack channel, through informational emails sent centrally, ad-hoc group web conferences and through face-to-face meetings at Cochrane events. The community is managed by Cochrane Central Executive Team staff, but is a space where members have the flexibility to interact with the community in a way that serves them in their varied roles throughout the organization.

You should attend this presentation if you:

- are interested in learning more about what a Community of Practice looks like at Cochrane
- would like to hear about our experiences on the benefits and challenges of sustaining a Community of Practice; or
- are curious to know how this Community of Practice is influencing dissemination at Cochrane

Patient or healthcare consumer involvement: While the act of disseminating evidence, and specifically the dissemination checklist, should help improve the way we share information with consumers, this particular community of practice is meant for staff who do dissemination as part of their job. However, questions about consumer involvement emerge in this community, and in such cases the consumer involvement team is consulted and involved.

Cultural adaptation of interventions to reduce pesticide exposure among farmworkers: a systematic review

Karimi-Shahanjarini A[1], Afshari M[1], Besharati F[2]

[1] Hamadan University of Medical Sciences, Iran; [2] Guilan University of Medical Sciences, Iran

Background: Pesticide exposure among farmers and their families is a significant health concern worldwide. The agricultural system is a dynamic and complex system in which multilevel factors from individual preferences and motives of farmers to socio-cultural structures have a unique contribution to farmers' behaviors. It seems that the success of health interventions in identifying and integrating cultural differences in farmers could inform the interventions' development, implementation, and evaluation. However, little is known about the effectiveness of cultural adaptations in such interventions.

Objectives: To systematically assess the cultural adaptation strategies in interventions addressing the reduction of pesticide exposure in farmworkers.

Methods: We searched PubMed, Embase, ISI Web of Science, Scopus, NIOSHTIC, Agricola, Agris, as well as the reference lists of identified articles for randomized controlled trials (RCTs) published in English from 2000 to 2019. Two independent review authors assessed trial quality and extracted data. Disagreements were resolved by discussion or referring to a third review author.

Results: A total of 7 trials (8 papers) involving 1371 participants met the review's inclusion criteria. The majority of these studies were conducted in the USA and had a low-quality score. All except one study employed cultural adaptation strategies to varying degrees. Socio-cultural strategies were widely used in the interventions. Four studies were tailored for subgroups or individuals and reported improvements in the outcomes. It seems that cultural adaptations were related to intervention effectiveness.

Conclusions: Our systematic examination revealed that cultural adaptation had a moderate effect on the reduction of farmworkers' pesticide exposure. To improve health outcomes, it is necessary to employ a deep level of cultural adaptation in future investigations.

Editorial processes: what exactly happens behind the scenes?

Dooley L[1], Jess C[2], Pardo Pardo J[3], Elstub L[3], Erskine A[2], Kellie F[2], Kelly R[2], Ovelman C[4], Stewart M[2], Paulsen L[5]

[1] Managing Editors Executive, Australia; [2] Managing Editors Executive, UK; [3] Managing Editors Executive, Canada; [4] Managing Editors Executive, US; [5] Managing Editors Executive, Norway

Background: In 2017 the Cochrane Managing Editors' Executive contacted all Managing Editors (MEs) and Assistant Managing Editors (AMEs) to arrange one-to-one interviews to give them an opportunity to discuss work-related issues. Of the 77 MEs/AMEs contacted, 64 (83%) participated in the interviews. Of the 13 (17%) MEs and AMEs who did not participate: six declined the invitation, three were on long-service leave and four had recently vacated their post. With the upcoming new Editorial Management System (EMS) and the new organization of Networks, we felt it was timely to repeat to interviews in 2020.

Objective: To conduct one-to-one interviews with all Cochrane MEs/AMEs in post in order to proactively identify MEs/AMEs who struggle to keep up to date with changes, objectives, policies, and best practice recommendations, and to understand how the new structures in the organization (Networks, Editorial and Methods Department (EMD) and People Services Department) can more effectively address our concerns to alleviate the pressure in our roles.

Methods: We consulted with Network Senior Editors, the EMD, the People Services Department, and the Informatics and Technology Services to suggest potential questions for this second survey. The MEs' Executive decided on the final list of questions. Using Archie, we identified 85 individuals with the role ME or AME. The invitation was extended to editorial assistants if the ME thought appropriate. We assigned one ME Executive member to interview the MEs/AMEs in each Network. After a general explanatory email, we sent a personalized invitation to each ME and AME to agree on a date, with reminders if necessary, to book a time. After the first round of interviews in 2017, 20 MEs/AMEs had left their posts. We interviewed 28 people for the first time.

Results: The results of the 2017 interviews are available on the MEs' Portal (https://community.cochrane.org/organizational-info/resources/resources-groups/managing-editors-portal/managing-editors-executive/mesames-capacity-building-survey). The second round of interviews are in progress, and we plan to complete this by April 2020.

Conclusion: Cochrane is at a critical stage regarding the development of a new EMS. It is essential that we have an accurate picture of how MEs/AMEs are currently managing workload and the ways in which the newly formed Networks are addressing their needs. To ensure a smooth collaborative development and maximize performance of the new EMS system, it is vital that the ME community is consulted and engaged at all stages of the development.

Patient or healthcare consumer involvement: Patients are an integral part of the editorial process in Cochrane, but they were not directly involved in this project.

Essential core competencies for health policy doctoral graduates

Doshmangir L[1], Mostafavi H[2], Takian A[3]

[1] Health Services Management Research Center, Iranian Center of Excellence in Health Management, School of Management and Medical Informatics, Tabriz University of Medical Sciences, Tabriz, Iran; [2] Vice-Chancellor's Office in Treatment Affairs, Health Economy Department, Shahid Beheshti University of Medical Sciences, Tehran, Iran; [3] Department of Global Health and Public Policy, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran

Background: Competency-based training has gained much relevance in recent years, because of its potential to contribute to knowledge and skills development. This paper explains a set of core competencies required for health policy graduates, aiming to prepare doctoral students for a range of future roles, academic and non-academic.

Methods: We used a three-phase qualitative study:

1) a critical review of the literature;
2) 74 face-to-face and email-based semi-structured interviews; and
3) validating the identified competencies through face to face consultations with qualified experts.

Results: We identified five core competencies without specific order including research, policy analysis, education, decision making, and communication.

Conclusions: The development of evidence-based and updated approaches are essential to identify core competencies for responsive and competency-based education, aiming to improve teaching and learning processes. As countries are gearing up towards sustainable development goals (SDGs), the role of health policy graduates is crucial to pave the way towards SDG 3 on health and well-being. An appropriate and contextually tailored curriculum, as outlined above, is pivotal, we envisage, to foster multi-dimensional competencies that are complementary to the specific disciplines of health policy scholars of future, those who can genuinely serve their health systems towards sustainable health development.

Patient or healthcare consumer involvement: No patient or healthcare consumer was involved in this study.

Evaluating the impact of the knowledge broker mentoring program on evidence-informed decision making using a qualitative approach

Dobbins M[1], Howarth C[1], Husson H[1], Steinberg M[2]

[1] National Collaborating Centre for Methods and Tools, Canada; [2] Marla Steinberg Consulting, Canada

Background: Evidence-informed decision making (EIDM) is important to ensure that public health practice is evidence-informed and resources are being used efficiently and effectively. However, public health professionals can face barriers to EIDM. Capacity for EIDM has improved in recent years, however, gaps still remain. In response, the National Collaborating Centre for Methods and Tools (NCCMT) developed the Knowledge Broker Mentoring program, a hands-on mentorship program to support organizational and individual capacity development for EIDM.

Objectives: To explore the impact of the Knowledge Broker Mentoring program on both program participants and their organizations with respect to EIDM knowledge, skill and behaviours.

Methods: We used a fundamental descriptive qualitative design to evaluate the program. We invited all public health units participating in cohorts one and two of the program to participate in the evaluation. Each organization identified up to five staff to participate in data collection. Eligible staff included program participants, managers, and senior decision-makers in the organization. We conducted telephone interviews using an interview guide. The interviews were recorded and transcribed verbatim. We used a general inductive approach to analyze the data which involved organizing the data into themes and sub-themes based on each area of inquiry.

Results: Fifty respondents from ten local public health departments participated in the study. Of those, 33 respondents were program participants, nine were managers and eight were senior decision-makers. Participants reported increased: confidence; EIDM knowledge and skills; and interpersonal connections. At the organizational level, the groups reported conducting rapid reviews, critically appraising evidence, and using evidence in program planning decisions. Additionally, organizations have put in place ongoing supports to build EIDM capacity. Additional strategies that were implemented across the organizations to support EIDM include dedicated staff positions, working groups, ongoing training, standardized processes, and resources such as guidebooks, frameworks and online portals.

Conclusions: The program was perceived as being highly successful in increasing capacity and furthering EIDM practices. The use of a qualitative design facilitated the identification of impact beyond what would have been observed if the evaluation had been limited to a quantitative approach. The knowledge broker mentoring program appears to be a promising approach to supporting EIDM.

Patient or healthcare consumer involvement: Not applicable.

Evidence-based practice educational program in nursing students' evidence-based practice knowledge and skills: a cluster randomized controlled trial

Cardoso D[1], Couto F[2], Cardoso AF[3], Bobrowicz-Campos E[4], Teixeira-Santos L[4], Rodrigues R[3], Coutinho V[4], Pinto D[4], Ramis M[5], Rodrigues M[4], Apóstolo J[3]

[1] Faculty of Medicine, University of Coimbra; Health Sciences Research Unit: Nursing, Nursing School of Coimbra, Portugal Centre for Evidence-Based Practice: A Joanna Briggs Institute Centre of Excellence, Portugal; [2] Alfena Hospital – Trofa Health Group; Health Sciences Research Unit: Nursing, Nursing School of Coimbra, Portugal; [3] Health Sciences Research Unit: Nursing, Nursing School of Coimbra, Portugal Centre for Evidence-Based Practice: A Joanna Briggs Institute Centre of Excellence, Portugal; [4] Health Sciences Research Unit: Nursing, Nursing School of Coimbra, Portugal; [5] Evidence in Practice Unit & Queensland Centre for Evidence-Based Nursing and Midwifery: A Joanna Briggs Institute Centre of Excellence, Australia

Background: Evidence-based practice (EBP) improves healthcare outcomes and quality. Nonetheless, there are significant gaps between research and practice which challenge use and sustainability of EBP by healthcare organizations and providers. To reduce this gap it is crucial to prepare future health professionals for EBP use in their daily care.

Objectives: To measure the effect of an EBP educational program on undergraduate nursing students' EBP knowledge and skills.

Methods: A cluster randomized controlled trial was performed (ClinicalTrials.gov Identifier: NCT03411668). From the 12 optional courses of the 8th semester (fourth and final year) of the Bachelor of nursing degree, an independent researcher randomly assigned three courses to an EBP educational program and three to education as usual. The EBP educational program was carried out over 17 weeks, including 12 hours of lessons (expositive method and practice method) and six hours of mentorship to small groups of students (three sessions of two hours). The outcomes of EBP knowledge and EBP skills (as assessed by an adapted Fresno Test) were measured at baseline and after the intervention.

Results: One hundred and forty-eight undergraduate nursing students with an average age of 21.95 years (standard deviation (SD) 2.25; range: 21 to 41) participated in the study. The groups were comparable at baseline regarding socio-demographic data and outcomes of interest. We found a statistically significant interaction between the intervention and time on EBP knowledge and skills (P = 0.002). From baseline to post-intervention, students' EBP knowledge and skills improved in both groups (intervention group: P < 0.001; control group: P < 0.001). At the post-intervention, the two groups showed a statistically significant difference in EBP knowledge and skills (P = 0.011) with the intervention group having better performance than the control group.

Conclusions: Both groups showed an improvement in EBP knowledge and skills from baseline to post-intervention. This result was probably due to the fact that all students receive training during the 8th semester to carry out the final year written work which is a literature review. Despite this, the undergraduate nursing students who received the EBP educational program showed higher levels of EBP knowledge and skills at completion of the program when compared with students who received only education as usual. Therefore, nurse educators could consider integrating the EBP educational program into curricula to promote EBP knowledge and skills of future nurses.

Patient or healthcare consumer involvement: Eight educators from different areas (nursing, psychology, education, and physiology) participated in development of the intervention through an expert opinion process, particularly, in terms of structure, content and educational strategies. The suggestions provided by them were carefully analyzed and included in the intervention design.

Evidence-informed decision making in public health: lessons from a journey to build organizational capacity

Mill C[1], Harding J[1], Figliano Scott J[1], Dhawan A[1], Wenger A[1], Hummel J[1], Gardham B[1], Sharma M[1]

[1] Toronto Public Health, Canada

Background: Rapid learning, at an organizational level, includes the timely production of research syntheses for improving internal program planning or decision making. Across Ontario, various public health units, including Toronto Public Health (TPH), contracted The National Collaborating Centre for Methods and Tools (NCMMT) from McMaster University to provide in-depth evidence informed decision making (EIDM) training. This training focused on using the seven steps for conducting a rapid review with the aim of making decisions on adopting, adapting, or eliminating services or programs.

Objectives: This presentation shares our model for training public health staff and management in EIDM, highlights processes to support training and ongoing skill sustainment, and outlines evaluation findings and lessons learned for others interested in building organizational capacity.

Methods: A multi-faceted approach was used to expose staff to EIDM principles and practices. Staff and management received an online self-assessment activity, five days of in-person interactive critical appraisal training, and consolidated their learning by completing a rapid review in teams. To support the training various resources, such as a centralized website, standardized tools and report templates were designed. In addition, we held regular meetings to support the research teams along their journey of completing rapid reviews and a knowledge translation event allowed for the sharing of final results and learnings. An evaluation of the training model revealed areas for improvement.

Results: Between 2017 and 2019, expert consultants from NCCMT trained 33 staff and 12 management, including senior managers. As a result of the training, 13 rapid reviews have been completed. A system has been created to support the rapid and critical appraisal of evidence. The evaluation identified key areas to support organizational capacity building, including:

1) investment in library resources, staffing, and technology;
2) ensure dedicated staff time to complete the training and the reviews;
3) train management to support staff in the process and understand resources required for rapid reviews;
4) use EIDM champions to shift culture; and
5) develop centralized supports (e.g. knowledge broker).

Conclusions: The EIDM journey continues; sustainment efforts are underway to ensure that the EIDM culture is cultivated in the organization. This include integrating EIDM skill sustainment into the annual professional development planners for staff; hosting journal clubs to provide opportunity to practice and maintain critical appraisal skills; creating mentoring opportunities for staff; providing opportunities to complete additional rapid reviews. Continued support from NCCMT staff has been the key to ensuring that staff have support when interpreting or appraising published literature.

Patient or healthcare consumer involvement: Ongoing support of EIDM internally ensures that evidence is leveraged when making programmatic or policy decisions that will impact population health.

Inspiring the next generation: The Cochrane UK and Ireland Trainees Group – a national model with international potential

Carter E[1], Gould R[1]

[1] Cochrane UK, UK

Background: The Cochrane UK and Ireland Trainees Advisory Group (CUKI-TAG) was founded in 2016 to improve trainee engagement and involvement with evidence-based medicine (EBM), systematic reviews in general, and in particular with the work of Cochrane. It is chaired by the Cochrane UK Fellow, is supported by Cochrane UK, and membership is open to all UK and Ireland medical and dental trainees. By encouraging membership from a range of medical specialities and geographical regions we aim to foster a relationship with a diverse group who will become the next generation of clinical leaders. The group is accountable to the Director of Cochrane UK. With four years' experience we have developed a model which could be replicated in other countries to promote trainee engagement with Cochrane internationally.

Objectives: To describe a successful model for engaging trainees nationally in EBM and the work of Cochrane. To discuss how this can be transferred to other countries to support our international colleagues.

Methods: Initial recruitment to CUKI-TAG was by advert. Via regular online meetings the committee has organized events at the UK and Ireland Cochrane Symposium and the 2018 Cochrane Colloquium. We host an annual trainees' conference to provide EBM skills to support journal clubs in local hospitals. We have used a variety of online methods to engage trainees including a detailed website of training resources linked to the Cochrane UK website, Evidently Cochrane blogs, Twitter journal clubs, Instagram, podcasts, and videos. Cochrane UK has provided travel expenses for meetings, Cochrane conference scholarships and training events.

Results: The annual trainees' conference is well attended and receives excellent feedback. There is demonstrable impact on the involvement of trainees in writing blogs, local journal clubs and online Twitter journal clubs. There is a positive self-reported use of online EBM resources, engagement with the Cochrane community and on personal development. CUKI-TAG meets regularly online and semi-regularly in person. Recruitment varies: we have had between 2 and 22 members over the four-year period. We have experienced the highest impact using our website, face-to-face events, Twitter journal clubs and Evidently Cochrane blogs. Other projects have involved creating videos and podcasts. We have experienced moderate impact with podcasts and videos.

Conclusions: Our successful model involves recruiting and retaining a national trainees' committee, hosting an annual trainees' conference and supporting trainee initiatives. The CUKI-TAG chair links the work of individual committee members to the Cochrane community and EBM resources. Future direction includes promoting speciality-dependant trainee engagement via diversifying and expanding our committee. We would encourage the transfer of this successful model to our international colleagues to engage trainees and inspire the next generation of clinical leaders.

Instruction on multiple uses of health research reporting guidelines throughout the systematic review process

VonVille H[1]

[1] University of Pittsburgh Health Sciences Library System, USA

Background: Authors of systematic reviews (SRs) have learned to indicate they follow the PRISMA reporting guidelines although many are not transparent nor clear in their reporting. New researchers, and especially student researchers, attempt to complete systematic reviews with little idea of the methodology, how to describe aims, or how to select dependent/independent variables. Faculty are often not able to advise them appropriately.

Objectives: The goal of this presentation is to describe and demonstrate multiple ways a wide variety of health research reporting guidelines can be incorporated into both the instructional and advising setting to assist researchers in developing coherent and transparent systematic reviews.

Methods: Faculty and students have been provided instruction on, and tools to assist with, using reporting guidelines to develop the SR protocol, data extraction forms, and evidence tables.

Results: While anecdotal, students have reported a better understanding of the SR process by walking through the protocol document. Data extraction form development has been simplified, and the tool for creating evidence tables aids in the visualization of study data.

Conclusions: Health research reporting guidelines should be incorporated throughout the research process, not saved until it is time to write the final draft.

Is evidence-based medicine a mirage in resource-constraint settings? A survey among resident doctors in selected teaching hospitals in Nigeria

Bello S[1], Ajayi D[1], Bamgboye E[1], Mpama E[2]

[1] College of Medicine, University of Ibadan, Nigeria; [2] College of Health Sciences, University of Calabar, Nigeria

Background: The adoption of evidence-based medicine (EBM) in clinical practice has demonstrated effectiveness in reducing mortality, shortening hospital stay and bringing down the overall costs of patient care. Despite the strides that EBM has achieved globally, the adoption in some parts of the world is at best rudimentary.

Objective: To assess the self-rated knowledge of, attitude towards, practice (KAP) of, and barriers to EBM among resident doctors in Nigeria.

Methods: This was a cross-sectional web-based survey among 238 resident doctors in four selected teaching hospitals in southern Nigeria. The survey questionnaire contained items assessing the KAP of EBM, familiarity with and understanding of key EBM terms, the use of EBM in decision making, barriers to EBM and perceived ways to improve EBM adoption. We report proportions and summary statistics for the distribution of survey items.

Results: Mean number of years in clinical practice was 9.3 ± 4.5 years. Respondents were uniformly distributed in major clinical specialties. The majority (70.5%) were senior registrars. Respondents' understanding of EBM components included: current best clinical evidence (98.3%), clinical expertise (65.5%), and patients' choices (36.6%). Self-rated familiarity with EBM terms was high while perceived understanding of the terms was lower. The least understood concept was heterogeneity (20.6%). The attitude towards EBM was generally positive. Only about half (53.6%) had used medical bibliographic databases within the last six months prior to survey. Barriers to EBM included lack of time (47.1%) and lack of requisite skills (32.4%). Suggestions to improve EBM adoption included training (58.1%), provision of free Wi-Fi, and free access to bibliographic databases (25.2%) and increased political will (23.1%).

Conclusion: A further understanding of the EBM concept, provision of enabling infrastructure, regular clinical audit and advocacy to hospital management and clinical consultants, may improve the level of adoption of EBM.

Knowledge in systematic reviews and meta-analysis in postgraduate students of a medical school in Colombia

Sanchez G[1], Plazas M[1]

[1] Foundation University of Health Sciences, Colombia

Background: Cochrane has developed activities in Colombia for approximately 20 years, however, it was only in 2019 that the Cochrane Colombia Geographic Center was officially established, with the participation of four associated centers. One of them is located in the University Foundation of Health Sciences (FUCS), which is an institution that belongs to Cochrane since 2012, and has been working to increase in its postgraduate students the knowledge and practical use of systematic reviews as a tool to inform decisions in clinical practice.

Objectives: To describe the knowledge about systematic reviews and meta-analysis in a sample of postgraduate students from a medical school belonging to Cochrane Colombia Geographic Center (Cochrane Associated Center-FUCS), and to assess the level of information that students have about Cochrane as an organization.

Methods: Design: Cross sectional study. Setting and participants: Between January and March 2020, a survey was sent to a sample of graduate students from a medical school in Bogotá. Students belong to a set of different clinical specialization programs (example: internal medicine, surgery, others), who in their formative process receive specific training in research methods, including systematic reviews. Sample size assumptions: To establish the sample size assumptions, a pilot survey was conducted with 30 randomly chosen students. Sampling frame of graduate students (N = 676), expected percentage of right interpretations of a pooled effect (89%), maximum expected difference (5%, two tailed test) and alpha value (0.05). The calculated sample size was 124 students. Variables and measurement tool: A structured questionnaire, looking for sociodemographic characteristics and knowledge about systematic reviews, meta-analysis and Cochrane as an organization. Analysis: We performed a descriptive analysis using absolute and relative frequencies. For numerical variables we used central tendency and dispersion measures. In order to make inferences to our population of graduate students, we calculated 95% confidence intervals

Results: A total of 142 students answered the survey, of which 52.8% were women. The average age was 28.2 years (standard deviation 4.1 years). 80.9% have previous training in epidemiology (master of specialization). 92.3% know the correct concept of a systematic review and 78.9% was able to identify the right steps for a systematic review (know the process), (see Table 1).

Conclusions: The results of this study allow us to have a baseline to strengthen the formative process of postgraduate medical students both in theoretical aspects and those related to the Cochrane Network. Future studies should explore aspects related to teaching methodologies, as well as strengths and barriers that students identify to develop knowledge of systematic reviews.

Patient or healthcare consumer involvement: Graduate students from a medical school in Bogotá are the consumers. They are developing competencies to critically understand and appreciate systematic reviews and meta-analyses.

Additional file: Table 1

Professionalization of health services management mapping and the role of the education sector: scoping review

Daire J[1]

[1] Curtin University, Australia

Background: There has been increasing recognition of the importance of healthcare management in improving health service performance and strengthening the capacity of healthcare systems. However, healthcare management has yet to be formally accepted as an occupation with clear professional standards. The pathways by which countries may develop to sustain a high performing healthcare management workforce are fragmented with focus on different aspects of management and professional competency requirements. In order to inform the establishment of health service management as a formal profession, the current efforts in providing training and development opportunities and requirements for health service managers need to be fully understood.

Objectives: Mapping the theoretical concepts, models and frameworks of management and leadership development in globally, this study aims to develop a guiding frame of what 'healthcare management profession' would look like and key aspects incorporated in advancing professionalization of healthcare management globally. The study will also clarify the role of education sector and professional associations in improving education and practice of health management.

Methods: Document analysis on Master of Health Administration programs globally and systematic review of both peer-reviewed and grey literature published in the past 10 years using a PRISMA methodology. Relevant peer reviewed articles will be searched from Emerald, ProQuest, Scopus and Google Scholar. The key words include; healthcare, health services, management, leadership, and health administration. Information will be extracted from eligible articles using a data extraction sheet both descriptively and thematically.

Results: Document analysis and systematic review will be conducted between April and June. Initial results of the analysis will be available by early July in time for presentation and discussion at the SHAPE Symposium. The purpose of the presentation is to discuss the initial findings and further research questions and directions. The experience from the co-author in conducting management competency studies in Australia and other Asia pacific countries and also the brief review of the literature conducted by the corresponding author confirm the strong demand for professionalizing the health service management workforce. Examples can be the increasing research interests in health service management and competency development, the establishment of health management professional associations, and the newly developed and validated management competency frameworks to guide management practice and education

Specific strategies to increase access to Cochrane training opportunities in South Africa and the African region

Hohlfeld AS[1], Kredo T[1], Davids E[1], Galloway M[1], Ndwadwe D[1], Oliver J[1], Pienaar E[1], Durao S[1]

[1] Cochrane South Africa, South African Medical Research Council, South Africa

Background: Cochrane South Africa (CSA), as a Geographic centre, provides training on conducting and disseminating high-quality Cochrane Reviews to inform healthcare decisions relevant to their context. Ensuring equitable access to training in South Africa and sub-Saharan Africa remains challenging, with the small CSA team unable to match the demand for face-to-face training. We thus developed a strategy to address inequity in access to Cochrane training in South Africa and the region.

Objectives: To describe CSA's approach to increasing access to Cochrane training in South Africa and sub-Saharan Africa.

Methods: We implemented two approaches. We identified and reached out to health sciences faculties in 'historically disadvantaged institutions' (HDIs), labelled so due to previous racial barriers to accessing adequate resources. We aimed to implement an introductory workshop on the principles of evidence-based healthcare. Each workshop was evaluated through a post-workshop written survey. We also shifted from face-to-face monthly methods training to monthly webinars on a variety of methods topics directed at novice authors and evidence users including the experienced ones. We evaluated the webinars through an online survey in December 2019.

Results: Of the six HDIs contacted, we delivered three workshops in 2019, with 82 health sciences post-graduates and practitioners attending (Table 1). Of these, 67 completed evaluations and all gave positive feedback (Table 2). Due to logistical matters, the remaining three HDIs postponed the workshops to the following year. We presented nine webinars in 2019 to about 450 delegates in sub-Saharan Africa. On average 42% of those who confirmed their attendance actually attended. Each webinar had a mean of 46 participants from at least 8 African countries. The response rate to the survey was 6% (27/450). Most participants were involved with evidence synthesis and noted that the webinars were important resources that allowed them to do their work using newly learnt skills (Table 3).

Conclusions: CSA as a regional Centre expanded the reach of its training through specific targeted approaches, to combat the historical inequities that still exist in attaining knowledge in the region. This process is in its infancy. Careful evaluation of these efforts will continue. This will allow us to adapt our approaches to ensure a greater reach and an improved attendance rate.

Patient or healthcare consumer involvement: We strive to reach out and share the importance of Cochrane methods and reviews for decision-making. Through the current training methods, CSA has attempted to ensure that healthcare consumers remain involved so that they may provide the best necessary care to patients.

Additional files: Tables

Standing out in an evidence-informed world: online training and credentialing for evidence-informed public health

Clark E[1], Dobbins M[1]

[1] National Collaborating Centre for Methods and Tools, Canada

Background: Evidence-informed decision making is emerging as a must-have skill in public health. Job candidates are expected to know how to integrate research and community-level evidence into practice, program and policy decisions. Public health organizations that champion evidence use in Canada have integrated evidence-informed decision-making skills assessments into the hiring process.

Objectives: As leaders in evidence-informed public health, the National Collaborating Centre for Methods and Tools (NCCMT) sought to develop a program that would allow public health students and professionals to build the skills they need to enter an evidence-informed workplace as well as achieve the certification required to distinguish themselves in pools of applicants.

Methods: The NCCMT developed an online suite of learning modules focused on developing skills for evidence-informed public health. The modules offer interactive, problem-based learning incorporating realistic public health scenarios. Through a partnership with McMaster University Continuing Health Sciences Education, successful completion of the suite of modules grants the learner a certificate in evidence-informed public health endorsed by McMaster University, a global leader in evidence-based practice.

Results: Since launching the certificate program in late 2017, over 300 learners have registered for the modules. The majority of registrants have been students, indicating the perceived value of a certificate issued by a globally recognized leader in evidence-informed practice when entering the workforce. Analysis of pre- and post-module tests of knowledge and skill indicate an increase in scores.

Conclusions: As the prioritization of evidence-informed practice continues to increase across Canada and worldwide, there is a need for accessible training in evidence use for public health professionals. The NCCMT has developed an online program that delivers this training and distinguishes successful learners with a globally-recognized certificate for evidence-informed public health.

US Cochrane Network:who are we?

Hill M[1], Haas DM[2], Kariyawasam D[1], Kumar N[2], Jahanfar S[1]

[1] US Cochrane affiliate-Central Michigan University, USA; [2] US-Satellite-PCG, USA

Background: The scientific community in the United States (US) is now linked with the international Cochrane body through the US Cochrane Network since the June 2019 launch in Washington, DC. Central Michigan University Affiliate of Cochrane US-Network is responsible for fulfilling training and knowledge translation gaps.

Objectives: We aimed to investigate the whereabouts of Cochrane affiliated individuals in the United States. We planned to determine how many authors, reviewers, translators, statisticians are self-registered in ARCHIE and their geographical distribution across the USA.

Methods: We used ARCHIE to extract data. Various filtering was used to generate data on the number of authors, reviewers, consumers, statisticians, editors and translators with generated addresses and affiliation is the US. We created frequency tables of descriptive characteristics of people registered using SPSS and produced mapping using StatPLanet software.

Results: Of 18,739 authors identified with a US-based address in the ARCHIE database, 10,617 members self-identified themselves to be a part of the Cochrane community. Figure 1 shows that 8122 members belonged to a primary group. Of those registered with a primary group, 1600 (19%) had at least two roles, such as authors and reviewers. Of those who had only one role within a primary group, 3353 41%) were authors, and 2818 (35%) were peer reviewers. Moreover, the following frequencies were found using a simple descriptive analysis: Editors (n = 116, 1.43%), managing editors (n = 4, 0.04%), consumer reviewers (n = 142, 1.74%), statisticians (n = 13, 0.16%), information technologists (n = 9, 0.11%) and handsearchers (n = 17, 0.2%). In addition, there were 50 translators (0.61%), but it was not clear what languages these translators can work with. From those who are found in ARCHIE, only 4650 (57%) are identified as active by ARCHIE administration staff. The mapping of addresses stated by the Cochrane members across the US shows a clear gap in many states (e.g. Figure 2). The remaining maps will be presented at the Colloquium.

Conclusions: There are still many areas in the US that do not have a Cochrane member or have very few members engaged in Cochrane activities. There is also a need to update ARCHIE and create a survey to identify the gaps and needs of Cochrane members across the US.

Additional files: Figures

Using children and adolescents as standardized patients in health: a scoping review and evidence map

Yu X[1], Hou X[2], Chen Y[3]

[1] WHO Collaborating Centre for Guideline Implementation and Knowledge Translation, China; [2] School of Public Health, School of Medicine, Shanghai Jiaotong University, China; [3] Evidence-Based Medicine Center, Lanzhou University, China

Background: Using children and adolescents as standardized patients play an increasingly important role in clinical teaching and practice. However, there are a lack of systematic descriptions of the current research status and existing problems on this topic.

Objectives: To evaluate children and adolescents as standardized patients, analyze the current research quantity, attribute and characteristics of children and adolescents as standardized patients, find the research gap and provide an evidence basis for developing guidelines and policies.

Methods: This review used Arksey and O'Malley's scoping review methodological framework as a guideline. We searched MEDLINE via PubMed, Embase, Web of Science, CBM, CNKI, Wanfang and retrieved all Chinese and English literature related to children and adolescents as standardized patients published from inception to 31 December 2019. We also searched Google Scholar, Baidu Xueshu, gray literature, references included in the studies and consulted experts in related fields. The research team is an interdisciplinary research team. Two review authors independently screened literature according to the inclusion-exclusion criteria formulated in advance and consulted a third-party when there was a dispute.

Results: We included 35 articles in this scoping review. Most of the children and adolescents as standardized patients were studied in the United States (13), Canada (14) and the United Kingdom (4). Four articles are review, of which 31 articles, children and adolescents as standardized patients was mainly used to educate (15) or evaluate (16) medical students or healthcare providers. Using children and adolescents as standardized patients are mainly used in clinical settings (such as diarrhea, headache, etc.), doctor–patient communication, and mental health.

Conclusions: After more than 30 years of research and experiments, it is feasible and effective to use children and adolescents as standardized patients. In future education and evaluation of the health setting, more children and adolescents need to actively participate in and kindly hear their voices.

Patient or healthcare consumer involvement: None.

Writing a Cochrane Review update: getting started and challenges to overcome

Carter E[1], Elstub L[2], Wallace S[2], Vale L[2]

[1] Cochrane UK, UK; [2] Cochrane Incontinence, UK

Background: Cochrane Reviews updates can appear a daunting undertaking for new review authors. There are multiple stages which may not be immediately apparent: revisiting the protocol, updating the search, piloting revised eligibility criteria, inclusion assessments, piloting (revised) data collection forms, application of GRADE to assess the certainty of the body of evidence and rewriting the main text. Updating reviews is vital to maintain the Cochrane Library. The promise of mentorship and support of a Cochrane Review Group (CRG) can act as a strong motivator for a variety of new review authors to get involved with Cochrane.

Objectives: To provide a troubleshooting guide and helpful tips for a new review author updating a review.

Methods: Using our current work on a Cochrane Review update as a case study, we discuss key components of the process, challenges faced, solutions, and lessons learned.

Results: The advantages of updating an intervention review include the opportunity, with the benefit of new patient and consumer involvement, to ensure that the study question, comparisons and outcomes remain clinically pertinent in the light of new research. New review methods may now be in use that were not in place for the original review, including risk-of-bias assessment and GRADE, and a relevant core outcome set may now be available. Targeting these issues will enhance the review quality. Challenges include: the time-consuming process of revisiting the original study papers to produce a full risk-of-bias assessment or extract data on a new outcome; loss of data such as original data extraction forms or translated papers; engagement of the original study authors with the update; and the potential volume of combined new and old studies. Synthesis of new and existing data in light of revised comparisons and outcomes requires careful planning.

Conclusion: Our troubleshooting guide and helpful tips focus on the specific challenges of review updates. These are a helpful way for new review authors to get involved with Cochrane, providing the opportunity for comprehensive systematic reviewing training with the support of a CRG within an established organization. Review authors will additionally benefit from recognized training courses and the Cochrane Library of resources, fostering a culture of lifelong involvement. By encouraging new review authors to get involved with Cochrane and supporting them through the review process, the organization will be diversified internationally, particularly in relation to the engagement of those with restrictions on their time, e.g. clinicians, who provide valuable guidance with steering the direction of future research. This enables continuing worldwide clinical input into CRGs, who may benefit from the use of technology initiatives such as Covidence and RevMan Web.

SEARCHING AND INFORMATION RETRIEVAL

A bibliometric analysis of Web of Science published literature on adverse drug effects

Diaz-Barrera ME[1], Pacheco-Mendoza J[2]

[1] Universidad Nacional de Trujillo, Perú; [2] Unidad de Investigación en Bibliometría, Universidad San Ignacio de Loyola, Perú

Background: Adverse effects during hospital admissions affect nearly one in 10 patients, 15.1% of these effects are drug-related, whereas 7.4% effects were lethal, all of which have a median percentage of preventability of 43.5%. Regarding scientific production (SP) in this field, it represents 10.8% of all drug-related literature, and despite the growth of this area in the last 20 years, the bibliometric analyses developed to describe this production have not analyzed the data from the organization types that produce these documents or the funding agencies.

Objectives: To perform a bibliometric analysis of Web of Science (WoS) published literature on drugs adverse effects.

Methods: We searched the WoS – Core Collection database the terms "adverse effects" and its synonyms or variants, using thesaurus from PubMed, Embase, and related articles that allow us to focus on adverse effects caused by drugs. Then, we refined the results according to document type (articles, meetings abstracts, reviews, letters, and proceeding papers) and year of publication: from 1981 to 2019. These metadata were processed using InCites (Clarivate Analytics). We analyzed the scientific production in general (by trends and journals), according to the InCites classification for organization type and funding agencies.

Results: We found 34,818 documents with the Emerging Sources Citation Index (ESCI) according to our search strategy. The scientific production in the past 10 years was 22,657 (65% of whole production found). Figure1 shows that 2019 had the highest production (2753). The higher production according to journals is presented in Figure 2, which shows Drug Safety (1607) leading the top ten. Regarding scientific production according to organization type (Table 1), academics have a higher presence (63.15%) in all analyzed documents. Regarding funding agencies (Figure 3), only 11,113 (31.91%) declare the funding source.

Discussion: Our results present a deep analysis of the scientific production of this area regarding the reports of previous reports. The scientific production in this area has been increasing over the years. Two specialized journals in pharmacovigilance and post-marketing surveillance focus on 13.16% of the total production (Figure 2). The role of academic organizations in this research field is bigger than that of other organization types. Reports of funding in our data are very low (31.91%) and do not match with previous works that globally assess the Science-Citation Index-Expanded (SCIE), which has 52% of this section filled (ESCI barely filled 20% of this section). Our results could suggest bad practices in scientific integrity regarding the declaration of funding sources. Despite this observation, the organization type that has reported the major number of funding papers is a government institution and near this in the top ten, we found three corporations (Pfizer, GlaxoSmithKline, and Novartis), which according InCites have funded 308, 221, and 173 investigations and take part as author in 41(13%), 37(17%), and 5(3%), respectively.

Additional files: Table 1; Figure 1; Figure 2; Figure 3

Achieving evidence interoperability in the computer age: setting evidence on the FHIR (Fast Healthcare Interoperability Resources)

Alper B[1], Munn Z[2], Salas N[3], Tristan M[3], Brand L[2], Lorio A[4], Schilling L[5]

[1] EBSCO Health, USA; [2] Joanna Briggs Institute, Australia; [3] Cochrane Central America – IHCAI institute, Costa Rica; [4] McMaster University, Canada; [5] University of Colorado Anschutz Medical Campus, USA

Background: Evidence standard practice requires the use of research results to inform the computers that can add capacity for evidence-based practice by making the information from research results, standard the summarizations searchable and re-usable without labour-intensive manual screening and repetition of data entry. Such interoperability can be achieved by establishing universal standards for data exchange for communicating evidence concepts in machine-interpretable formats.

Objectives: The "Extending FHIR for Evidence-Based Medicine Knowledge Assets" (EBMonFHIR) project is defining a standard for computable, interoperable expression of evidence. As part of a universal architecture for the Evidence Ecosystem that aligns with the FAIR data principles (making knowledge Findable, Accessible, Interoperable, and Reusable), the EBMonFHIR Work Group is in active development with a substantial coalition of international organizations and co-ordination with other standards development groups.

Methods: We are using the HL7 standards development methodology, including weekly open meetings, and three Connect-a-thons per year, to extend FHIR to create resources for exchanging descriptive, statistical and certainty of the evidence.

Results: The EBMonFHIR project was approved in June of 2018. As of February 2020, we have defined new Resources (Evidence Resource and EvidenceVariable Resource) and new Data Types (Statistic Data Type, Ordered Distribution Data Type) to define exchangeable interoperable evidence results entirely. The project website (https://confluence.hl7.org/display/CDS/EBMonFHIR) includes multiple examples and information on how to participate.

Conclusions: Working together we can achieve interoperability for evidence in the electronic era to realise the technological breakthroughs we see in other domains such as navigation support. Typical information architecture will also facilitate the harmonization of 'Real World Evidence' and 'Evidence-Based Medicine' which collectively represent a clear understanding of evidence and its certainty, regardless of evidence source. Extending the solutions achieving interoperability for healthcare services provide a means to not only solve this challenge for the Evidence Ecosystem but also to keep it well connected with healthcare services delivery.

Patient or healthcare consumer involvement: None.

Cochrane Madrid and BiblioMadSalud: an example of co-operation between Cochrane and librarians

Campos-Asensio C[1], Estrada Lorenzo JM[2], Rebollo Rodríguez MJ[3], Stallings E[4], Monge D[5], Zamora J[4], López-Alcalde J[6]

[1] Biblioteca del Hospital Universitario de Getafe, Madrid; Comité Ejecutivo de BiblioMadSalud, Spain; [2] Biblioteca del Hospital Universitario 12 de Octubre, Madrid; Comité Ejecutivo de BiblioMadSalud, Spain; [3] Biblioteca del Colegio de Médicos de Madrid; Comité Ejecutivo de BiblioMadSalud, Spain; [4] Instituto Ramón y Cajal de Investigación Sanitaria (IRICYS); CIBER Epidemiology and Public Health, Spain; [5] Universidad Francisco de Vitoria (UFV), Madrid, Spain; [6] Cochrane Madrid; Universidad Francisco de Vitoria, Madrid; Instituto Ramón y Cajal de Investigación Sanitaria (IRYCIS), Madrid; Institute for Complementary and Integrative Medicine, University Hospital Zurich and University of Zurich; CIBERESP, Spain

Background: A rigorous search is a key stage in the synthesis of the best available evidence. BiblioMadSalud, founded in 2019, is an association that brings together professionals from libraries and documentation health science centres in the Community of Madrid, Spain. Its objective is to promote the co-ordination and collaboration between its institutions and professionals and achieve greater visibility of library professionals. Cochrane Madrid is located at Hospital Universitario Ramón y Cajal (Madrid). It was founded in 2015 through an agreement between Instituto Ramón y Cajal de Investigación Sanitaria (IRYCIS) and Universidad Francisco de Vitoria (UFV), Madrid. Cochrane Madrid has detected a shortage of librarians with experience in the design and execution of searches for Cochrane Reviews.

Objectives: This presentation aims to describe our experience in fostering the collaboration between Cochrane Madrid and a network of health science librarians (BiblioMadSalud).

Methods: A descriptive study reporting the approach followed in Madrid to speed up the collaboration between Cochrane and BiblioMadSalud.

Results: During 2019, Cochrane Madrid and BiblioMadSalud collaborated in two training activities on literature searches for systematic reviews (a total of 90 participants). In addition, an oral presentation was delivered during the presentation of the Cochrane Library at the Spanish Ministry of Health, Consumer affairs and Social Welfare (Madrid). In February 2020, the collaboration agreement was signed, which aims to achieve the following objectives, among others:

1) offer training focused on the search for evidence in the framework of Cochrane Reviews;
2) generate a consensus document detailing the potential functions of the librarian within a Cochrane Review; and
3) initiate collaborative methodological projects;
4) Collaborate on key projects for Cochrane, such as the manual search for clinical trials.

Conclusions: The collaboration between Cochrane Madrid and a network of health science librarians (BiblioMadSalud) represents a unique opportunity to effectively involve the librarians in Cochrane and, thus, to improve search process and collating literature which are crucial steps for systematic reviews.

Patient or healthcare consumer involvement: The agreement will consider the training of patients' associations on gathering health information based on the best available evidence.

Contribution of Indian randomized controlled trials in Cochrane Reviews

Ezekiel JPB[1]

[1] The Brain Vascular Malformation Consortium (BVMC), Christian Medical College, Vellore, India

Background: With the evidence-informed healthcare movement spreading across the world contextualization and applicability of evidence is becoming a key area of concern.

Objectives: To evaluate the contribution of Indian randomized controlled trials (RCTs) in Cochrane Reviews.

Methods: We searched for active Cochrane Reviews in all 45 Cochrane Review Groups (during the period from January 2019 to June 2019. We identified Indian studies from lists of included studies, excluded studies, ongoing studies or studies awaiting classification.

Results: We retrieved 8995 Cochrane Reviews out of which 325 had at least one Indian RCT marked as either included, ongoing, or awaiting classification. This accounts for only 4% of the data. India is a tropical region and its connectivity to the infectious health hazards alone contributes to the highest amount of data (12%) in the Cochrane Reviews of interventions. There are nearly 427 Indian studies contributing data to the meta-analysis. The contribution of Indian data to the global evidence ranges from 1% to 12% in various Cochrane Reviews groups.

Conclusions: There is a need to conduct high-quality RCTs in India. Global health funders need to fund more trials in India to improve the generalizability of evidence.

Development and validation of two sensitive and precise search filters for retrieval of clinical practice guidelines

Lunny C[1], Salzwedel D[1], Kitchin V[2], Liu T[3], Ramasubbu C[3], Puil L[1], Vis-Dunbar M[2], Koepfler L[4], Rishar R[4], Gerrish S[1], Wright JM[1]

[1] Cochrane Hypertension Group, University of British Columbia, Canada; [2] University of British Columbia, Canada; [3] Pharmaceutical Sciences, University of British Columbia, Canada; [4] ECRI, USA

Background: Given the difficulty in navigating large volumes of information and limited time for searching the literature, clinical practice guidelines (CPGs) are important sources of evidence for healthcare providers. If large bibliographic databases are to be helpful, clinicians must be able to retrieve relevant references, without missing key citations or retrieving excessive numbers of irrelevant reports. Search filters designed for MEDLINE may provide a more efficient way to retrieve CPGs. Search filters can maximise the number of relevant results while minimizing the number of irrelevant ones.

Objectives: To develop validated search filters in Ovid MEDLINE using text mining techniques and measure their performance according to sensitivity and precision. We aim to develop a sensitivity-maximizing filter, and a sensitivity-and-precision maximizing filter for retrieval of CPGs.

Methods: We derived two samples of CPGs: a "test set" of CPGs (n = 100), and a validation set of CPGs (n = 100). Using the test set, we conducted text mining to determine the frequency of terms (one word, bigrams and trigrams) in the titles, abstracts, ti/ab and full text. Candidate terms were combined iteratively and tested in Ovid MEDLINE. Development of the search filters focused on precision (without compromising sensitivity) as this will help users to cut back on screening time and resources. Using the most frequent terms and MeSH, two researchers developed the strategies independently, then compared, refined, and finalized the optimal strategies for each filter type. If the strategies changed based on the comparisons, we recalculated the precision and sensitivity. We used transparent instructions to create the strategies to increase standardization of procedures. Finally, we validated our final filters on an external validation set of guideline citations and calculated the sensitivity and precision.

Results: To our knowledge, this is the first study to develop validated search strategies using text mining for identifying CPGs in Ovid MEDLINE. We developed semi-objective search filters: a sensitivity-maximizing strategy, and a sensitivity-and-precision maximizing strategy to retrieve CPGs. The text mining software enables large amounts of n-grams to be sorted by frequency into matrices, allowing for a more objective choice of single and multiple terms used in testing search algorithms. We used different text mining applications and software to identify key terms and synonyms for guidelines on different topics. The sensitivity-maximizing filter should be used when there is need for comprehensiveness and when the filter is appended to search terms for specific conditions or interventions. When the sensitivity-maximizing filter is appended to search terms for conditions or interventions, it is unlikely to result in an unacceptably large number of citations to screen.

Conclusions: The search filters enable more efficient identification of CPGs in Ovid MEDLINE.

Patient or healthcare consumer involvement: None.

Development of a bibliographic database in response to COVID-19

Allen T[1], Garnica Carenno J[1], Norris SL[1]

[1] WHO, Switzerland

Background: The World Health Organization (WHO) requires real-time data and information from the peer-reviewed scientific literature to inform daily decisions regarding the COVID-19 pandemic.

Objectives: To describe the development, maintenance, and dissemination of search results and a bibliographic database related to COVID-19 for WHO staff and for public access.

Methods: WHO senior management requested daily searches for COVID-19 to support the COVID-19 Incident Management Team (IMT) on 20 Jan 2020. Information scientists and systematic review and guideline experts at WHO developed search strategies and selected relevant journals and publisher websites. Screening and dissemination started on 26 Jan 2020: manual searches were conducted each weekday, and citations were compiled using EndNote, then imported and deduplicated in Covidence. Simple inclusion and exclusion criteria were developed, and initially two screeners reviewed titles and abstracts for inclusion. Each included citation was then tagged with study design (e.g. case series) and with a broad topic area (e.g. clinical characterization). The daily output was a list of citations in Microsoft Word, including DOIs, and disseminated to key IMT members. We then developed a searchable database in Microsoft Power BI, implemented 12 February 2020 on the WHO public website, while daily feeds to IMT continued. The database was downloadable in various formats. We improved the search interface using BIREME, launched on 9 April 2020. BIREME is a Specialized Center of PAHO/AMRO (www.paho.org/en/evidence-and-intelligence-action-health).

Results: The daily number of citations increased progressively, starting with 8 on 26 January growing to 1120 on 14 April. Daily handsearches were maintained in order to avoid indexing delays with proprietary bibliographic databases. Public usage was monitored and steadily increased, with approximately 100,000 visitors daily to the Microsoft Power BI database. Given demands on staff time during the outbreak, a formal evaluation of use by WHO staff was not performed. We encountered challenges during the development and implementation of the database. First, the project was very resource intensive, taking a full-time library staff, technical support and screener time. This decreased as streamlining and augmented technological approaches were implemented. Second, a large amount of research was published in Chinese and in Chinese-based journals which initially we did not have the capacity to search. Third, we had difficulty identifying key studies to draw to the attention of the IMT. Fourth, the database had to serve two audiences: WHO's staff dealing with the outbreak and a diverse external audience.

Conclusions: Daily searches and a searchable database were achieved to help meet the needs of WHO's COVID-19 IMT as well as an external audience. The lessons learned will facilitate the development and implementation of a bibliographic database in future public health emergencies.

Development of a Chinese-language bibliographic database in response to COVID-19

Yu X[1], Guo Q[1], Wu S[1], Ren M[1], Wang Q[2], Chen Y[1], Norris SL[3]

[1] Evidence-Based Medicine Centre, Lanzhou University, Lanzhou; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation, Lanzhou; Lanzhou University, Cochrane China Network, Lanzhou, China; [2] Department of Health Research Methods, Evidence and Impact, Faculty of Health Sciences, McMaster University, Hamilton, Canada; [3] WHO, Switzerland

Background: The COVID-19 pandemic started in Wuhan China in December 2019 and as a result critically important scientific literature has emerged from Chinese clinicians, healthcare workers and researchers. The World Health Organization (WHO) therefore began working with the WHO Collaborating Centre for Guideline Implementation and Knowledge Translation (WHO CC) of Lanzhou University, China to develop and maintain a repository of relevant Chinese-language scientific publications related to COVID-19.

Objectives: To describe the approach for developing, maintaining and disseminating search results and a bibliographic database of Chinese-language publications related to COVID-19.

Methods: WHO approached the WHO CC on 31 January 2020 regarding the possibility of developing a Chinese-language bibliographic database. A multi-disciplinary team of researchers at the WHO CC worked to establish the database, including one methodologist, one librarian, and several researchers. The librarian and methodologist developed search strategies and selected two Chinese Core databases (China National Knowledge Infrastructure and the Chinese Medical Journal Network Database). In addition, they manually searched 12 Chinese core journals. Data were extracted for each publication, including title, author, journal, DOI, abstract, keywords, categories, and full-text link. When English abstracts were not provided by the journal, the WHO CC provided a brief summary in English. The extracted information was presented in Microsoft Word and sent to WHO daily, along with an EndNote library to facilitate searching.

Results: The WHO CC Center started to deliver COVID-19 Chinese literature to WHO on 7 February 2020. As of 12 April 2020, 2057 publications have been delivered over 66 days. The number of daily publications steadily increased to a peak on 18 March (n = 66), after which the number trended downward, although with significant fluctuation. The main types of publications included peer-reviewed journal articles (including pre-online publications) and guidelines or expert consensus publications. The WHO CC team needed approximately 16 to 18 hour per day for searching, screening, information extraction, translation and editing. At the time of writing, efforts are being made to enhance the search capabilities and to move the database into the public domain, combining it with WHO's main COVID-19 publication database. While a formal evaluation has not as yet been feasible given the outbreak situation, anecdotal reports from the WHO Incident Management Team indicate that the Chinese COVID-19 literature database is reviewed daily and informs decision-making. The Cochrane Network and teams performing reviews for WHO used this database. A formal evaluation is planned once the outbreak situation allows.

Conclusions: Daily searches of the Chinese-language literature complement the primarily English-language bibliographic searches conducted by WHO. This resource appears to be highly valued by WHO. Work is underway to make the database public so that clinicians, researchers and institutions globally can access it.

Differences in database performance for retrieving reports of randomized and qualitative studies: experiences from a mixed-methods review

Stovold E[1], Dennett E[1], Janjua S[1]

[1] St George's, University of London, UK

Background: As part of a NIHR programme grant we are conducting a Cochrane Review of tailored interventions for people with COPD and comorbidities. The review topic was prioritized by our patient steering group. The review will include both randomized controlled trials (RCTs) and qualitative studies. Our literature search of multiple databases yielded thousands of results. So that our time could be directed as efficiently as possible, we wanted to find out which databases the included studies were retrieved from to help us decide which sources to continue searching.

Objectives: For a mixed-methods review on interventions for people with COPD and multi-morbidities including both RCTs and qualitative studies we aimed to:

1) analyse which databases the included studies were found in;
2) identify the overall coverage of each database;
3) identify which databases contained unique studies.

Methods: We searched the following databases in June 2019: Cochrane Airways Trials Register; Cochrane Central Register of Controlled Trials (CENTRAL); MEDLINE; Embase; CINAHL; PsycINFO; Web of Science Core Collection; ClinicalTrials.gov; WHO ICTRP. We ran separate searches to identify RCTs and qualitative studies. We searched all sources for both sets of searches. Following the initial screen and study selection we used the search summary table format proposed by Bethel to record which database each study came from, calculate the overall coverage of each database, and identify any unique studies. We only checked for the retrieval of the primary publication, not for any additional publications related to the study.

Results: We found 17 RCTs and six qualitative studies for inclusion. The number of databases an RCT report was found in ranged from 1 to 6. CENTRAL had the highest overall coverage with 88% (n = 15) included studies. We did not find any of our included studies in CINAHL. We found unique studies in CENTRAL (N = 4); Cochrane Airways Trials Register (n = 1); and Embase (n = 1). For qualitative studies, the number of databases in which we found a report ranged from 1 to 5. Embase had the highest overall coverage with 100% (n = 6) of our included studies. We did not find any of the included studies in ClinicalTrials.gov or the WHO ICTRP. We found unique studies in Embase (n = 1).

Conclusions: We found differences in the best sources for retrieval of RCT reports and qualitative study reports. CENTRAL performed very well for retrieving reports of RCTs. All of our included qualitative studies were found in Embase. We will use the information in the search summary table to streamline the number of databases to search for future updates. We will continue to update the search summary table each time we re-run our searches for this review.

Patient or healthcare consumer involvement: The review was prioritized by a patient steering group.

Evidence syntheses for assessing the health impacts of climate change: the key role of developing search strategies

Thomson D[1], Dennett L[2], Noga A[3], Sebastianski M[3]

[1] Department of Pediatrics, University of Alberta, Canada; [2] John W. Scott Health Sciences Library, University of Alberta, Canada; [3] Alberta SPOR SUPPORT Unit Knowledge Translation Platform, Canada

Background: Background Climate change has been described as the greatest health threat of the 21st century. Adaptation involves adjusting to the current and predicted effects of climate change, to reduce vulnerability to harmful effects. Strategies for adaptation can have both direct and indirect health benefits. However, research to date includes limited recognition of health impacts and limited study incorporating a health perspective. Attributing health outcomes to adaptation interventions can be challenging for many reasons. Evidence synthesis collating data on the effectiveness of adaptation interventions is therefore of strong interest to the health community and the public at large. Effective search strategies are a cornerstone of high-quality syntheses so appropriate methods are crucial.

Objective: To present the challenges in creating a search strategy for a scoping review to evaluate the effectiveness of climate change adaptation measures in creating health impacts.

Methods: The primary research question is, 'Globally, what adaptation measures have been effective in reducing the negative impacts of climate change on human health?' We designed search strategies for OVID MEDLINE, OVID Embase, and Web of Science, as well as for numerous grey literature sources, EBSCO Open Dissertations and Google Scholar.

Results: Numerous iterations of the search strategy were required for this scoping review. Challenges included:

- developing a working definition and list of search terms to adequately cover the broad concept of 'health impacts';
- incorporating search terms for a wide range of study designs (not just clinical studies) that are measuring effectiveness of adaptation initiatives;
- encompassing the wide range of activities that can be considered as adaptation initiatives with direct or indirect health co-benefits;
- capturing studies that are relevant but that are not explicitly described as being related to climate change;
- teasing out health benefits from other social, economic and/or environmental benefits.

Conclusions: The impacts of climate change will be more deeply felt around the world in the coming decades, so it is essential to establish what is effective for protecting human health. Relevant interventions are planned, delivered and implemented in a highly complex space, given that the long-term impacts of climate change are still unknown and the huge variability of the social and economic systems in which the interventions are unfolding. Developing high-quality evidence syntheses about the health impacts of climate change will be an important contribution to global adaptation planning. Well-designed search strategies are a fundamental component of this work, and the complexities of designing them must be thoroughly explored.

Patient and healthcare consumer involvement: None.

Giving European SLRs the upper hand: are manual searches of local journals necessary?

Bobrowska A[1], Evans JS[1], Steeves S[1]

[1] Costello Medical, UK

Background: Literature reviews in medicine and healthcare are usually performed by searching English language databases such as MEDLINE and Embase. As a consequence, there may be a bias towards English language publications in literature reviews that have not supplemented such database searches with country or language-specific searches. A possible method of accounting for this includes handsearching of local journals.

Objectives: We investigated the extent to which handsearches of European journals might be advantageous alongside English language database searches.

Methods: We retrieved a list of journals categorized under the "Medicine" subject area in the Scimago Journal & Country Rank database and with a 2018 ranking for EU/EEA countries and Switzerland. For countries with more than 100 journals, we identified which journals were not indexed in MEDLINE. We analysed the country with the most non-indexed journals to see which subject categories those journals were assigned to. We then calculated the overall count for each category to determine whether any specific areas would be particularly susceptible to data being missed (in that country).

Results: We found seven countries with more than 100 journals: Netherlands (585), Germany (498), Switzerland (196), France (172), Italy (164), Spain (157) and Poland (117). The proportion of journals not indexed in MEDLINE ranged from 4% to 16% in all countries but Germany, where 33% (162/498) journals were not indexed in MEDLINE. The German non-indexed journals were classified under 38 different categories within the "Medicine" subject area (each journal could be assigned more than one category). The most represented individual categories were "Medicine (miscellaneous)" (40), "Surgery" (17) and "Public health, environmental and occupational health" (17). When grouped thematically, the most represented topic areas were public health (21), internal and emergency medicine (20), surgery (17), psychiatry (16) and orthopaedics/sports medicine (16). These non-indexed journals made up 29% to 56% of all German journals in these topic areas and 5% to 9% of all European Journals indexed in Scimago in these topic areas.

Limitations: Firstly, we based our list of journals on those indexed in Scimago, which may itself not represent all available journals. Secondly, our analysis did not consider the relative contribution that non-indexed journals make to the literature, either in terms of the number of articles published in these journals or their quality. Nevertheless, we suggest that our results give an indication of where it may be valuable to include additional journal handsearches in literature reviews.

Conclusions: Our results indicate that over most European journals, the majority are indexed in MEDLINE and failing to conduct country-level specific searches is in general unlikely to result in large numbers of articles being missed. However, specialized searches may be necessary where a review's scope is limited to certain countries or on topics where there is a higher proportion of non-indexed journals.

How do Cochrane Reviews use information from trial registers? Findings from a cross-sectional audit

Tan-Koay A[1], Berber S[1], Opiyo N[2], Glanville J[3], Dwan K[2], Lasserson T[2], Willson M[1]

[1] Cochrane Breast Cancer Group, The University of Sydney, Australia; [2] Editorial & Methods Department, Cochrane Central Executive, UK; [3] York Health Economics Consortium, University of York, UK

Background: Trial registers fulfil an important function by providing a transparent scientific record of clinical trials. Their role has been widely acknowledged as being a mechanism for reducing publication bias and detecting reporting issues around study outcomes by study investigators. We conducted this study to assess the extent to which Cochrane authors are incorporating this important source of trial information into their reviews and to determine whether guidance is required for Cochrane authors.

Objectives: To assess current Cochrane Review practice in identifying and incorporating information from clinical trial registers.

Methods: We conducted a cross-sectional study to assess a sample of new or updated intervention reviews from all Cochrane Review Groups up to February 2017. Two assessors independently extracted data from each review using a pre-tested audit questionnaire. We analysed data relating to the frequency of reporting (i) the source and search strategy; (ii) results of trial register searches; and (iii) the use of trial register information in the review.

Results: Over 90% (236/260) of Cochrane Reviews reported searching a trial register (e.g. ClinicalTrials.gov) or registers via the World Health Organization's International Clinical Trial Registry Platform (WHO ICTRP). In reviews that reported trial register searches, 39% (92/236) indicated the number of trial records retrieved and 56.7% (134/236) used information from the trial register record in the review. Trial record information was incorporated into the results (39.6%; 53/134), risk of bias assessments (53.7%; 72/134), discussion (24.6%, 33/134) and conclusion sections (25.4%, 34/134). In the discussion section, trial register information was used mostly to describe the overall completeness and applicability of the evidence (48.5%, 16/33), and potential biases in the review process (30.3%, 10/33). Table 1 provides examples of how trial register information was incorporated in the review.

Conclusions: The majority of audited reviews searched trial registers. Information identified from these sources was most commonly integrated into results, but there are opportunities to consider their relevance in formulating implications for research. Based on this audit, we suggest Cochrane authors may benefit from guidance on how to incorporate information from registers into Cochrane Reviews. This could be useful to inform decisions on planning additional primary research and updating systematic reviews.

Patient or healthcare consumer involvement: Although patients or healthcare consumers were not involved in this audit, the findings of this work indicate that some improvements are needed in the uptake of trial register information in Cochrane Reviews. By using trial register information more effectively, systematic reviews will be able to provide a more complete picture of the current and emerging evidence to its consumers.

Additional file: [Table](#)

Identifying digital intervention research: using bibliometric analyses to inform searching and mapping

Stansfield C[1], Burchett H[2]

[1] EPPI-Centre, University College London Institute of Education, UK; [2] London School of Hygiene & Tropical Medicine (LSHTM), UK

Background: Bibliometric analyses, such as word co-occurrence networks, topic modelling and research field networks can provide an overview of trends and patterns of a research area. Digital technology for promoting health behaviour change is a growing interdisciplinary field. Interactive digital interventions for health promotion and treatment use a range of modes such as smartphone apps, websites, videoconferencing, and websites. Systematically identifying research evidence within this area is challenging owing to the variety of digital modes and the heterogeneous terminology used to describe them. As part of undertaking a systematic map of research on digital interventions for prevention, treatment and recovery of alcohol and drug misuse, a readily available set of primary studies and systematic reviews provided an opportunity to explore how bibliometric tools could describe these studies in order to inform search term development and mapping.

Objectives: To explore ways that digital interventions have been described in primary studies and systematic reviews and their references and citation relations. To compare differences between them in order to explore the utility of highlighting the different conceptual areas within this domain to inform search strategies and mapping.

Methods: We analysed the sets of citations using text analysis, topic modelling and text visualization tools for keyword co-occurrence: 1241 primary studies and 23 systematic reviews and their references. The distribution of research areas of 22 systematic reviews and references were mapped using Web of Science research fields.

Results: We will present the trends and differences across the sets of citations and an overlay map of research areas. We will group key terms and phrases into broad categories to distinguish different facets of the digital components of the intervention: nature of technology (digital, electronic, voice recognition, automated); hardware (e.g. computer, smartphone); technology interface (internet, app, text-message); strategy of the technology intervention (eCBT, e-health).

Conclusions: This exploratory study informs the potential of applying bibliometric tools for identifying research described by diffuse terminology. It contributes to increasing the breadth of research evidence available to users of reviews, including healthcare consumers.

medrxivr: a new tool for searching for and retrieving records and PDFs from the medRxiv preprint repository

McGuinness L[1], Schmidt L[1]

[1] Department of Population Health Sciences, University of Bristol, UK

Background: The medRxiv [med-archive] repository, which hosts copies of health-related manuscripts uploaded prior to formal peer review and publication, represents a key source of grey literature for systematic reviewers. However, the current web-based interface to this repository presents some key challenges for systematic reviewers. Only relatively simple searches may be performed, and there is no option to bulk-download either the meta-data (e.g. title, abstract, subjects) or a copy of the full-text PDF for records identified by the search. We sought to address these issues via a new tool, medrxivr.

Objectives: To develop a new tool that allows users to search the medRxiv preprint repository data using complex search strategies, and to download metadata (e.g. title, abstract, authors, subject category) and full-text PDFs for records identified by their search.

Methods: The R programming environment was used to create a snapshot of the medRxiv preprint repository, and to develop a new R package, medrxivr, and associated web-based application that allow users to query the snapshot.

Results: The baseline snapshot of the medRxiv database was created in November 2019. This snapshot is updated daily to capture new records added to the repository. The medrxivr R package and associated web application (package bit.ly/medrixvr-package:, app: bit.ly/medrxivr-app) were made available in March 2020 and allow users with varying levels of ability in R to search the snapshot for relevant articles. Search strategies using Boolean logic (AND, OR, NOT) and regular expression syntax (e.g. "[Tt]est" finds both "Test", and "test") can both be used, while results can be filtered by date of publication. As records on medRxiv can be updated, users can also choose to retrieve all versions of a given record or only the most recent one. Identified records can then be passed to a helper function in the R package which will automatically download the full-text PDF of each record to aid with the full-text screening process.

Conclusions: medrxivr enables systematic reviewers to search for and retrieve relevant metadata and full-text PDFs for articles in the medRxiv preprint repository.

Patient or healthcare consumer involvement: As this was methods focused project, no patients or healthcare consumers were directly involved in the tool's development. However, a tool that helps users search (and retrieve data from) a health-focused preprint repository such as medRxiv will promote the production of systematic reviews that incorporate this source of grey literature, resulting in a more accurate and up-to-date summary of all available evidence, thus maximizing the relevance of these reviews to patients and the public.

Poor reporting of reasons for termination in clinical trials registered in the Clinical Trial Registry of India

Charan J[1], Dutta S[1], Kaur R[2], Bhardwaj P[1]

[1] All India Institute of Medical Sciences (AIIMS), Jodhpur, India; [2] SN Medical College, Jodhpur, India

Background: Out of all the new clinical trials initiated to explore the efficacy and safety of new interventions, 10% to 20% terminated prematurely. Looking at the altruistic motive of patients' involvement in these trials and the use of huge resources, premature termination of trials raises an important ethical issue. There may be some genuine reasons for premature termination of clinical trials and it is important to evaluate the reasons for termination of such trials to get an understanding of the current situation. There is a scarcity of data related to the reasons for termination of clinical trials being conducted in India.

Objectives: To evaluate the reporting of reasons for termination of clinical trials registered in the Clinical Trial Registry of India.

Methods: This study was based on a publicly available database, the Clinical Trial Registry of India (CTRI). We extracted from the CTRI data related to all the terminated trials. The data related to the trials' characteristics like type of intervention, phase of trial, sponsor, single centre/multicentre etc. Two investigators separately assessed the reasons for the termination of clinical trials and categorized them in subheadings. We resolved discrepancies by mutual consultation.

Results: We included in the analysis all clinical trials from the CTRI database from inception to 18 December 2019. Of the 16,579 clinical trials of interventions posted on the results database, 243 trials were listed as terminated. Out of 243 terminated trials, reasons for termination were given in only 50 (20.58%) of them. Out of these 50, 15 (30%) reasons were related to the scientific data, 27 (54%) were related to non-scientific issues and eight reasons were unclear. In reasons related to the scientific data, the most common reason was safety issues with the development of the product (14%) and among non-scientific issues, significant delay in recruitment was most prevalent reason (16%).

Conclusions: Reporting of reasons for termination of clinical trials registered in the Clinical Trial Registry of India is very poor. The administrators of the registry need to make appropriate changes in the system of registration so that most of the reasons can be reflected on the website for better transparency of data.

Patient or healthcare consumer involvement: This study is based on secondary data hence no patients or consumers were involved directly.

Additional file: Tables

Potential impact of English-language limits in systematic reviews

de Kock S[1], Noake C[1], Ross J[1], Stirk L[1], Duffy S[1], Misso K[1], Kleijnen J[1]

[1] Kleijnen Systematic Reviews Ltd, UK

Background: Despite the intention of a systematic review (SR) to identify and evaluate as much relevant evidence as possible with as little bias as possible, it is common for English-language limits to be applied at searching or screening stage.

Objectives: To see how many SRs which do not use English-language limits, go on to have non-English included studies. We want to investigate connections between an SR which does not impose language restrictions and its risk of bias in Domain 2 (D2) of the ROBIS (Risk of Bias) tool (Identification and Selection of Studies). We also want to see if there are patterns in relation to topic and the inclusion of non-English studies which might indicate when English-language limits would be inappropriate.

Methods: We propose taking a random sample of SRs published between 2016 and 2018 from KSR Evidence (a database of SRs) to investigate the use of English-language limits. We intend to analyse SRs which do not apply an English-language limit to see how many include non-English studies. We will investigate if there is a relationship between searching quality of SRs, as assessed by D2, and language limits. We also want to see if there are connections to topics which include non-English studies.

Results: In a pilot study, we analysed a random sample of 168 SRs published in 2016. Ninety-four SRs did not have English-language limits. Twenty-four (25%) of these SRs included at least one non-English study. Fifty-three (56%) SRs which did not apply English-language limits had a low risk of bias in D2. Conversely, 47/74 (64%) of studies which applied an English language limit were judged as high risk of bias in D2. The SRs which included non-English studies were varied in topic and we could not identify a connection to topic and the likelihood of including a non-English study.

Conclusions: Based on preliminary findings, we argue that if 25% of SRs which did not apply an English-language limit, went on to include at least one non-English study, the application of language limits means there is a risk of missing relevant evidence which, in the worst scenario, could change outcomes. At a minimum the inclusion of relevant non-English studies increases sample size and the certainty of outcomes. Preliminary findings suggest that topics which included non-English studies are varied and therefore English-language limits cannot be justified at searching stage. We conclude that SRs which did not apply a language limit are likely to have search methods which are low risk of bias and thus have searched widely with a comprehensive strategy. Findings may change.

Patient or healthcare consumer involvement: None.

Precision searching: an innovative search strategy for retrieving the newest and optimal systematic reviews from PubMed

Lin PC[1], Su YC[2], Hung PL[3], Tu SC[3]

[1] Department of Medical Education and Research, Kaohsiung Veterans General Hospital, Kaohsiung, Taiwan; Department of Pharmacy, Kaohsiung Medical University, Kaohsiung, Taiwan; [2] Department of Pharmacy, Kaohsiung Veterans General Hospital, Kaohsiung, Taiwan; Institute of Clinical Pharmacy and Pharmaceutical Sciences, College of Medicine, National Cheng-Kung University, Tainan, Taiwan; [3] Department of Pharmacy, Kaohsiung Veterans General Hospital, Kaohsiung, Taiwan

Background: Medical knowledge grows quickly and is often accessed via online databases such as PubMed to solve clinical questions. PubMed is the most popular database due to the characteristics of being easy to use, free and continuing to improve to meet user requirements especially in evidence-based medicine. Clinical queries are online search filters designed to improve the retrieval of relevant and evidence-based articles for the purpose categories of therapy, diagnosis, prognosis, etiology, clinical prediction guides and systematic review (SR) from the PubMed database. The research methodology filters use MeSH and tags to improve the sensitivity and specificity of search results. However, the usage of MeSH and tags may lose the newest evidence just supplied by publishers because the indexing is in process or because the data are not in MEDLINE format.

Objectives: To develop an innovative search filter for SR and compare sensitivity, specificity and precision to the clinical queries for SR in PubMed.

Methods: We created a customized search strategy to accommodate the newest not-indexed or in-process systematic reviews. We used COVID-19 as the reference to compare the sensitivity, specificity and precision of retrieval of SRs by clinical queries in PubMed or our customized search strategy because it was a trending topic recently. Sensitivity and specificity were defined as the proportion of SRs detected and non-SRs excluded by the given search filters, respectively. Precision was defined as the proportion of retrieved articles that are SRs.

Results: We searched for the text word COVID-19 in PubMed on 5 April 2020 and found 2575 articles. Fourteen (0.5%) and 25 (1.0%) articles were recognized as systematic reviews using PubMed clinical queries of SR and our customized filter, respectively. The PubMed search strategy for SR had a sensitivity of 52.0% (95% confidence interval (CI) 33.5% to 70.0%) and a specificity of 99.96% (95% CI 99.80% to 100%). Our customized search strategy for SR had a sensitivity of 96.0% (95% CI 80.5% to 99.3%) and a specificity of 99.96% (95% CI 99.80% to 100%). The precision of these two search strategies were 92.9% and 96.0%, respectively. Most of the missed articles (91.7%) by using PubMed clinical queries for SRs were retrieved by our customized search strategy of not-indexed filter.

Conclusions: Systematic reviews can be retrieved from PubMed by this innovative and optimal search strategy with perfect sensitivity or specificity, and high precision even for the newest topics or articles. Sometimes, the newest and not-indexed articles in PubMed may be missed if clinical queries or the "limit" function was used.

Patient or healthcare consumer involvement: Yes. The precision searching provided in this study can aid the retrieval of systematic reviews in PubMed. Clinical healthcare professionals can easily search the best evidence of systematic reviews for solving patient-centered questions.

Searching LILACS in Cochrane Reviews: methods used and user's feedback

Escobar Liquitay CM[1], Vergara-Merino L[2], Verdejo C[2], Kirmayr M[2], Schuller-Martínez B[2], Madrid E[3], Meza N[3], Pérez-Bracchiglione J[3], Franco JVA[1]

[1] Research Department, Cochrane Associate Centre, Italian Hospital University Institute, Argentina; [2] Medicine School, University of Valparaíso, Associated Cochrane Centre, Chile; [3] Interdisciplinary Centre for Health Studies (CIESAL), Cochrane Associate Centre, School of Medicine, University of Valparaíso, Chile

Background: LILACS, the main reference database in Latin America and the Caribbean, contains health literature published since 1982. It includes a network of collaborating centres, located across the region, which feed the database with locally produced scientific literature and co-ordinated by the Regional Library of Medicine (BIREME in Spanish). The LILACS methodology developed by BIREME brings together a set of standards, manuals, guidelines and applications, aimed at selecting, describing, indexing documents and, thereby, generating the database. However, the impact of the inclusion of LILACS in the search of Cochrane systematic reviews (CSR) remains unknown.

Objectives:

1) To estimate the proportion of CSR that used LILACS as a source of information during 2019. –
2) To analyse the methods used to search this database.
3) To explore the perceptions of the Cochrane community regarding LILACS.

Methods: We included CSR of interventions published in 2019. We inspected the search methods section for each CSR and selected those that specified LILACS as a source of information, including specialized registries. After selecting those reviews that mentioned LILACS in their methods, we looked whether their search strategy was available in the appendices. We extracted: search dates, use of basic/advanced search, use of free text and controlled language (DECS), languages (English +/- other languages), use of filters (web-based or validated filter) Furthermore, we will circulate a survey among key stakeholders: Information Specialist / Assistant Information Specialist / Medical Librarian, Cochrane Review author, non-Cochrane systematic review author, Cochrane Review Group Editor. The survey includes descriptive data, perceptions regarding what they expect from a database during a search, recovery and information management, the reasons for using LILACS and perception regarding documents retrieval using LILACS.

Results: We analysed 545 CSR (see Figure 1); only 182 CSR (33%) included LILACS as a source of information; 36 of those retrieved articles through the Cochrane Review Group specialized registry and the remaining 146 conducted independent searches in the database (13 of which did not report the search strategy). An analysis of the 133 reported search strategies indicates that there is substantial heterogeneity in the use of advanced search methods (controlled language, truncation, etc.). The survey was approved by the Cochrane Knowledge Translation Department and will be sent to the above-described community and will be submitted to the institutional review board.

Conclusions: The Methodological Expectations of the Cochrane Intervention Reviews (MECIR) considers LILACS as a regional database with a highly desirable standard of consultation. This study contributes to the understanding regarding the use of this resource. This input will allow the Cochrane community to establish explicit criteria and guidance on when and how to use LILACS.

Patient or healthcare consumer involvement: None

Additional file: Figure 1

Synthesized HIV/AIDS research evidence (SHARE): an online database to capture, link, and share current HIV-related systematic reviews

Gogolishvili D[1], Andruszkiewicz N[1], Giliauskas D[1], Wasdell M[1]

[1] The Ontario HIV Treatment Network (OHTN), Canada

Background: The use of high-quality research evidence is necessary to support human immunodeficiency virus (HIV) prevention efforts and treatments. Timely use of evidence can improve the experiences and health outcomes of HIV positive individuals. However, individual studies are often not enough to drive policy or treatment changes, as only the strongest study designs (e.g. systematic reviews) can yield robust recommendations. Stakeholders often lack the capacity to navigate research databases and rapidly identify relevant research evidence to support their initiatives. Effectively connecting community organizations, researchers, and policymakers to high-quality research evidence in a timely manner is imperative in the development of evidence-informed programs and policy change towards the end of the HIV epidemic.

Objectives: To describe the features of a free online searchable database of HIV-related systematic reviews called Synthesized HIV/AIDS Research Evidence (SHARE).

Methods: The Ontario HIV Treatment Network (OHTN) has developed SHARE, an online database which provides access to timely high-quality HIV research for stakeholders available at www.hivevidence.org. A search of MEDLINE/PubMed and the Cochrane Library is conducted quarterly to locate articles for screening and inclusion. Reviewers then categorize included articles according to four HIV-relevant domains: population(s) of interest; prevention, engagement and treatment cascade; health systems; and syndemics and determinants of health. SHARE also includes rapid response reviews on novel topics produced by the OHTN.

Results: As of 1 February 2020, we have screened over 43,000 articles for SHARE with 3225 peer-reviewed systematic reviews and rapid responses categorized and included in the database. Articles focus on 15 population groups and 29 different syndemic and social determinant categories. A total of 72% of articles (n = 2,322) relate to the prevention, engagement, and treatment cascade. We send email blasts with the latest updates made to the database quarterly to stakeholders which include 2364 HIV-related systematic review authors, 502 librarians and listservs, and 112 HIV research organizations worldwide. Since March 2017 (when tracking of website visitors using Google Analytics was initiated), there have been 2383 unique visitors to the database from 70 countries (Figure 1).

Conclusions: The SHARE database has the potential to support research, decision-making, and rapid learning related to HIV/AIDS at all levels of the healthcare system by disseminating high-quality and timely HIV relevant research to stakeholders worldwide. Efforts to promote the SHARE database are being strengthened to continuously expand its reach and to ensure it remains useful for HIV/AIDS researchers, policymakers, and community organizations.

Patient or healthcare consumer involvement: The database aims to support evidence-informed decision making to improve the lives of those living with and at risk of HIV.

Additional file: Figure 1

Testing the sensitivity and precision of the Cochrane MEDLINE randomized controlled trial search filters

Glanville J[1], Kotas E[1], Featherstone R[2], Dooley G[3]

[1] York Health Economics Consortium, UK; [2] Cochrane Editorial and Methods Department, UK; [3] Metaxis Ltd, UK

Background: The Cochrane Handbook for Systematic Reviews of Interventions contains two search filters to find randomized controlled trials (RCT) in Ovid MEDLINE: a sensitivity maximizing RCT filter and a sensitivity and precision maximizing RCT filter. The RCT search strategies were originally published in 1994 and have been adapted and updated, most recently in 2008.

Objectives: To determine whether the Cochrane filters are still performing adequately to inform Cochrane Reviews, we tested the performance of the two Cochrane RCT filters and 36 other MEDLINE filters in a very large gold standard set of RCT records.

Methods: We identified a gold standard set of 27,617 RCT reports published in 2016 from the Cochrane Central Register of Controlled Trials (CENTRAL) database. We then retrieved these RCT records in Ovid MEDLINE using PubMed identifiers. Each RCT filter was run in MEDLINE and combined with gold standard set of records, to determine its sensitivity, precision and f-score.

Results: We tested each filter against the gold standard on 16 July 2019. The most sensitive RCT filter was Duggan (sensitivity 0.99). The Cochrane sensitivity maximizing RCT filter had a sensitivity of 0.96, but was more precise than Duggan (0.14 compared to 0.04 for Duggan). The most precise RCT filters were Chow, Glanville/Lefebvre, Royle/Waugh, Dumbrique (precision 0.97, sensitivity 0.83). The best precision Cochrane filter was the sensitivity and precision maximizing RCT filter (precision 0.46).

Conclusions: This study used a very large gold standard to compare the performance of all known RCT filters. We concluded that the Cochrane MEDLINE sensitivity maximizing RCT filter can continue to be used with confidence by Cochrane Review authors and CENTRAL compilers. Slightly more sensitive filters are available, but with lower precision. Using the information from this research, searchers can make better informed decisions about which filters to use for their own searches.

Patient or healthcare consumer involvement: This was a searching methods investigation and no patients or healthcare consumers were involved.

The RobotSearch RCT classifier: ready for use?

Hausner E[1], Sow D[1], Waffenschmidt S[1]

[1] Institute for Quality and Efficiency in Health Care (IQWiG), Germany

Background: The option of automating individual working steps when creating systematic reviews has been discussed for many years (1,2). The use of machine-learning algorithms for the identification of randomized controlled trials (RCTs) appears to be particularly promising. The publication by Marshall et al. (3), for example, shows that the RCT classifier provided by RobotSearch (4) can be reliably used (in terms of sensitivity) instead of conventional database search filters. But what does this mean in practice?

Objectives: To determine whether the use of an RCT classifier could be integrated into the standard working process at the German Institute for Quality and Efficiency in Health Care (IQWiG).

Methods: In two projects (1. Biomarkers in breast cancer; 2. Osteoporosis drugs), we compared the results yielded using the RCT classifier (sensitive model) with those yielded using our standard working process (see Table 1) with conventional search filters. In addition, we tested the practical applicability of the RCT classifier in our daily project work.

Results: Using the sensitive CHSSS filter in MEDLINE and an optimized filter in Embase for Project 1 (non-drug assessment), the number of hits with the RCT classifier was lower than the number with the standard process. However, using the optimized filters for MEDLINE and Embase for Project 2 (drug assessment), the number of hits increased considerably with the RCT classifier (4803 vs. 2589 in MEDLINE). Applying the standard process together with the RCT classifier would have significantly reduced the number of hits (- 29%); see Table 2. In addition, in the practical application of the RCT classifier, a number of new issues arose: longer processing times due to a high number of hits; higher workload due to increased documentation and error-proneness due to media disruption; and less information in the data file due to the Research Information System (RIS) format required.

Table 1. Standard use of study filters at IQWiG

Standard process	MEDLINE	Embase	Central
Drug assessment	CHSSS sensitivity- and precision-maximizing version [5]	Wong strategy, minimizing difference between sensitivity and specificity [6]	No study filter applied
Non-drug assessment	CHSSS sensitivity maximizing version [5]		

Table 2. References to be screened (after check for duplicates)

	Number of references in MEDLINE, Embase and Central to be screened (after check for duplicates)		
	Standard process	RCT classifier	Standard process + RCT classifier
Project 1	325	288	167
Project 2	3938	6342	2781

Conclusions: We explicitly support the automation of work processes and the RobotSearch interface is easy to use. However, routine use of the RCT classifier at IQWiG is unlikely in the near future, as the advantages (e.g. fewer hits to screen) do not seem to outweigh the disadvantages. Further testing of the reliability of the RCT classifier as a complementary tool to the standard study filters would be useful.

Patient or healthcare consumer involvement: None.

References:

1. Tsafnat G. Systematic review automation technologies. *Systematic Reviews* 2014; 3:74.
2. Beller E. Making progress with the automation of systematic reviews. *Systematic Reviews* 2018; 7(1):77.
3. Marshall I. Machine learning for identifying randomized controlled trials: an evaluation and practitioner's guide. *Research Synthesis Methods* 2018;9(4):602-14.
4. RobotSearch RCT classifier. https://robotsearch.vortext.systems.

Updating recommendations in guidelines using new tools. Which approach is more efficient?

Torres LA[1], Rojas MX[1], Gonzáles L[1], Reyes D[1], Barrera E[1], Mancilla N[2]

[1] Fundación Cardioinfantil – Instituto de Cardiología, Colombia; [2] Facultad de medicina- Universidad del Rosario, Colombia

Background: A major challenge in updating clinical practice guidelines (CPGs) is to efficiently identify new, relevant evidence. As part of a CPG implementation project for non-communicable diseases (NCDs) in Colombia, we evaluated the efficiency and feasibility of three approaches to identifying up-to-date literature. This abstract presents the preliminary results of an updating process followed for the selected recommendations from the type 2 diabetes mellitus (T2DM) Colombian CPG.

Objectives: To identify, among three different approaches, which resulted in 1) less time consuming and 2) more accurate results (i.e. identifying relevant up-to-date evidence) using fewer steps.

Methods: We compared three approaches: (1) based on PubMed for MEDLINE: a restrictive search strategy using the minimum number of Medical Subject Headings (MeSH) terms and text words required from the original search strategies published in the original CPG, plus a narrow filter for systematic review (SR) identification; (2) based on the Epistemonikos database: a broad search strategy using only population and intervention key words; and (3) based on the recently launched new Living overview of evidence (L·OVE) platform which uses artificial intelligence searching: a revision of SR references included in the correspondent PICO-specific T2DM-L·OVE platform.

Two people independently ran searches and applied predefined selection criteria following strategies 1 and 2, a third person performed approach 3. We compared the number of references retrieved and the number of steps required to find out relevant up-to-date evidence. In case of identifying no up-to-date SRs relevant for recommendation updating, we plan to search the Cochrane Central Register of Controlled Trials (CENTRAL) for Randomized Controlled Trials (RCTs) and compare these results with the primary studies references included in the T2DM-L·OVE platform.

Results: We updated searches for a total of eight recommendations; four recommendations had original searches published in the CPG. Approach 2 retrieved fewer references than approach 1. Additionally, approach 2 identified the majority of key references for the recommendation updating. Compared to approach 3, all key new SR identified with 1 or 2 were included in the corresponding PICO-specific T2DM-L·OVE. Approach 3 allowed fewer steps as it was not necessary to run searches. However, the new L·OVE platform is still completing the reference classification process, therefore several references included under each PICO set were not specifically related to the PICO. In these preliminary results we did not need to search for new RCTs, as we found key SR in all cases.

Patient or healthcare consumer involvement: The L·OVE platform-based approach promises to be a very efficient way to update CPG recommendations. Healthcare consumer involvement is essential to support this collaborative project that could greatly facilitate the task of updating CPGs.

Experience of public involvement in the implementation process of guidelines for non-communicable diseases

Torres LA[1], Barrera E[1], Muñoz AM[1], Rojas MX[1]

[1] Fundación Cardioinfantil – Instituto de Cardiología, Colombia

Background: Non-communicable diseases (NCDs), including heart disease, stroke, cancer, type 2 diabetes mellitus (T2DM) and chronic lung disease, are responsible for almost 70% of deaths worldwide. In Colombia, more than 110,000 people die due to NCDs each year. Adequate self-management presents a potentially scalable way of mitigating the growing burden of NCDs. As part of an NCD clinical practice guidelines (CPG) implementation project in Colombia, we seek to identify strategies that allow the patient to be empowered by current recommendations about their disease. This abstract presents our experience with patient/caregiver/physician involvement in CPG implementation.

Objectives: To assess the patient/caregiver/physician involvement strategy as part of the NCDs guideline implementation process.

Methods: Before starting the CPG implementation process, we decided to prioritize those recommendations that required special efforts to be implemented. Criteria for priority were: recommendations based on high quality evidence, potential feasibility, and low frequency of implementation. The study team selected the recommendations based on the first criteria. To apply all other criteria, we involved expert physicians, patients and caregivers. We invited patients/caregivers using local media campaigns and presentations in patient groups. Physicians were invited by personal communication considering their background in internal medicine and ambulatory practice. Criteria for participants were: being a patient or caregiver for one of the NCDs of interest (high blood pressure (HBP) or T2DM); having the time for completing the project activities; and giving informed consent. Participants attended a four-session training workshop in the fundamentals of evidence-based practice; different sessions were planned for patients/caregivers and for physicians. Participants completing the workshop successfully took part of one of the two selection panels (one panel for each condition). Panels met three times with specific objectives: (1) to evaluate the understanding of the recommendation; (2) to introduce modifications in the presentation and wording of the recommendation according to public perceptions and suggestions; and (3) to approve the final version of modified recommendations.

The experience was assessed using questionnaires completed by all types of participants at the end of the process.

Results: HBP and T2DM national guidelines present a total of 68 recommendations, 40 based on high quality evidence. A total of 12 patients, 8 caregivers and 15 physicians were involved. Twenty-one people completed the training workshops and took part in the selection panels. The study results included: demographic characteristics of participants, type of contributions made by patients/caregivers/physicians, perceptions of feasibility, suggestions for making recommendations clearer and more understandable from the patient/caregiver/physician perspective.

Patient or healthcare consumer involvement: Early public involvement seems to be a strategy to achieve better results in the local implementation process of CPGs.

Using Microsoft Academic Graph in a rapid review of reviews on vaccine uptake

Stansfield C[1]

[1] EPPI-Centre, University College London Institute of Education, UK

Background: Review databases are promising sources for rapid reviews of reviews to provide a breadth of studies with a relatively low yield of results. A novel option is Microsoft Academic Graph (MAG) to identify related publications and citation searching of relevant studies. A rapid map of systematic reviews of interventions and views on vaccine uptake provided an opportunity to analyse the utility of sources. Searches were undertaken in Epistemonikos, specialist registers, PubMed, CINAHL, Social Sciences Citation Index, Microsoft Academic Graph and other sources.

Objectives: To explore the utility of using Epistemonikos combined with Microsoft Academic Graph for identifying reviews. Methods Retrospective analysis of sources used in the map of reviews. All sources apart from MAG were searched using Boolean searching techniques and 1438 citations were screened on title and abstract after duplicate removal. MAG was searched in a beta interface within EPPI-Reviewer, using 178 relevant reviews identified from Epistemonikos as seed studies to identify related publications and bi-directional citation links. The results from MAG were filtered by searching within the title and abstracts, followed by screening on title. All reviews included in the map, that were additional to those identified in Epistemonikos, were checked to determine if present in the MAG results. The analysis of reviews identified from the other sources were based on the initial searches used to populate the map.

Results: A total of 197 reviews were included in the map, of which 154 (78%) were identified in Epistemonikos, after a date threshold was applied. Of the remaining 43 reviews: 4 were uniquely identified by MAG through the related publication feature, of which one was also found from citation links. A further 18 identified from MAG were also identified in other resources, though 2 were missed from the initial filtering process. The remaining 21 reviews were only identified from the other sources, including 10 from the SYSVAC register and 7 each from other review registers and bibliographic databases (each category provided some unique citations).

Conclusions: While Epistemonikos provided the majority of reviews, there was benefit in searching other sources as time allowed. MAG offers a promising approach to identify new studies. However, it required a two-step approach as it uses relevant seed studies from screening the initial searches, plus further deduplication steps. Filtering approaches are needed to make the MAG results manageable to screen.

Patient or healthcare consumer involvement: Efficient ways to expand the breadth of relevant search results improves the research available to stakeholders including healthcare consumers.

Using text-mining to develop a US-specific geographic search filter to facilitate systematic reviews in Ovid MEDLINE

Cheung A[1], Popoff E[1], Szabo SM[1]

[1] Broadstreet Health Economics and Outcomes Research, Canada

Background: Bibliographic databases like MEDLINE are crucial for healthcare researchers to access the latest evidence. As such databases index an ever-increasing volume of research, tools supporting information retrieval are valuable for identifying relevant evidence efficiently. Geographic search filters have been developed for jurisdictions such as the United Kingdom, Spain, and Africa.

Objectives: To develop a geographic search filter for accurately identifying research from the United States (US) in Ovid MEDLINE.

Methods: Citations indexed in MEDLINE were collected from bibliographies of reviews by the US Preventive Services Task Force, which publishes evidence-based recommendations in various disease areas. An algorithm was developed to select US-based publications meeting ≥ 2 of the following 3 criteria – US-based: author affiliation, place of publishing, or grant funding. Text mining identified one- and two-word terms in title/abstract fields, and the frequency compared between US and non-US citations. We used the findings to develop a preliminary search filter. We performed analyses in R.

Results: We collected 22,280 citations, of which 8243 were US-based according to the algorithm. US citations were published between 1980 and 2019; therapeutic areas included cardiovascular disease (9.9%), obesity (6.5%), and HIV infection (5.0%). Common US-related terms included (expressed as ratio of frequency in US to non-US citations) US cities/states/regions ("Pennsylvania" (64.0), "Miami" (26.3), "midwest" (23.0)), and words related to US populations ("African American" (22.2), "Medicare beneficiaries" (14.0)). The search filter was developed by combining these and other key terms in title/abstract fields (Table 1).

Conclusions: This development of a MEDLINE-based search filter will streamline the systematic identification of evidence from US studies. As the validity of the filter would be impacted by changes in controlled vocabulary in MEDLINE, periodic updates will be necessary. The above algorithm assumes that publications meeting at least 2/3 of the stated criteria are US-based; although this was not formally tested, an audit confirmed the relevance of select publications. Future work will include validation of the filter and refinement to improve sensitivity/specificity.

Patient or healthcare consumer involvement: This search filter will allow healthcare researchers to access information driving evidence-based decision making in a more targeted and efficient manner.

Additional file: Preliminary US search filter for Ovid MEDLINE

Validating Cochrane Neonatal's standard search databases: is it okay to stop searching Embase?

Ovelman C[1], Eckert C[2], Friesen C[3]

[1] Cochrane Neonatal, USA; [2] University of Vermont, USA; [3] Cochrane Neonatal, Canada

Background: Methodological Expectations of Cochrane Intervention Reviews (MECIR) standards require Cochrane Reviews to search Cochrane's Central Register of Controlled Trials (CENTRAL), MEDLINE and Embase (if Embase is available). In July 2019, Cochrane Neonatal (CN) lost access to Embase. Previously, we searched MEDLINE, Embase, CENTRAL and CINAHL for all intervention reviews. CENTRAL is comprised of randomized controlled trials (RCT) and controlled clinical trials (CCT) from Embase and MEDLINE, identified via Cochrane's centralized search and screening processes. Cochrane Information Specialists add included RCTs from their reviews to CENTRAL. CN added included RCTs to CENTRAL through December 2017. For intervention reviews, CN uses these standard search methods: RCT filters created using Cochrane's highly sensitive search strategies for identifying RCTs and neonatal filters created and tested for each database by our Information Specialist.

Objectives: to assess the effectiveness of conducting searches for CN intervention reviews in MEDLINE, CENTRAL and CINAHL, without searching Embase.

Methods: We searched for all new and updated CN intervention reviews published from January 2018 through April 2019. For new reviews, we compiled a list of included studies. For updated reviews, we compiled a list of only new studies added in the current update. If an included study had multiple publications, we only included the primary publication. We then used our standard search methods to search for each study in MEDLINE and CENTRAL. When a study was not identified in MEDLINE or CENTRAL, we searched Embase and CINAHL.

Results: We found 13 new intervention reviews. One new review was outside the neonatal scope and was excluded. The remaining 12 new reviews had 64 included studies. We found 17 updated intervention reviews, with 66 new studies included in the recent updates. The new and updated reviews included only RCTs and CCTs, and represented 12 of 24 active neonatal topic areas. For the 130 included studies we tested, 4 did not appear in either MEDLINE or CENTRAL. Of these 4 studies: none were found in Embase; 1 study appeared only in CINAHL; 2 studies were conference abstracts and were not identified in any of the databases searched; 1 study was in a journal not indexed by MEDLINE, Embase, or CINAHL. This study had a conference abstract that was identified in CENTRAL.

Conclusions: For new and updated neonatal intervention reviews, which plan to include only RCTs and CCTs, it is adequate to conduct searches in: MEDLINE, CENTRAL and CINAHL, without additionally searching Embase. Authors should search the reference lists of any studies selected for inclusion in their reviews in order to identify additional relevant articles. Searches for conference abstracts could be conducted separately. Authors may consider if searching CINAHL is necessary for their review topic.

Patient or healthcare consumer involvement: Healthcare consumers contribute to the creation of CENTRAL through Cochrane Crowd.

INVESITIGATING DIFFERENT TYPES OF BIAS

A tool to assess Risk Of Bias In Non-randomized Studies – of Exposures (ROBINS-E)

Morgan R[1], Taylor K[2], Higgins J[3], Rooney A[2], Thayer K[4], Schünemann H[1], Sterne J[3]

[1] McMaster University, Canada; [2] National Institutes of Health, USA; [3] University of Bristol, UK; [4] Environmental Protection Agency, USA

Background: Systematic reviews should include rigorous risk-of-bias assessments of eligible studies. Although a large majority of Cochrane Reviews address the effects of interventions, there is increasing interest in systematic reviews of other types of problem. One important category is systematic reviews on the effects of environmental and occupational exposures, for which eligible studies are almost always observational in nature. The ROBINS-I tool (1) has become widely used to assess the risk of bias in non-randomized studies of interventions but is not appropriate for studies of exposures.

Objectives: To develop a new tool for assessing risk of bias in non-randomized studies of exposures, for application to studies of environmental and occupational exposure.

Methods: Between 2017 and 2020 a workgroup of methodologists, content experts and systematic review authors with experience of relevant topic areas adapted ROBINS-I to create a new tool, ROBINS-E, for assessing risk of bias in non-randomized studies of exposures other than interventions, including environmental and occupational exposures. The preliminary instrument underwent piloting during a meeting in Bristol, UK, in October 2019.

Results: ROBINS-E features preliminary considerations of risk of bias within the review protocol; specification of a 'target experiment'; use of signaling questions leading to algorithm-guided risk-of-bias judgements; assessment of bias within seven domains (confounding, measurement of exposure, selection of participants into the study, post-exposure interventions, missing data, measurement of outcomes, selection of reported results); and derivation of an overall risk-of-bias judgement for the assessed result. Pilot testing identified areas requiring additional refinement, including signaling questions within the domains of confounding, exposure assessment, and selection of participants into the study, to expand on the concept of 'post-exposure interventions', and modifications to the algorithm-guided risk-ofbias judgements. Piloting served as a mechanism to collect examples of instrument application across a wide range of environmental and occupational health topics.

Conclusions: Proposed changes to ROBINS-I will make ROBINS-E suitable for assessing studies of exposures. Further revisions will incorporate feedback from pilot testing and application across a spectrum of environmental and occupational health topics.

Patient or healthcare consumer involvement: No

Reference:

1. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016;355:i4919.

Active placebo versus standard placebo control interventions in pharmacological randomized trials: a systematic review

Laursen DRT[1], Hansen C[1], Paludan-Müller AS[2], Hróbjartsson A[1]

[1] Centre for Evidence-Based Medicine Odense (CEBMO), Odense University Hospital, Denmark; [2] Nordic Cochrane Centre, Rigshospitalet, Denmark

Background: Placebos are used as control interventions in randomized trials to enable blinding of participants and personnel. This can help reduce the risk of bias due to, for example, observer bias, reporting bias and bias related to placebo effects. Standard placebos in pharmacological trials imitate the appearance of the experimental intervention, without containing the active component under investigation. They can be, for instance, saline infusions or lactose tablets. However, the experimental drug may have perceptible adverse effects or psychotropic effects that can lead to unblinding, and thus cause bias. This is suspected to be a potential issue with tricyclic antidepressants and selective serotonin reuptake inhibitors (SSRIs) for depression, as well as methylphenidate for attention deficit hyperactivity disorder (ADHD). As a potential solution, some trials have employed 'active placebo' as a control intervention instead of standard placebo. The active placebo contains a substance designed to imitate some of the physiological effects of the experimental treatment under investigation, but without being therapeutically active. One example is atropine as an active placebo for tricyclic antidepressants. In theory, the use of active placebos as control interventions can reduce the risk of bias due to unblinding, but this has yet to be investigated systematically.

Objectives: To estimate the impact on treatment effects in pharmacological randomized trials when using active placebo as compared to standard placebo.

Methods: A systematic review of randomized trials with both standard and active placebo control groups. We will search for eligible randomized trials in PubMed, CENTRAL and Embase, as well as Google Scholar, reference lists and citations of relevant papers. For included trials, we will extract basic trial data and outcome results. From each trial, we will extract one of each of the following outcome types: a patient-reported outcome (preferably continuous), an observer-reported outcome (preferably continuous), a dichotomous harms outcome. We will select two time points for each outcome: earliest post-treatment and latest follow-up. Also, we will assess attrition rate and co-intervention use. Based on the selected outcomes above, we will convert the outcome data to standardized mean differences (SMD) for active placebo versus standard placebo. We will then summarize the individual trials' estimates for each outcome type in random-effects inverse-variance meta-analyses. The primary analysis is the analysis for patient-reported outcomes at earliest post-treatment. The remaining analyses are secondary analyses.

Results and conclusions: Will be presented at the Colloquium.

Patient or healthcare consumer involvement: None

Artificially defined transformation methods for obtaining dose points in dose–response meta-analysis should be more standardized: a cross-section study

Lao Y[1], Lu Z[2], Yong R[2], Wei Q[2], Ma Y[1], Wang Y[2], Fan J[1], He G[1], Yao L[3], Yang K[4]

[1] Second Clinical Medical College, Lanzhou University, Lanzhou, China; [2] First Clinical Medical College, Lanzhou University, Lanzhou, China; [3] Health Research Methodology, Department of Health Research Methods, Evidence and Impact, McMaster University, China; [4] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China

Background: Dose–response meta-analyses (DRMA) have been published increasingly over the past few years. Their basic purpose is to reveal the relationship between disease risk and exposure dose. In order to meet the data condition for dose–response estimation, the open-ended dose intervals reported in original studies should be transformed to a certain dose point. And multiple units of exposure/intervention reported in original studies are also often unified to one. However, the transformation methods artificially defined by the review authors of DRMA were not always the same and this might cause misleading results.

Objectives: To investigate how DRMA authors defined transformation methods to obtain dose points for dose–response estimation.

Methods: We searched PubMed for meta-analyses published in 2019 that explicitly combined dose–response estimates from multiple original studies and reported the results of dose–response meta-analyses. Paired authors selected eligible studies and extracted related information independently. We resolved disagreements between two authors by discussion with a third author.

Results: We included 247 DRMA. The main outcomes were the risk of cancer and cardiovascular disease. As for open-ended interval, 32.8% studies reported methods to obtain the lower boundary of the lowest category. The two most common statements were: width of the open-ended interval was the same as adjacent category (58%) and the lowest category was assumed to zero (29.6%). Additionally, 4.9% articles assumed the lowest boundary to a certain nonzero dose based on specific clinical background. In total, 7.5% studies reported another six different methods to obtain the lower boundary. Furthermore, 39.3% of studies reported methods to obtain the higher boundary of the highest open-ended category. In this case, 71.2% studies assumed the width of the highest category based on different multiples of adjacent category: equality (65%), 1.5 times (5.2%), or half (1%). 27.8% studies assumed the higher boundary of the highest category based on different multiples of the lower bound of the same interval: 1.2 times (13.4%), 1.5 times (9.3%), 1.25 times (2.1%), 1.4 times (1%), 2 times (1%), and equality (1%). Another 1% of articles reported a certain nonzero dose based on the specific clinical background. As for unifying multiple units of reported in original studies, 14.6% of studies reported transform methods, among which 47.2% intervention/exposure were diet-related factors. In studies considering certain diet related factors, 77.3% of studies involved transformation methods between nine different units and "gram" while 22.7% studies reported methods between "cup" and "ml". But these methods are heterogeneous from each other. The overall results were presented in Figure 1.

Conclusions: Artificial defined transformation methods were frequently used for obtaining dose points in DRMAs. These methods were heterogenous which might cause misleading results and they should therefore be more standardized.

Patient or healthcare consumer involvement: None.

Additional file: Figure 1

Assessing risk of performance and detection bias in Cochrane Reviews as a single domain is less accurate compared to assessment of two separate domains

Barcot O[1], Boric M[1], Dosenovic S[2], Puljak L[3]

[1] Department of Surgery, University Hospital Split, Split, Croatia; [2] Department of Anesthesiology and Intensive Care, University Hospital Split, Croatia; [3] Center for Evidence-Based Medicine and Health Care, Catholic University of Croatia, Croatia

Background: The initial version of the Cochrane risk-of-bias (RoB) tool included one joint domain for "blinding of participants, personnel and outcome assessors". In the 2011 version of the tool, this domain was split into two domains: blinding of participants and personnel (performance bias) and blinding of outcome assessors (detection bias). There are no studies that have compared whether the decision to split the domain was justified.

Objectives: We analyzed prevalence of usage of the joint blinding domain, and proportion of inadequate assessments made in the joint versus single RoB domains for blinding in Cochrane Reviews.

Methods: This was a primary methodological study, in which we analyzed the methodology of Cochrane Reviews published in 2016/2017 in the Cochrane Database of Systematic Reviews (CDSR). From Cochrane Reviews we extracted information about assessment of blinding from RoB tables, including judgment and comment. For data extraction we used a custom-designed software. We assessed prevalence of using joint blinding domain for performance and detection bias, and split domain where performance bias and detection bias were analyzed separately. We assessed whether judgments made by Cochrane authors were adequate by comparing judgments and comments with instructions from the Cochrane Handbook for Systematic Reviews of Interventions. We compared frequency of inadequate judgments in reviews with two separate domains for blinding versus those with a single domain for blinding.

Results: We analyzed 728 Cochrane Reviews, with 10,523 trials included. In those Cochrane Reviews, we found 6918 assessments for performance bias, 8656 for detection bias, and 3228 for the joint domain. Prevalence of adequate assessments was 73% for performance bias, 78% for detection bias and 59% for the joint domain. The lowest prevalence of adequate assessments was found when Cochrane authors made judgment of low risk: 47% in performance bias, 62% in detection bias and 31% in joint domain.

Conclusions: The decision to split the single RoB domain for blinding of key individuals was justified, as Cochrane authors more frequently make adequate judgments with a split domain for blinding. In Cochrane RoB 2.0 tool, blinding is assessed in three separate domains. We anticipate that this should result in even higher adequacy of judgments of blinding of key individuals in Cochrane RoB 2.0 tool, but this will need to be confirmed after its full implementation in Cochrane Reviews.

Patient or healthcare consumer involvement: This was a research methodology study and it did not include patient or healthcare consumers.

Assessment of language bias and indexing bias among Chinese-sponsored randomized controlled trials

Jia Y[1], Huang D[1], Wen J[1], Wang Y[1], Rosman L[2], Chen Q[3], Robinson K[4], Gagnier J[5], Ehrhardt S[1], Celentano D[1]

[1] Bloomberg School of Public Health, the Johns Hopkins University, USA; [2] Welch Medical Library, the Johns Hopkins University, USA; [3] Institute of Information & Medical Library, Peking Union Medical College & Chinese Academy of Medical Sciences, China; [4] School of Medicine, the Johns Hopkins University, USA; [5] Michigan Medicine, School of Public Health, University of Michigan, USA

Background: Language bias and indexing bias may exist among Chinese-sponsored randomized controlled trials (CS-RCTs). Such bias may threaten the validity of systematic reviews.

Objectives: To evaluate the existence of language bias and indexing bias among CS-RCTs on drug interventions.

Methods: In this retrospective cohort study we retrieved eligible CS-RCTs from trial registries and searched bibliographic databases to determine their publication status. The search and analysis were conducted from March to August 2019. Trial registries included the primary trial registries recognized by the World Health Organization (WHO) and the Drug Clinical Trial Registry Platform (DCTRP) sponsored by the China Food and Drug Administration (China FDA). Eligible CS-RCTs were on drug interventions and conducted between January 2008 and December 2014. Exposure was defined as individual CS-RCTs with positive (versus negative) results. For assessing language bias, the main outcome was the language of the journal where CS-RCTs were published (English versus Chinese). For indexing bias, the main outcome was the language of bibliographic database where the CS-RCTs were indexed (English versus Chinese).

Results: We identified 891 eligible CS-RCTs. Four hundred and seventy CS-RCTs were published by August 2019, of which 368 (78.3%) were published in English. Among CS-RCTs registered in the Chinese Clinical Trial Registry (ChiCTR), positive CS-RCTs were 3.92 (95% confidence interval (CI) 2.20 to 7.00) times more likely to be published in English than negative CS-RCTs; among CS-RCTs registered in English registries, positive CS-RCTs were 3.22 (95% CI 1.34 to 7.78) times more likely to be published in English than negative CS-RCTs. These findings suggest the existence of language bias. Among CS-RCTs registered in ChiCTR, positive CS-RCTs were 2.89 (95% CI 1.55 to 5.40) times more likely to be indexed in EBDs than negative CS-RCTs; among CS-RCTs registered in English registries, positive CS-RCTs were 2.19 (95% CI 0.82 to 5.82) times more likely to be indexed in EBDs than negative CS-RCTs. These findings support the existence of indexing bias.

Conclusions: Our study indicates the existence of language bias and indexing bias among registered CS-RCTs on drug interventions. This may distort evidence synthesis towards more positive results of drug interventions.

Patient or healthcare consumer involvement: None.

Assessment of publication bias in breast cancer systematic reviews

Pompeu AC[1], Porfírio GJM[1]

[1] Universidade Municipal de São Caetano do Sul, Brazil

Background: Publication bias is related to aspects other than quality that can influence the viability of a publication. In general, there is a greater probability of publishing studies with positive results not showing the totality of the available data. Considering the large number of reviews published in breast cancer in relation to the modern treatments, we decided to assess publication bias in these studies.

Objective: To assess the methods used to avoid or consider risk of publication bias in systematic reviews published on breast cancer.

Methods: We conducted an electronic search identifying systematic reviews of breast cancer published in MEDLINE in the last five years. We extracted data using a structured form and analysed the studies with calculation of frequency and 95% confidence interval.

Results: We included 185 reviews. A mean of 4 (± 2.1) databases were used in the included reviews. Of these, 37% (CI 95% 30% to 44%) reported a plan to use funnel plot analysis with 18% (CI 95% 13% to 24%) of the total reviews using it. Only 30% (CI 95% 24% to 37%) used sources other than electronic databases.

Conclusion: There is inadequate planning and reporting of the assessment of risk of bias in systematic reviews in breast cancer.

Baseline heterogeneity: a method to identify trials with bias arising from randomization in meta-analyses

Wang R[1], van Wely M[2], Mol B[1], Li W[1]

[1] Monash University, Australia; [2] University of Amsterdam, The Netherlands

Background: Meta-analyses of a baseline variable should have no heterogeneity as differences between treatment groups from true randomization are due to chance. Heterogeneity of baseline variables has therefore been proposed to evaluate the robustness of randomization of randomized controlled trials (RCTs) in meta-analyses. Trials resulting in baseline heterogeneity should be excluded to reduce potential bias in meta-analyses. However, this method has not been applied widely and it is unclear whether this method is comparable to other objective tools to assess bias arising from randomization such as Monte Carlo simulations. Recently, we assessed the randomization in RCTs included in two Cochrane Reviews by using Monte Carlo simulations and found that the randomization of RCTs in one review was likely to be robust, while in the other it was not.

Objectives: To assess whether the two methods assessing randomization, i.e. investigating the heterogeneity of baseline variables and Monte Carlo simulations, coincide in conclusions; and to assess the changes in risk estimates by excluding trials resulting in heterogeneity of baseline variables in meta-analyses.

Methods: We applied the two methods assessing randomization to two Cochrane Reviews evaluating endometrial scratching: one for in-vitro fertilization (IVF) and the other for intrauterine insemination (IUI)/natural intercourse. We extracted baseline age and body mass index (BMI) across treatment arms from trials with full-text publications included in the two Cochrane Reviews. Next, we performed a fixed-effect meta-analysis of the baseline age and BMI, respectively, and compared the results from baseline heterogeneity with those from Monte Carlo simulations. When baseline heterogeneity is observed, we excluded trials with the largest t-statistic and repeated the meta-analysis until $I^2 = 0$.

Results: For RCTs included in the Cochrane Review on endometrial scratching for IVF, there was no heterogeneity for baseline age or BMI, indicating that these RCTs were likely to be properly randomized, which is in line with the results of Monte Carlo simulations (P = 0.8654). For trials included in the Cochrane Review on endometrial scratching for IUI/intercourse, the heterogeneity for baseline age and BMI were high (I^2 = 80% and 92%, respectively), indicating that some of these trials were unlikely to be properly randomized, which also agreed with the results of Monte Carlo simulations (P = 1.754*10-5). After excluding trials resulting in heterogeneity of baseline age or BMI, the effect size of meta-analysis changed from 2.02 (95% confidence interval (CI) 1.52 to 2.68) to 1.76 (95% CI 0.74 to 4.21).

Conclusions: Assessing baseline heterogeneity could be an alternative method to evaluate bias arising from the randomization process for trials included in a meta-analysis. Excluding trials contributing to baseline heterogeneity could result in a more accurate estimate of effect size in a meta-analysis.

Patient or healthcare consumer involvement: None

Comparing quality assessments of systematic reviews/meta-analyses in nutrition using AMSTAR-2 and ROBIS: betwixt and between

Swierz MJ[1], Storman D[1], Koperny M[2], Zajac J[2], Tobola P[1], Staskiewicz W[3], Gorecka M[3], Skuza A[3], Wach A[3], Kaluzinska K[3], Johnston B[4], Bala MM[1]

[1] Chair of Epidemiology and Preventive Medicine, Department of Hygiene and Dietetics, Jagiellonian University Medical College, Krakow, Poland; [2] Agency for Health Technology Assessment and Tariff System, Poland; [3] Students' Scientific Group of Systematic Reviews, Jagiellonian University Medical College, Krakow, Poland; [4] Department of Nutrition, Texas A&M University, College Station, Texas, USA; Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada

Background: The quality of studies published as systematic reviews/meta-analyses (SR/MAs) in many fields of health care is low. AMSTAR-2 (A Measurement Tool to Assess Systematic Reviews) and ROBIS (Risk of Bias in Systematic Reviews) are independent instruments that focus on different but overlapping aspects of SR/MA quality.

Objectives: To assess and compare the quality of studies published as SR/MAs on nutritional interventions in cancer prevention using AMSTAR-2 and ROBIS.

Methods: We performed a systematic survey of MEDLINE, Embase and the Cochrane Library for SR/MAs published between January 2010 and August 2018. We included studies identified as SR/MAs in the title and/or abstract which examined the effects of any nutritional intervention for cancer prevention. Following a calibration exercise, two independent review authors completed study selection, data extraction and AMSTAR-2/ROBIS assessments. We resolved conflicts by discussion or consultation with a third review author. AMSTAR-2 is comprised of 16 questions for which yes, partial yes, no, or not applicable can be applied. ROBIS consists of 21 signalling questions for which yes, partial yes, partial no, no, or no information can be applied. Our protocol was registered in PROSPERO: CRD42019121116.

Results: Out of 743 included articles, we selected a random sample of 101 for detailed analyses. Overall, the quality of SR/MAs in nutrition was low on AMSTAR-2 and ROBIS. We made 11 comparisons between AMSTAR-2 and ROBIS questions assessing the same construct. Some questions measuring separate constructs could not be compared (i.e. explanation for selection of study design in AMSTAR-2; appropriateness of eligibility criteria in ROBIS), and in some cases multiple questions were combined for comparison (i.e. comprehensiveness of searches and validity of statistical methods used). Comparability of reviewers' judgments between AMSTAR-2 and ROBIS questions is presented in Table 1. In 9 out of 11 comparisons the assessments were comparable, ranging from 78.2% to 99.0% agreement. For two comparisons, including comprehensive literature search and publication bias, the assessments were poorly comparable (59.4%).

Conclusions: AMSTAR-2 and ROBIS instruments mostly address similar aspects of SR/MA quality and our assessments using different instruments were similar. However, AMSTAR-2 uniquely addresses reporting of excluded studies, sources of funding, conflict of interest within individual studies and reasons for selection of study designs for inclusion, while ROBIS uniquely addresses adherence to predefined analyses, appropriateness and restrictions within eligibility criteria. Potential users should be aware of the considerably large overlap and the small but unique differences.

Patient or healthcare consumer involvement: We did not have a patient or healthcare consumer involved, however information about the quality of SR/MAs is important for consumers when interpreting results.

Table 1. Comparability of AMSTAR-2 and ROBIS judgements

AMSTAR-2 Items	ROBIS signaling questions	Comparability of judgments	Cohen's kappa
Item 1	1.3	91.1%	0.716
Item 2	1.1	99.0%	0.928
Item 4	2.1 & 2.2 & 2.3 & 2.4	59.4%	0.247
Item 5	2.5	92.1%	0.838
Item 6	3.1	94.1%	0.882
Item 8	3.2	78.2%	0.555
Item 9	3.4	93.1%	0.861
Item 11	4.3 & 4.4	89.6%	0.773
Item 12 & Item 13	4.6	85.1%	0.600
Item 14	4.4	83.2%	0.574
Item 15	4.5	59.4%	0.222
Total			**0.660**

Current Cochrane acupuncture reviews might be biased without searching Chinese, Korean and Japanese databases

Lu ZX[1], Lao YF[2], Yao L[3], Yan PJ[4], Hui X[4], Han X[2], Feng XW[2], Fu YW[2], Ma YX[2], Li J[4], Yang KH[5]

[1] First Clinical Medical College, Lanzhou University, Lanzhou, China; [2] Second Clinical Medical College, Lanzhou University, Lanzhou, China; [3] Health Research Methodology, Department of Health Research Methods, Evidence and Impact, McMaster University, Canada; [4] Evidence-Based Social Science Research Center, School of Public Health, Lanzhou University, Lanzhou, China; [5] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China

Background: Acupuncture is widely used as an alternative therapy and evaluated by different Cochrane Reviews. Some acupuncture reviews searched Chinese, Korean and Japanese databases to include more potential studies published in local language, whereas others did not. In this study we want to investigate whether Chinese, Korean and Japanese databases should be considered in Cochrane Reviews of acupuncture.

Objectives: To explore the value of Chinese, Korean and Japanese databases to Cochrane acupuncture reviews.

Methods: We searched for and included all Cochrane Reviews and protocols related to acupuncture. We extracted information on countries, databases searched, and included studies. Local databases were defined as Chinese, Korean and Japanese databases. Local studies were defined as studies that could only be found by searching local databases.

Results: We included 134 acupuncture reviews and 20 protocols. The authors are mainly from China, Australia, USA, UK, Korea and Canada. Eighty-six reviews did not search any local databases, and 48 reviews searched local databases (median = 4, range: 1 to 13). We found 85% of reviews that searched local databases included at least one Chinese, Korean or Japanese author. Recruiting local authors was associated with searching local databases in the review (odds ratio (OR) 35.79, 14.28 to 89.67). Cochrane Reviews of acupuncture searched local databases included 61% non-English studies, and 60% could only be obtained from local databases. In addition, those reviews that did not search any local databases included 7% non-English studies from PubMed, Embase, CENTRAL and other databases.

Conclusions: There is no consensus of searching local databases in current Cochrane acupuncture reviews. Our study shows it is valuable to do that, otherwise, the results might be affected by selection bias. The possible solution is to included Chinese, Korean and Japanese authors in the review.

Patient or healthcare consumer involvement: There was no patient or healthcare consumer involved in this project.

Additional files: Table 1; Figure 1; Figure 2

Duplicate patterns, bias, and inconsistencies among Chinese-sponsored randomized controlled trials

Jia Y[1], Huang D[1], Wen J[1], Qureshi R[1], Wang Y[1], Rosman L[2], Chen Q[3], Robinson K[4], Gagnier J[5], Ehrhardt S[1], Celentano D[1]

[1] Bloomberg School of Public Health, the Johns Hopkins University, USA; [2] Welch Medical Library, the Johns Hopkins University, USA; [3] Institute of Medical Information & Library, Chinese Academy of Medical Sciences & Peking Union Medical College, China; [4] School of Medicine, the Johns Hopkins University, USA; [5] Michigan Medicine, School of Public Health, University of Michigan, USA

Background: Previous studies implied the existence of duplicates among research publications from China.

Objectives: To explore the patterns of duplicates, evaluate the existence of duplicate publication bias, and assess the inconsistencies between duplicates and the original publications produced from Chinese-sponsored randomized controlled trials (CS-RCTs).

Methods: We searched trial registries for eligible CS-RCTs which evaluated drugs and were conducted between 1 January 2008 and 31 December 2014. We identified journal articles produced from CS-RCTs from both English and Chinese bibliographic databases. We identified the main article and duplicates of each CS-RCT. Duplicates were classified as four patterns: IMALAS (interim analysis without self-identification or cross-reference), re-publication (re-publication of the main article without self-identification or cross-reference), SALAMI (subgroup analysis without self-identification or cross-reference), and partial duplicates which shared a subset of participants with other articles and had no cross-reference. Duplicate publication bias was evaluated following a retrospective cohort study design. We hypothesized that a CS-RCT was more likely to have duplicate(s) if its initial journal article was positive. We also assessed the inconsistencies in reporting between the main articles and the duplicates in terms of treatment, efficacy outcomes, and adverse events.

Results: Among 470 CS-RCTs published as journal article(s), 55 (11.7%) had 75 duplicates. Fifteen (20.0%), 33 (44.0%), 25 (33.3%), and 2 (2.7%) out of 75 duplicates were IMALASes, re-publications, SALAMIs, and partial duplicates, respectively. After adjusting for covariates, among CS-RCTs that were initially published in Chinese, CS-RCTs were 2.35 (95% confidence interval (CI) 1.04 to 5.30) times more likely to have duplicate(s) if their first article was positive. Among CS-RCTs that were initially published in English, CS-RCTs were 0.99 (95% CI 0.31 to 3.15) times more likely to have duplicate(s) if their first article was positive. Among 51 eligible duplicates, 14 (27.5%) and 3 (5.9%) reported inconsistent doses and schedules compared with main articles, respectively. Among 25 eligible duplicates, 9 (36%) reported inconsistent outcomes comparing with main articles. Among 15 eligible duplicates, 11 (73.3%) reported inconsistent adverse events comparing with main articles, of which (26.7%) reported completely different adverse events comparing with main articles.

Conclusions: At least 11.7% of CS-RCTs registered in trial registries have at least one duplicate. The most prevailing duplicate pattern was re-publication. There was evidence supporting duplicate publication bias among CS-RCTs initially published in Chinese. The inconsistencies between the main articles and duplicates implied the inaccurate reporting of CS-RCTs.

Patient or healthcare consumer involvement: None.

Evaluation of the treatment integrity as part of the risk of bias assessment: a proposal is needed

Yakoub NS[1], Wolf M[2], Munder T[2], López-Alcalde J[1], Witt CM[1], Barth J[1]

[1] Institute for Complementary and Integrative Medicine, University Hospital Zurich and University of Zurich, Zurich, Switzerland; [2] Institute of Psychology, University of Zurich, Zurich, Switzerland

Background: Risk-of-bias (RoB) assessment is essential in systematic reviews (SRs). The assessment of the risk of performance bias requires to determine the degree to which the interventions were delivered as intended (1). The successful implementation of an intervention is also known as treatment integrity (TI) (2). The determination of TI is challenging, particularly in studies of face-to-face non-pharmacological interventions, such as psychotherapy, or other complex interventions.

Objectives: To develop a framework for the assessment of TI in each arm of a study evaluating the effects of a face-to-face non-pharmacological psychotherapeutic intervention. To identify core domains of TI that the Cochrane RoB 2 tool should consider.

Methods: Descriptive study of an ongoing project. We generated a list of potential TI domains. We collated these domains from the quality/risk of bias tools used in two sources: a sample of SRs of psychotherapeutic interventions published in Clinical Psychology Review and one SR of quality indicators of psychotherapy outcome studies (3). We will share this list of domains and related sub-domains with a small group of experts to check if all relevant domains were considered. This preliminary list will be used in an upcoming survey with approximately 300 invited people (covering psychotherapists, researchers, Cochrane Review authors and patients). The survey will inform about the adequacy of each domain (or sub-domain) and the feasibility to rate each sub-domain. We will conduct a consensus meeting with key stakeholders (preferably Cochrane Review authors, meta-analysts of psychotherapy outcome studies and trialists) to agree on the TI domains and sub-domains that the RoB 2 tool should include.

Results: We identified a total of nineteen different quality/risk of bias tools. Eleven tools (58%) considered TI in at least one domain. Our preliminary analysis suggests that the tools considered heterogenous aspects concerning TI, such as the competence of the provider, the compliance by the patients, or the description of co-interventions.

Conclusions: To assess TI is critical in the risk of bias assessment. However, SRs of face-to-face non-pharmacological psychotherapeutic interventions assessed TI inconsistently. There is an urgent need to define a pragmatic, explicit and reproducible approach to determine TI in SRs.

Patient or healthcare consumer involvement: Valid assessment of the TI is a critical step of the risk of bias assessment. Besides, to assess the TI can also be useful for providers and healthcare consumers to know the acceptability and feasibility of an intervention. Therefore, we plan to integrate psychotherapists and patients' views in our project.

References:

1. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6 (July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook.
2. Perepletchikova F, Treat TA, Kazdin AE. Treatment integrity in psychotherapy research: analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology* 2007;75(6):829-41. https://doi.org/10.1037/0022-006X.75.6.829.
3. Liebherz S, Schmidt N, Rabung S. How to assess the quality of psychotherapy outcome studies: A systematic review of quality assessment criteria. *Psychotherapy Research* 2016;26(5):573-89.

Exploring the effect of risk of bias on the evaluation of knowledge translation in the context of health policy: systematic review and meta-epidemiology analysis

Ahmadi A[1], Doshmangir L[2], Yazdizadeh B[3], Majdzadeh R[3], Asghari S[4]

[1] Knowledge Utilization Research Center, Tehran University of Medical Sciences, Iran; [2] Health Services Management Research Centre, Tabriz University of Medical Sciences, Iran; [3] Knowledge Utilization Research Centre, Tehran University of Medical Science, Iran; [4] Memorial University of Newfoundland, Health Science Centre, St. John's, Canada

Background: Activities that try to fill the "knowledge to action gap" can be considered as a Knowledge Translation (KT) intervention. Most of the KT interventions do not have enough compatibility, usually because of different levels of methodological quality over individual studies. Most of the time, it is responsible for observed heterogeneity in estimated effects among individual studies. Therefore, applying methods to reduce risk of bias (ROB) in individual KT studies can increase comparability and promotes validity of the conclusions of the related review studies.

Objectives: (1) To describe how different methods (sequence generation, allocation concealment, blinding, incomplete outcome data, selective outcome reporting) were applied to reduce ROB in KT interventional studies; (2) To estimate the contribution of ROB to observed heterogeneity among individual study' estimated effects of KT interventions in health policy decision making.

Methods: Systematic review and meta-epidemiology design. Participants: policy makers. Intervention: structural, managerial or individual strategies. Outcomes: measures of research knowledge uptaking. Type of studies: randomized and non- randomized controlled trials. Search: MEDLINE, the Cochrane Central Register of Controlled Trials, DARE, SCOPUS, Web of Science (see Table 1). Strategy for synthesis: narrative synthesis of the findings from the included studies which was structured around the type of the methodological technique and type of study intervention, outcome (according to EPOC) and audience was provided. We obtained scores for each ROB (selection bias, performance bias, attrition bias, detection bias) and the overall ROB for every single included study. We estimated the relationship between the type of intervention/method/audience with ROB via meta-epidemiology analysis. We estimated the contribution of each method to the observed heterogeneity among individual studies through meta-regression.

Results: Out of 1633 retrieved independent studies, 31 were eligible and 17 were included in quantitative analysis. The risk of bias was higher in group level studies, consensus process intervention and non-randomized trial. Studies which applied sequence generation techniques in the intervention allocation process showed lower total score (Mean (standard deviation): 38 (18) vs 76 (31)). Meta-epidemiology analysis showed that SMD in studies with higher ROB was 0.17 (CI: 0.05 to 0.29) greater than studies with lower ROB over all types of KT interventions (Figure 1). Studies that showed stronger effect have higher risk of bias in meta-regression analysis (B = .013, p.007) (figure 2).

Conclusion: Risk of bias can distort the observed studies' results in a way that they would show more exaggerated resulted values. This distortion is seemingly higher in more complex interventions than simple interventions and when there are higher levels of subjectivity in study measures.

Patient or healthcare consumer involvement: Policy makers have shared their opinion on the study's objectives and the results interpretation so they have been modified accordingly.

Additional files: Search strategy; Figure 1; Figure 2

Identifying publication bias in meta-analyses of continuous outcomes in the presence of baseline risk

Doleman B[1], Freeman S[2], Sutton A[2]

[1] University of Nottingham, UK; [2] University of Leicester, UK

Background: In meta-analyses, funnel plot asymmetry can be considered evidence of small study effects (possible publication bias (PB)). Egger's test is a linear regression of the treatment effect estimates on their standard errors weighted by their inverse variance and is conducted as a formal statistical test for funnel plot asymmetry. The performance of Egger's and related tests has been widely studied for binary outcomes but not for continuous outcomes. Baseline risk (BR) is an interaction of treatment effect with the severity of a condition measured by the observed treatment effect in the control group. In postoperative pain meta-analyses it has been shown that, on average, studies with higher BR (i.e. higher morphine consumption in the control group) will have larger standard deviations. If the treatment effect estimates are also dependent on BR then this may cause correlation between the outcome measure and standard errors which could result in funnel plot asymmetry even in the absence of PB. To overcome this, we propose a new test for funnel plot asymmetry based on meta-regression residuals. The new test is a two-stage process in which BR is included as a study-level covariate in a meta-regression (MR) model before a regression-based asymmetry test using the MR residuals as the outcome and inverse sample size as the exploratory variable is performed.

Objectives: To evaluate and compare the performance of Egger's test and the test of MR residuals for identifying funnel plot asymmetry.

Methods: 1) Application of Egger's test and the test of MR residuals to 9 meta-analyses of postoperative analgesics measuring 24-hour morphine consumption. 2) Simulation study to formally evaluate the test of MR residuals considering each combination of BR and PB being present or not.

Results: 1) Egger's test and the test of MR residuals identified funnel plot asymmetry in 6 and 2 (of 9) meta-analyses respectively. 2) Based on 10,000 simulated meta-analyses the test of MR residuals had similar power to Egger's test when no BR and PB were simulated (63% versus 63%) and reduced type I errors when BR and no PB were simulated (60% versus 6%). It also had modest power to detect funnel plot asymmetry in the presence of treatment effects interacting with BR (40%).

Conclusions: Continuous outcomes are commonly measured on an absolute (mean) difference scale and it is not uncommon for the magnitude of the intervention effect to be related to response in the control arm (i.e. baseline risk). When this is the case funnel plots can appear highly asymmetric, even when PB is not present, since correlations exist between outcome and both effect size and standard error. We have shown that Egger's test is potentially misleading for continuous outcomes and a test which regresses the residuals from a MR model, including BR as a study-level covariate, has better statistical properties.

Patient or healthcare consumer involvement: There was no patient involvement in this research however implications from this research will be discussed with patient representatives.

Ignoring non-significant factors without data may bias the results of meta-analysis of prognostic studies

Wang L[1], Guyatt G[2], Thabane L[2], Busse J[1]

[1] Department of Anesthesia, McMaster University, Canada; [2] Department of Health Research Methods, Evidence and Impact, McMaster University, Canada

Background: Meta-analysis of prognostic studies usually pools reported measures of association for common factors across studies. Failure to consider non-significant factors that were excluded or not reported in regression models may overestimate pooled measures of association.

Objectives: Using systematic reviews of predictors of persistent postsurgical pain or unemployment after breast cancer surgery, we explored the impact of imputing data for missing non-significant data on overall associations of risk factors.

Methods: We pooled predictors to explore their association with persistent pain or unemployment after breast cancer surgery using random-effects models and the DerSimonian-Laird method. For our primary analyses, we imputed an odds ratio (OR) of "1" for predictors that were excluded or not reported in multivariable analyses due to non-significant association. We acquired the associated variance using the hot deck approach. We performed sensitivity analysis by excluding the imputed data for non-significant predictors (analysis of adjusted data reported). We calculated the ratio of odds ratio to estimate the difference between these two approaches.

Results: We included 56 studies with 66,740 patients. Most of the studies either excluded factors that were not significant in bivariate analysis (32%; 18 of 56) or failed to present data for non-significant predictors in their final regression models (68%; 38 of 56). Twenty-four of 27 risk factors contained missing data for non-significant factors (Table 1). The median ratio of odds ratio of pooling analyses using imputed data vs. only reported data in the final multivariable models was 1.07 (range: 1.01 to 1.15) for 9 poolable risk factors for persistent pain and 1.05 (range: 1.01 to 1.80) for 15 poolable predictors of unemployment (Figures 1 to 4). All pooled associations were larger in meta-analyses based on reported data only vs. when imputed missing data were considered, which exaggerated the magnitude of association by 1% to 55%.

Conclusions: Primary studies exploring prognostic factors often fail to report data for non-significant predictors. Failure to impute for missing non-significant predictors in meta-analyses systematically overestimates pooled measures of association. Systematic review authors should acquire missing data for non-significant predictors from study authors when possible, and impute data when not.

Patient or healthcare consumer involvement: No

Additional files: Tables and figures

Impact of excluding non-English studies on network meta-analysis: an empirical study

Wang Q[1], Zhang Q[2], Pan B[3], Tian J[4], Ge L[5]

[1] Department of Social Medicine and Health Management, School of Public Health, Lanzhou University, China; [2] School of Nursing, Lanzhou University, China; [3] Gansu Provincial Hospital, Lanzhou, China; [4] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, China; [5] Evidence-Based Social Science Research Center, School of Public Health, Lanzhou University, China

Background: Clinical decision-making requires synthesis of evidence from literature reviews on a specific theme. Meta-analysis and network meta-analysis are two important tools that provide an efficient and comprehensive way of solving conflicts on one health problem statistically. Researchers from different countries are using English, which is generally considered as the universal language of science when publishing studies. A systematic review of empirical studies found no differences between summary treatment effects in English-language restricted meta-analyses and non-English inclusive meta-analyses. To date, the effect of excluding non-English studies of network meta-analysis has not been studied.

Objectives: To assess the impact of excluding non-English studies on the network geometry and conclusions in Cochrane network meta-analysis.

Methods: We searched the Cochrane Library for network meta-analysis. We extracted the original data from the attached forest plots and also the publication language of included studies. We then re-conducted network meta-analysis with same approach to the original Cochrane network meta-analysis after excluding non-English studies. We compared the network geometry and results between network meta-analysis including all studies and those only including English studies. We will report the change of network plots and results, and also calculate the relative ratio of network estimates between including all studies and including only English studies. We will use Stata 15.1 software for all statistical analysis.

Results: This study is ongoing, and results will be presented at the Cochrane Colloquium as available.

Conclusions: This study is ongoing, and results will be presented at the Cochrane Colloquium as available.

Patient or healthcare consumer involvement: No patient or healthcare consumers are involved in this study.

Impact of training and guidance on the inter-rater and inter-consensus reliability of risk of bias instruments for non-randomized studies

Jeyaraman MM[1], Robson RC[2], Pollock M[3], Copstein L[1], Balijepalli C[4], Hofer K[5], Xia J[6], Al-Yousif N[1], Mansour S[7], Fazeli MS[5], Ansari MT[8], Tricco AC[9], Rabbani R[1], Abou-Setta AM[1]

[1] University of Manitoba, Canada; [2] Li Ka Shing Knowledge Institute, St. Michael's Hospital, Canada; [3] Institute of Health Economics, Canada; [4] Pharmalytics Group, Canada; [5] Evidinno Outcomes Research Inc., Canada; [6] Nottingham Ningbo GRADE Centre, China; [7] University of Montréal, Canada; [8] University of Ottawa, Canada; [9] University of Toronto, Canada

Background: In 2016, the Cochrane Bias Methods group developed a risk of bias (ROB) tool (Risk of Bias in Non-Randomized Studies of Interventions [ROBINS-I]). In 2019, this tool was adapted for non-randomized studies of exposures (Risk of Bias Instrument for Non-Randomized Studies of Exposures [ROB-NRSE]).

Objectives: To evaluate the impact of training and customized guidance on inter-rater reliability (IRR), inter-consensus reliability (ICR; comparison of consensus assessments across reviewer pairs), and evaluator burden of ROBINS-I and ROB-NRSE.

Methods: An international team of seven review authors from six review centers appraised the ROB using either ROBINS-I (n = 44) or ROB-NRSE (n = 44) in two stages. Stage one was ROB assessments before training or customized guidance, and stage two was ROB assessments after training and customized guidance. Two pairs of reviewers independently assessed the same sample of study publications in both stages. After completion, each pair resolved conflicts through consensus. Reviewers also recorded the time taken for completion of each step. For analysis of the IRR and ICR, we used Gwet's AC1 statistic. Agreements among the reviewers were categorized as: poor (0 to 0.09), slight (0.10 to 0.20), fair (0.21 to 0.40), moderate (0.41 to 0.60), substantial (0.61 to 0.80), near perfect (0.81 to 0.99), or perfect (1.00).

Results: For ROBINS-I, the IRR (Table 1) improved after training and customized guidance for all domains except "bias in classification of interventions", which showed a decrease in IRR (from moderate to slight agreement). For ROB-NRSE, the IRR for all domains showed improvement after training and customized guidance (Table 2), except for the "bias due to missing data" domain, for which there were no improvements, and the "bias in classification of exposures" domain, for which there was a slight decrease in IRR (from moderate to fair agreement). For ROBINS-I, the ICR improved for all domains (Table 3). For ROB-NRSE, all domains improved except for "bias due to confounding", for which there was no improvement after guidance and training (Table 4). For both tools, the overall bias assessments for both IRR and ICR showed improvements after training and guidance. The evaluator burden (time taken to read article + adjudication + consensus) decreased after guidance and training for ROBINS-I (before training and guidance: 48.45 min (95% confidence interval (CI) 45.61 to 51.29) vs. after training and guidance: 35.6 min (95% CI 32.77 to 38.33), whereas there was a slight increase for ROB-NRSE (before training and guidance: 36.98 min (95% CI 34.80 to 39.16) vs. after training and guidance: 40.5 min (95% CI 37.30 to 43.66)).

Conclusions: In our cross-sectional study, the IRR and ICR of ROBINS-I and ROB-NRSE improved overall after training and customized guidance. While conducting systematic reviews of non-randomized studies, we highly recommend additional training and customized guidance to reviewers prior to ROB assessments.

Patient or healthcare consumer involvement: Patients or healthcare consumers were not involved in this project.

Additional files: Tables

Key methods reported in protocols and publications of core outcome sets were inconsistent: a comparative study

Liu M[1], Gao Y[1], Yuan Y[2], Shi S[1], Yang K[1], Wu J[3], Zhang J[4], Tian J[1]

[1] Lanzhou University, China; [2] Gansu University of Chinese Medicine, China; [3] Beijing University of Chinese Medicine, China; [4] Tianjin University of Traditional Chinese Medicine, China

Background: Outcomes reporting bias is a prevalent and easily overlooked problem in clinical research. Since 1992, the Outcome Measures for Rheumatology Clinical Trial (OMERACT) initiative has worked to identify and promote core outcome set in the field of rheumatic disease. At the same time, the number of completed studies with core outcome sets is rapidly increasing. As of November 2019, the COMET database contained 1327 references of planned, ongoing and completed work.

Objectives: To assess inconsistencies between protocols and full-text publications in developing core outcome sets, and to help improve the methodology.

Methods: We retrieved protocols and subsequent full-text publications by searching electronic bibliographic databases. We summarized the changes of general and methodology characteristics by comparing the protocols with the full-text publications and represented the information using frequency and proportion.

Results: Our search identified 24 protocols and 32 corresponding full-text publications . All included studies involved 14 different study topics. The earliest protocol was published in 2013. The gap of published years between the full-text publications and the protocols was one to four years. As for general characteristics, the first authors of nine studies (37.5%) and the correspondent-authors of eight studies (33.3%) have changed. As for methodological characteristics, in 11 (45.8%) studies, the literatures review was the only source of the initial list of outcomes. In six (25.0%) studies the initial list of outcomes was identified from literatures reviews and interviews. In four (16.7%) studies, the initial list of outcomes was identified from literature reviews and interviews and consultation. Four studies (16.7%) added study types of the literature reviews in full-text publications. Six studies (25.0%) added surveyed participants in full-text publications, and participants of 11 studies (45.8%) were found to be inconsistent in a consensus meeting.

Conclusions: Key methods reported in protocols and publications of core outcome sets were inconsistent. Methodologists should work with medical editors and scientific journals to encourage the registration of protocols before developing core outcome sets.

Patient or healthcare consumer involvement: None.

Table 1. General characteristics comparison between protocols and publications

Characteristics	Protocol	Publication	P value
Number of authors			
1 to 5 authors	8 (0.33)	2 (0.08)	0.033
6 to 10 authors	7 (0.29)	7 (0.29)	1.000
11 to 15 authors	5 (0.21)	8 (0.33)	0.330
16 or more authors	4 (0.17)	7 (0.29)	0.303
Journal impact factor			
0.0-3.0	21 (0.88)	10 (0.31)	0.000
3.1-6.0	1 (0.04)	14 (0.44)	0.001
6.1-9.0	0 (0.00)	6 (0.19)	0.032
>9.0	0 (0.00)	1 (0.03)	1.000
Non-SCI*	2 (0.08)	1 (0.03)	0.575
Funding sources			
None	7 (0.29)	7 (0.29)	1.000
Yes	17 (0.71)	17 (0.71)	1.000
Competing interests			
None	24 (1.00)	19 (0.79)	0.050
Yes	0 (0.00)	5 (0.21)	0.050
Source of initial outcomes			
Literatures	21(0.88)	21(0.88)	1.00
Literatures&Interviews	10(0.42)	10(0.42)	1.00
Literatures&Interviews&Consultations	4(16.7)	4(16.7)	1.00

*SCI= Science Citation Index

Lessebo effects in RCTs of SSRIs for panic disorder in adults

Ernst M[1], Hagmayer Y[1]

[1] Georg-Elias-Mueller Institute for Psychology, University of Goettingen, Germany

Background: Lessebo effects describe decreased effects of active treatment and placebo in randomized controlled trials (RCTs) with objectively lower chances of receiving active treatment. These effects may bias estimations of effect sizes in RCTs and systematic reviews.

Objectives: To investigate the presence of lessebo effects in RCTs of the treatment of panic disorder with selective serotonin reuptake inhibitors (SSRIs) in adults.

Methods: We searched PubMed to find relevant studies up to 27 February 2018. We also handsearched reference lists of relevant studies, systematic reviews and meta-analyses. SSRI and placebo arms from randomized, double-blind, placebo- or active-controlled trials of the treatment of panic disorder with SSRIs in adults were eligible for inclusion. We extracted the proportion of patients free from panic attacks and changes in panic disorder symptom severity. We pooled outcomes by trial design: outcomes in SSRI arms were pooled for active-controlled, and for placebo-controlled trials; outcomes in placebo arms were pooled for trials with multiple active treatment arms and for trials with only one active treatment arm.

Results: We identified 773 records and screened 68 abstracts. We assessed 49 full-text articles for eligibility. We included in this study 53 trial arms consisting of 36 SSRI arms and 17 placebo arms from 27 studies. Random-effects meta-regressions demonstrated that the proportion of patients free from panic attacks in SSRI arms was lower in placebo-controlled trials compared to active-controlled trials (48.3% vs 63.0%, P = 0.007) in the intention-to-treat (ITT) sample. However, this lessebo effect was accounted for by systematically lower completion rates in placebo-controlled trials compared to active-controlled trials. No further lessebo effects were identified.

Conclusions: Our study indicates that lessebo effects are quite small, if they exist in RCTs of the treatment of panic disorder with SSRIs in adults. Our data also suggest that lessebo effects found in ITT-analyses may be, at least partially, due to systematic differences in completion rates across trial designs. Analyses of higher power are needed to clarify the size and the robustness of lessebo effects. We emphasize that trial design should be considered whenever treatment effects in RCTs are interpreted.

Patient or healthcare consumer involvement: Even though no patients or healthcare consumers were directly involved in this investigation, this study's analyses were based on previously collected data from patients participating in RCTs. Since lessebo effects may bias estimations of effect sizes derived from RCTs and systematic reviews, patients and healthcare consumers referring to such studies may over- or underestimate the beneficial effect of a treatment, just like investigators and clinicians. Furthermore, lessebo effects may contribute to understanding subjective experiences made by patients while participating in a RCT.

Limitations of combining ROBINS-I and Cochrane RoB 2 in a systematic review with mixed study designs

De Brier N[1], Borra V[2], O D[1], Vandekerckhove P[3], De Buck E[4]

[1] Centre for Evidence-Based Practice, Belgian Red Cross, Belgium; [2] Centre for Evidence-Based Practice, Belgian Red Cross; Cochrane First Aid, Belgium; [3] Belgian Red Cross; Department of Public Health and Primary Care, Faculty of Medicine, KU Leuven, Belgium; [4] Centre for Evidence-Based Practice, Belgian Red Cross; Department of Public Health and Primary Care, Faculty of Medicine, KU Leuven; Cochrane First Aid, Belgium

Background: ROBINS-I is the preferred tool to assess risk of bias in Cochrane Reviews for non-randomized studies (NRS). Whereas the GRADE Working Group typically assigns a low certainty to the body of evidence in these studies, the underlying study design is not considered as a risk of bias feature in ROBINS-I. Thus, the initial certainty of the body of evidence from NRS would be high allowing for a better comparison of evidence from RCTs and NRS. However, there is to date little if any methodological guidance on how to combine and assess the body of evidence originating from both RCTs and NRS in GRADE assessments in one systematic review.

Objectives: To combine results from RCTs and NRS for assessing the body of evidence in the framework of a systematic review evaluating the effectiveness of storage techniques for an avulsed tooth.

Methods: We searched for experimental and observational studies in the Cochrane Library, MEDLINE and Embase. Two review authors independently assessed the risk of bias of the included studies using the Cochrane RoB 2 tool for RCTs and ROBINS-I for NRS. We appraised the certainty of the body of evidence according to the GRADE methodology.

Results: Out of 4118 references, we included 19 RCTs, 7 non-RCTs and 7 prospective cohort studies. The RCTs suffered from serious concerns regarding the randomization process, measurement of the outcome and selection of the reported results. The certainty of evidence was graded as very low after accounting for risk of bias (-1), indirectness (-1) and imprecision (-1). In the cohort studies, there were also serious concerns regarding measurement of the outcome and selection of the reported results. Most of these studies controlled for confounding factors and there were only some concerns regarding selection or information bias. The certainty of evidence was downgraded for risk of bias (-1) and imprecision (-1). The cohort studies alleviated the concerns about indirectness. The cohort studies were hence rated as low certainty evidence and, counterintuitively, provided the best available evidence. To tackle this discrepancy, we decided to assess the limitations in study design in cohort studies as very serious (-2). Thereby, we recognized that randomization is the only way to fully protect against confounding. No incompatibilities were noticed between the GRADE assessment of the RCTs and non-RCTs since the identified limitations were very similar.

Conclusions: ROBINS-I poses a number of challenges to summarize risk of bias in GRADE assessments when the results from RCTs and NRS are evaluated together. New guidance should take into account that the evaluation of RoB 2 and ROBINS-I tools cannot simply be brought together in GRADE evidence profiles and should be evaluated against the background of the RCT design.

Patient or healthcare consumer involvement: The interpretation of systematic review findings by experts and patients in formulating recommendations is only valid when they can rely on proper assessments of the body of evidence.

Meta-analysis as a simultaneous inference problem: a novel approach to assess replicability of evidence

Panagiotou O[1], Voorhies K[1], Jaljuli I[2], Schmid C[1], Heller R[2]

[1] Brown University, USA; [2] Tel-Aviv University, Israel

Background: Replicability of treatment effects protects patients, clinicians, and policy makers from claiming conclusive evidence solely based on a single study which may be a false-positive due to chance or bias.

Objectives: To assess the extent of replicability in Cochrane meta-analyses and characterise non-replicable bodies of evidence.

Methods: We included all meta-analyses of binary outcomes with n > 4 studies. We applied the partial conjunction hypothesis test to quantify the evidence for replicability. The method establishes that the treatment effect is replicated in at least u out of n studies by testing the u/n-replicability null hypothesis, i.e. at least n-(u-1) of the component hypotheses in a meta-analysis simultaneously hold true. It calculates a summary measure (r-value) which is the P value of the aforementioned null replicability hypothesis. Replicability is established if the r-value is less than the type I error α = 0.05. Using the same meta-analytical methods as the Cochrane Reviews, we computed the r-value for u = 2 and u = 3 to determine whether the treatment effect is replicated in at least 2 and at least 3 studies. For each meta-analysis, we computed the u-max, i.e. the maximum u for which the u/n-replicability null hypothesis is rejected; u-max is the 1-α lower confidence bound on the number of studies with effect in the same direction.

Results: A total of 23,561 meta-analyses with 258,948 individual trials were eligible. The median number of studies per meta-analysis was 8 (interquartile range (IQR) 6 to 12) and the median sample size was 2984 (IQR 1,231 to 7,722). Replicability for u = 2 was not met (r > 0.05) in 15,482 (66%) meta-analyses and for u = 3 in 17,738 (75%) meta-analyses. There were 9863 statistically significant meta-analyses. Among those, replicability for u = 2 was not met in 2970 (30%), i.e. 1 study driving the meta-analysis significance; for u = 3, replicability was not met in 4493 (46%) with 2 studies driving the significance. The median u-max was 3 (IQR 1 to 5) and the median ratio of u-max to the total number of studies was 33% (IQR 14% to 60%). In total, 5078 (22%) meta-analyses had evidence of small study effects and the treatment effect was replicated in in at least two studies in 2684 (53%) of those meta-analyses. Among statistically significant meta-analyses whose treatment effect was replicated in at least two studies (n = 6,893), the treatment effect between the replicated studies and the overall meta-analysis was greater than 10% for 3518 (51%) meta-analyses; differences in treatment effects between the replicated studies and the overall meta-analysis were statistically significant in 34 cases. Results were similar when using α = 0.005 and α = 0.001.

Conclusions: Treatment effects are replicated in at least 2 trials in two-thirds of statistically significant meta-analyses with small variations in effect estimates. For many meta-analyses, statistical significance is sensitive to a small number of studies relatively to the number of synthesized studies.

Patient or healthcare consumer involvement: None

Monte Carlo simulations: an objective tool to identify problematic randomization in Cochrane Reviews

Li W[1], Suke S[1], Wertaschnigg D[2], Lensen S[3], Wang R[1], Gurrin L[3], Mol B[1]

[1] Monash University, Australia; [2] Paracelsus Medical University, Austria; [3] The University of Melbourne, Australia

Background: Monte Carlo simulations, which are computational algorithms that use random sampling to generate numerical results, have been used by Carlisle et al. to prove the extremely low probability of randomization for fabricated randomized controlled trials (RCTs). This objective method could also be used in Cochrane Reviews where the quality of random sequence generation, at present, is only assessed by the subjective description of randomization from the RCT authors.

Objectives: To demonstrate that Monte Carlo simulations could be used to evaluate the quality of random sampling in Cochrane Reviews by applying this method to two published Cochrane Reviews about the effectiveness of endometrial scratching for in-vitro fertilization (IVF) and intrauterine insemination (IUI)/natural intercourse.

Methods: We extracted all the baseline characteristics across intervention groups from RCTs with full-text included in the two Cochrane Reviews. We used Monte Carlo simulations to generate a P value for differences between means for each baseline continuously valued variable or proportions for each baseline categorical variable. If randomization has been done correctly then the set of P values from all baseline variables in studies should follow a uniform [0,1] distribution, that is, they should be randomly drawn values between 0 and 1. Stouffer's method was used to combine the P values for all baseline variables in a study to generate a single combined P value for that study. We then used the Kolmogorov–Smirnov test, against a uniform distribution [0,1], for the P values of baseline variables and RCTs, to check for the effectiveness of randomization across studies.

Results: For RCTs included in the Cochrane Review for IVF, there was no evidence against the assertion that P values from all baseline variables followed the expected uniform distribution, P = 0.8654; whereas there was strong evidence against the null hypothesis that the P values followed the uniform distribution in RCTs concerning IUI/intercourse, P = $1.754*10^{-5}$ (Figure 1). Similarly, the distribution of pooled P values for RCTs with respect to IVF was likely to follow the expected uniform distribution, P = 0.5825, in contrast, RCTs regarding IUI/intercourse did not follow the expected uniform distribution, P = $7.707*10^{-5}$.

Conclusions: Monte Carlo simulations could be used to evaluate the probability of randomization across RCTs in Cochrane Reviews. In the case of a low probability, additional quality assessments such as acquiring and analyzing individual participant data should be considered before pooling RCTs.

Patient or healthcare consumer involvement: None.

Additional file: [Figure 1](#)

Reliability and applicability of RoB 2: experience within the pilot review "Cannabis and cannabinoids for people with multiple sclerosis"

Minozzi S[1], Dwan K[2], Borrelli F[3], Moore T[4], Filippini G[5]

[1] Cochrane Review Group on Drugs and Alcohol; Department of Epidemiology, Lazio Regional Health Service, Italy; [2] Cochrane Editorial and Methods Department, London, UK; [3] Department of Pharmacy, School of Medicine and Surgery, University of Naples Federico II, Naples, Italy; [4] Cochrane; Population Health Sciences Department, Bristol Medical School, University of Bristol, UK; [5] Scientific Director's Office Fondazione, Istituto di Ricovero e Cura a Carattere Scientifico, Istituto Neurologico Carlo Besta, Milan, Italy

Background: The revised Cochrane risk-of-bias tool for randomized trials (RoB 2) is being piloted in 13 Cochrane Reviews (SRs) and it is recommended for use in new SRs since 2020. In 2019 we assessed the reliability and applicability of RoB 2 on the primary outcome of a random sample of 70 individually randomized parallel-group trials (RCTs) covering very different topics, so limiting the generalizability of our results within the context of a SR.

Objectives: To measure the inter-rater reliability (IRR) of RoB 2 and assess the difficulties and the time required to implement it within a SR.

Methods: Four raters with medium-high expertise in risk of bias assessment of RCTs independently applied RoB 2 to 18 individually randomized parallel-group trials included in the pilot review "Cannabis and cannabinoids for people with multiple sclerosis". We first performed a calibration exercise on five RCTs. Then we prepared a structured document on how to implement the tool within the SR (how to answer to signalling questions (SQ) considering the types of outcomes and the clinical context). Finally, we applied the tool to the remaining studies included in the SR. We calculated Fleiss' k for multiple raters for individual domains and overall risk of bias. We classified agreement as poor (≤ 0.00), slight (0.01 to 0.20), fair (0.21 to 0.40), moderate (0.41 to 0.60), substantial (0.61 to 0.80), almost perfect (0.81 to 1.00). We calculated the IRR separately for the first five studies assessed during calibration and for the remaining studies assessed after calibration. We measured the time to complete RoB 2 as the mean time spent in minutes by each rater on each study. We also measured the mean time in hours spent for the discussion during calibration and the definition of the criteria to answer SQs in our SR).

Results: Preliminary results on the first five RCTs are reported. The IRR was poor for overall risk of bias (-0.15), domain 2 (-0.15) and 4 (-0.24), fair for domain 1 (0.30), slight for domain 3 (0.08) and 5 (0.12). The mean time to complete RoB 2 was 168.5 minutes (standard deviation 68.7). The mean time to complete the whole calibration exercise (including the preparation of the document) was about 55 hours over a three-month period.

Conclusions: The analysis on the first five RCTs showed poor agreement for the overall RoB and highlighted difficulties in the comprehension and applicability of some SQs, particularly in domains 2 (deviations from the intended interventions), 3 (missing outcome data) and 5 (selection of the reported result). The application of RoB 2 and the completion of calibration exercise required a significant amount of time. The tool appears to be complex and requires a sound background in clinical epidemiology and statistics, as well as a proof knowledge of the subject matter. We will present results of the assessment on the total sample. We will discuss the implications of the use of RoB 2 for the work of the Cochrane Review Group editorial bases.

Patient or healthcare consumer involvement: The project focuses on methods to assess RoB, so we couldn't involve consumers.

Reporting bias in imaging diagnostic accuracy research: update on recent evidence

Frank RA[1], Cherpak L[1], Treanor L[2], Dehmoobad Sharifabadi A[2], Korevaar D[3], McGrath TA[1], Salameh JP[4], Bossuyt PMM[5], McInnes MDF[6]

[1] Department of Radiology, University of Ottawa, Canada; [2] Faculty of Medicine, University of Ottawa, Canada; [3] Clinical Epidemiology and Biostatistics and Bioinformatics, Public Health Research Institute, University of Amsterdam, The Netherlands; [4] Department of Medical Imaging, The Ottawa Hospital | Ottawa Hospital Research Institute, Canada; [5] Clinical Epidemiology and Biostatistics and Bioinformatics, Public Health Research Institute, University of Amsterdam, Netherlands; [6] Department of Radiology, University of Ottawa | Ottawa Hospital Research Institute, Canada

Background: Reporting bias occurs when research dissemination – such as publication (publication bias), time-to-publication (time-lag bias) or citation rate (citation bias) – is influenced by the magnitude and direction of findings. Understanding the impact of reporting bias in imaging diagnostic test accuracy (DTA) research could inform policy on evaluating publication bias in DTA systematic reviews.

Objectives: To summarize findings from recent evaluations of reporting bias in imaging DTA research.

Key findings: Summarized in Table 1. Publication Bias: We evaluated 405 abstracts of DTA primary research presented at the Radiological Society of North America meeting. Multivariable logistic regression revealed an odds ratio of 3.6 (95% confidence interval (CI) 1.9 to 6.7, P < 0.001) for full text publication in favour of conference abstracts with positive (vs. negative/neutral) conclusions. Diagnostic accuracy showed no significant association with publication. Time-Lag Bias: We evaluated 781 primary DTA studies. Multivariable cox regression showed that positive conclusions were associated with shorter time from study completion to publication, with a hazard ratio (HR) of 1.32 (95% CI 1.03 to 1.69, P = 0.030); median time-to-publication was 7 months less for studies with positive vs. negative conclusions. Higher Youden index (YI) was also associated with shorter time to publication (rho = −0.11, P = 0.009), with HR of 1.07 (95% CI 1.01 to 1.13; P = 0.021) per unit increase in logit-transformed YI, independent of conclusion positivity; median time-to-publication was 1 month less for studies with YI above (vs. below) the median. Citation Bias: We evaluated 1016 primary DTA studies. Negative binomial regression showed that positive conclusions were associated with higher citation rate (regression coefficient [r] = 0.19 (95% CI 0.03 to 0.35; P = 0.03)); mean citations per month was 0.54 for studies with positive conclusions vs. 0.34 for negative conclusions. A positive association between YI and citation rate (r = 0.35, P = 0.011) was muted when adjusting for conclusion positivity (r = 0.22, P = 0.12).

Conclusions: Reporting bias may be present in imaging DTA research. Studies with positive conclusions are published more often, published faster, and cited more often than those with negative or neutral conclusions. Studies with higher accuracy estimates are not more likely to be published; however, they are published more quickly and cited more often. In general, the magnitude of effect was stronger for conclusions than for accuracy.

Implications: Preferential dissemination of positive findings may lead to overestimation of imaging test accuracy, contributing to misinterpretation and adverse patient outcomes. Clinicians and reviewers should consider reporting bias when interpreting and synthesizing literature. However, since no association between diagnostic accuracy estimates and publication was identified, formal assessment of publication bias in imaging DTA systematic reviews may not be warranted.

Additional file: [Table](#)

Reporting bias in imaging diagnostic accuracy studies: are studies with positive conclusions or titles submitted and published faster?

Treanor L[1], Frank R[1], Atyani A[1], Dehmoobad Sharifabadi A[1], Hallgrimson Z[1], Fabiano N[1], Salameh J[1], McGrath T[1], Korevaar D[2], Bossuyt P[3], McInnes M[1]

[1] Department of Radiology, University of Ottawa, Canada; [2] Department of Respiratory Medicine, University of Amsterdam, The Netherlands;

[3] Clinical Epidemiology and Biostatistics and Bioinformatics, University of Amsterdam, The Netherlands

Background: Time-lag bias is a form of reporting bias resulting from the delayed publication of studies with 'negative' findings. If imaging diagnostic test accuracy (DTA) studies with negative conclusions have a delayed publication course, this may impede the ability to evaluate all evidence on a topic (e.g. in a systematic review). This can lead to the misallocation of healthcare resources towards diagnostic tests and detrimental downstream effects on patient outcomes.

Objective: To evaluate whether imaging DTA studies with positive conclusions or titles have a shorter time to publication than those with non-positive (i.e. negative or neutral) conclusions or titles.

Methods: We included primary imaging DTA studies from systematic reviews published in 2015. We extracted conclusion and title positivity independently in duplicate for each study, based on a previously published classification scheme. We extracted and calculated time from study completion to publication. A Cox regression model was used to evaluate associations of conclusion and title positivity with time to publication, adjusting for potentially confounding variables.

Results: We included 774 imaging DTA studies; time from study completion to publication could be calculated for 516 studies. Median time from completion to publication was 18 months (interquartile range (IQR) 13 to 26; 413 studies) for studies with positive conclusions, 23 months (IQR 16 to 33; 63 studies) for those with neutral conclusions, and 25 months (IQR 15 to 38; 40 studies) for those with negative conclusions (as seen in Table 1). Conclusion positivity was associated with a shorter time from completion to publication for studies with positive conclusions, compared to those with non-positive conclusions (hazard ratio HR 1.31, 95% confidence interval (CI) 1.02 to 1.68), as seen in Figure 1. Positive titles were not significantly associated with a shorter time to publication (HR 1.12, 95% CI 0.75 to 1.69). When subdividing time from study completion to publication, conclusion positivity was associated with shorter time from study completion to submission (HR 1.31, 95% CI 1.00 to 1.74), but no association was identified between conclusion positivity and time from submission to publication (HR 0.95; 95% CI 0.61 to 1.48).

Conclusions: Positive author conclusions (but not titles) were associated with a shorter time to publication. Imaging DTA studies with positive conclusions may be over-represented in the literature, potentially leading to overly optimistic perceptions of the performance of diagnostic imaging tests.

Patient or healthcare consumer involvement: There were no patients or healthcare consumers involved in this project.

Additional file: Figure 1 and Table 1

Reporting of ITT analysis in reproductive studies

Watari N[1], Porfírio GJM[1]

[1] Municipal University of São Caetano do Sul, Brazil

Background: Missing outcomes and non-compliance are complications that most randomized controlled trials suffer after randomization and during a treatment. This problem ends up creating groups that cannot be compared, leading to a risk of bias. One method used to estimate the effect of the intervention in randomized samples, regardless of whether they have finished treatment, is intention-to-treat analysis (ITT). Therefore, as assisted reproduction has been gaining prominence and consequently, the search for new treatments has increased, this article will report the number of randomized controlled trials with that theme that have performed intention-to-treat analysis.

Objectives: To evaluate the quality of description of ITT analysis in randomized controlled trials (RCT) of female infertility.

Methods: We performed a search strategy in MEDLINE via PubMed to identify published RCTs of female infertility. The identified studies were selected by two researchers using Rayyan including RCTs that mentioned ITT analysis.

Results: We included 68 RCTs. Of these, 68% clearly stated the analysis of all randomized participants (95% confidence interval (CI) 56% to 78%). A total of 34% of included RCTs mentioned the use of "per protocol" analysis (CI 95%, 24% to 46%) with only one detailing if "as treated" or "naive per protocol". Four per cent of included RCTs reported the use of a modified ITT analysis.

Conclusion: There is a relevant proportion of RCTs that do not report intention-to-treat analysis properly.

Risk of Bias 2 in Cochrane Intervention Reviews: update on the pilot and implementation in Cochrane

Flemyng E[1], Dwan K[1], Moore TH[2], Hall R[3], Higgins JP[4]

[1] Cochrane, UK; [2] Cochrane; Population Health Sciences Department, Bristol Medical School, University of Bristol, UK; [3] Cochrane, Denmark; [4] Population Health Sciences Department, Bristol Medical School, University of Bristol, UK

Background: Cochrane Reviews are required to assess the risk of bias in the studies that they identify and summarise. In 2008, the publication of the Cochrane Handbook of Systematic Reviews of Interventions and simultaneous rollout of the Review Manager 5 software launched the first version of the Cochrane Risk of Bias tool. Following formal evaluations of its use, a new version of the tool, Risk of Bias 2 (RoB 2), was developed and included in version 6 of the Cochrane Handbook of Systematic Reviews of Interventions that was published in 2019. In 2019 and 2020, as part of a supported, gradual implementation process for RoB 2, Cochrane has been piloting its use in reviews via an opt-in process. The Cochrane Editorial and Methods and Information and Technology Services Departments have been working closely with all pilot review groups to ensure the necessary guidance, training, tools and support are available to authors and editors. The pilot also helps us understand and overcome any dependencies, gather evidence on the usability of the tool and understand the impact it has on technology, publishing and presentation. Fundamentally, the pilot is helping us work towards a streamlined and efficient process from writing the protocol to publishing the full review in the Cochrane Library.

Objectives: To provide an update on developments and next steps on the RoB 2 roll-out in Cochrane, including:

1) details on the number of Cochrane Reviews in the pilot and their feedback on using RoB 2;
2) details on available guidance, training and support;
3) updates on review production tools and developments to support RoB 2 use, including RevMan Web; and
4) details on how RoB 2 will be presented in full published reviews.

The poster will also provide feedback on the approach used to introduce RoB 2 within Cochrane; the gradual, roll-out of a new method by starting with keen early-adopters in a pilot and scaling-up. This will be of use to Cochrane as it considers plans to include more diverse and complex methods within its reviews.

Conclusions: RoB 2 helps authors produce Cochrane Reviews with more concrete conclusions on the reliability of the evidence and therefore has the potential to improve the relevance and quality of reviews generally. The pilot and roll-out plan are observing what challenges the improved tool may pose to authors and editors so that the necessary guidance and infrastructure can be put in place to alleviate them before changes to policy are made.

Patient or healthcare consumer involvement: As a methods implementation project within Cochrane processes, we have not involved patients. However, our 'consumers' are Cochrane Review author and editorial teams, who are involved as pilot teams within the project. Cochrane hosts monthly web clinics for the pilot teams to give them the opportunity to ask questions and so we can gain their feedback on guidance and developments as part of the pilot. We also actively seek their feedback via email

Risk of bias and PRISMA compliance in systematic reviews of rare disease

Leadley R[1], de Kock S[1], Swift S[1], Kleijnen J[1]

[1] Kleijnen Systematic Reviews Ltd, UK

Objective: To investigate the risk of bias (RoB) in systematic reviews (SRs) conducted in rare diseases using a random sample of SRs (and their corresponding quality appraisals using the risk of bias in systematic reviews (ROBIS) tool) and assess their compliance with PRISMA reporting standards.

Methods: We identified two random samples of SR publications from 2016 and 2017 from the KSR Evidence database of systematic reviews (ksrevidence.com). We screened these for systematic reviews of rare diseases and confirmed rare disease status using ORPHANET (a unique database for rare diseases and orphan drugs). We extracted data on review type, study characteristics, ROBIS quality assessment and PRISMA compliance.

Results: From 1026 SRs, we identified 25 (2.4%) that reported on rare diseases. These included: systemic sclerosis, cystic fibrosis, Guillain-Barré syndrome, myasthenia gravis, Kawasaki disease, Churg Strauss syndrome, Maroteaux-Lamy syndrome and rare cancers such as Burkitt's lymphoma, sarcoidosis and non-Hodgkin lymphoma. Of the 25 SRs, ten (40%) were interventional studies, two (8%) diagnostic, two (8%) epidemiological and 11 (44%) investigated other or multiple research questions. Two (of 25) SRs (8%) were at low risk of bias, two (8%) unclear risk of bias and 21 (84%) high risk of bias. The main areas of concern were not reporting search strategies, language limitations and no or inappropriate risk-of-bias assessment of included studies. Forty per cent of studies stated that they were PRISMA complaint; however, 80% of these did not report a search strategy and 40% did not assess risk of bias.

Conclusion: Systematic reviews of rare diseases represented approximately 2% of all reviews. Most studies were at high risk of bias. Given the paucity of research in this area, it is important to encourage good quality research and highlight the areas of concern. There are five relatively simple ways to potentially improve the risk of bias of rare disease systematic reviews:

1) report a full search strategy;
2) do not apply language restrictions;
3) include conference abstracts;
4) use an appropriate risk-of-bias tool; and
5) avoid inappropriate pooling. It is not sufficient to state a study is PRISMA compliant; compliance must be demonstrated.

Patient or healthcare consumer involvement: None

Risk of bias in trials of chronic medical conditions: a meta-epidemiologic study

Murad MH[1], Alahdab F[1], Wang Z[1]

[1] Mayo Clinic, USA

Background: Previous research has suggested that increased risk of bias may be associated with exaggerated effect size. However, this association is unclear in trials about treatments of chronic medical conditions. These conditions are associated with significant morbidity, mortality and burden. Knowing the status and direction of bias can affect the confidence of end users in this body of evidence.

Objectives: To evaluate the risk of bias in randomized trials evaluating a drug or a device used to treat chronic medical conditions and to determine if this risk is associated with the magnitude of treatment effect.

Methods: We identified meta-analyses with at least five randomized controlled trials (RCTs) published between 2007 and 2015 in 10 high impact general medical journals. Meta-analyses had to evaluate a medication or device for chronic medical conditions. We used Cochrane's risk-of-bias tool to evaluate the included RCTs. We used mixed-effects random intercept regressions to evaluate the association of bias judgments (high risk, unknown risk and low risk) and the effect size of individual RCT.

Results: We analyzed risk of bias of 930 RCTs [average of 13 RCTs (5 to 48) and 922 patients (10 to 20,536) per meta-analysis]. Only a small proportion of the RCTs received a clear judgment of high risk of bias (2% to 14% across domains) but a substantial proportion had unknown risk of bias judgment (notably, allocation concealment: 62%). Despite the large number of included RCTs in regression analysis, there was no statistically significant association between any of the seven items of risk of bias and the effect size (ratios of odds ratios with 95% confidence intervals overlapping 1.0, Table 1).

Table 1. The association between risk of bias and effect size

Items in Cochrane Risk of Bias tool	High risk	Unknown risk	Low risk	High risk vs. Low risk (ratio of odds ratio and 95% CI)	Unknown risk vs. Low risk (ratio of odds ratio and 95% CI)
Sequence generation	2.04%	35.91%	62.04%	0.94 (95% CI: 0.66 - 1.35)	0.92 (95% CI:0.82 - 1.04)
Allocation concealment	2.80%	61.83%	35.38%	0.93 (95% CI: 0.66 - 1.31)	0.99 (95% CI:0.88 - 1.11)
Participants and personnel blinding	13.66%	33.23%	53.12%	0.89 (95% CI: 0.75 - 1.05)	0.91 (95% CI:0.80 - 1.03)
Outcome assessment blinding	13.12%	40.97%	45.91 %	1.00 (95% CI:0.83 - 1.20)	0.95 (95% CI:0.83 - 1.09)
Incomplete outcome data	4.84%	26.67%	68.49%	0.87 (95% CI: 0.68 - 1.12)	0.84 (95% CI: 0.74 - 0.96)
Selective reporting	2.58%	36.13%	61.29%	1.15 (95% CI: 0.84 - 1.58)	0.96 (95% CI: 0.85 - 1.09)
Other sources of bias	7.53%	41.94%	50.54%%	0.83 (95% CI: 0.66 - 1.04)	0.91 (95% CI: 0.79- 1.04)

Conclusions: The reporting of randomized controlled trials continues to show a substantial amount of unknown or unclear ratings. Meta-analyses about treatments of chronic medical conditions depend on trials with a small proportion of high risk-of-bias ratings. The lack of association between ratings and the effect size suggests that the direction of bias remains unpredictable.

Patient or healthcare consumer involvement: Not performed.

Risk-of-bias judgments for blinding of outcome assessors in Cochrane Reviews were often inadequate

Barcot O[1], Dosenovic S[2], Boric M[1], Poklepovic Pericic T[3], Cavar M[4], Jelicic Kadic A[5], Puljak L[6]

[1] Department of Abdominal Surgery, University Hospital Split, Split, Croatia; [2] Department of Anaesthesiology and Intensive Care, University Hospital Split, Split, Croatia; [3] Department of Research in Biomedicine and Health, University of Split School of Medicine, Split, Croatia; [4] Department of Radiology, University Hospital Split, Split, Croatia; [5] Department of Pediatrics, University Hospital Split, Split, Croatia; [6] Center for Evidence-Based Medicine and Health Care, Catholic University of Croatia, Croatia

Background: Cochrane's risk-of-bias (RoB) tool is used for assessment of randomized controlled trials (RCTs). The 2011 version of the tool had seven domains, including blinding of outcome assessors (detection bias). This domain, with modifications, is used in the new RoB 2.0 tool as well. Our research group has previously shown that Cochrane authors do not assess other RoB domains adequately in Cochrane Reviews, but there were no prior studies about adequacy of judgments for detection bias in Cochrane Reviews.

Objective: To analyze whether judgments about the risk of bias associated with blinding of outcome assessors in Cochrane Reviews of RCTs were adequate, i.e. in line with recommendations from the Cochrane Handbook for Systematic Reviews of Interventions (the Handbook).

Methods: This was a primary methodological study, in which we analyzed the methodology of Cochrane Reviews published in 2016/2017 in the Cochrane Database of Systematic Reviews (CDSR). We extracted judgments and supporting comments for detection bias from RoB tables reported in Cochrane Reviews. For data extraction we used customized software. We categorized comments, and then compared judgment and supporting comment with instructions from the Handbook.

Results: We analyzed 8656 judgments for detection bias from 7626 trials included in 575 Cochrane Reviews. In those 7626 trials there were 8656 domains (judgments) for detection bias, because in some Cochrane Reviews this domain was split, i.e. had multiple assessments for various types of outcomes. Overall, 1909 judgments (22%) were not in line with the Handbook. The highest prevalence of inadequate judgments was found for trials judged with low risk of detection bias (1287 of 3374; 38%), followed by those judged with high risk (239 of 1679; 14%) and those judged with unclear risk (383 of 3603; 11%). In 9% of trials authors split the detection bias domain according to outcomes. Here, prevalence of inadequate judgments was 19%.

Conclusions: We found that Cochrane Reviews frequently had inadequate judgments for risk of detection bias. It would be worthwhile to explore interventions that would help ensure adherence to methodological guidance among systematic review authors. Risk-of-bias judgments are incorporated into systematic review conclusions, and it is in the interest of the entire medical community to have trustworthy evidence.

Patient or healthcare consumer involvement: This was a research methodology study and it did not include patient or healthcare consumers.

Spin in systematic reviews and health technology assessments about robotic thoracic surgery: a systematic review

Cömert D[1], Laane AR[1], Hartemink KJ[2], Langendam MW[1]

[1] Department of Clinical Epidemiology and Biostatistics, Amsterdam UMC, University of Amsterdam, The Netherlands; [2] Department of Surgery, Netherlands Cancer Institute – Antoni van Leeuwenhoek Hospital, The Netherlands

Background: Quality of reporting of studies is essential to maintain accuracy and transparency. However, researchers can choose which data to report and how to report them. Spin is defined as a specific way of reporting, implying that the beneficial effect of the experimental treatment is greater than that shown by the results. The presence of spin in research articles can negatively impact the development of further studies, clinical practice and health policies. Systematic reviews (SRs) summarize all available evidence and are the cornerstones of therapeutic evaluations. However, the interpretation of the findings is vulnerable to spin.

Objectives: To give an overview of the prevalence of spin in SRs about robotic thoracic surgery.

Methods: We searched MEDLINE, Embase, Epistemonikos and the Centre for Reviews and Dissemination Database (University of York) for eligible studies . SRs about robotic thoracic surgery published between 2000 and February 2020 were eligible for inclusion. We used the checklist published by Yavchitz et al. (2016) to identify spin.

Results: Out of a total of 2480 articles, we included 13 thoracic surgery SRs. Most SRs (10/13) investigated robotic surgery for lobectomy. In 10 SRs we found spin items for misleading reporting and misleading interpretation. In the abstract, selective reporting of, or overemphasis on efficacy outcomes favoring in the beneficial effect of the experimental intervention and a conclusion focusing selectively on statistically significant efficacy outcomes were the most frequently scored items (46%). Misleading interpretation was mostly scored as a result of a conclusion claiming equivalence, comparable effectiveness and/or safety for non-statistically significant results with a wide confidence interval (62%). Additionally, authors tend to focus on a relative effect when the absolute effect is small (62%). In the main text, misleading reporting was mostly a result of changing the scale of the forest plot to magnify the results (54%) and a conclusion focusing selectively on statistically significant efficacy outcome (46%). Misleading interpretation was a result of focusing on a relative effect when the absolute effect is small (69%) and a conclusion claiming equivalence or comparable effectiveness (69%) or safety (54%) for non-statistically significant results with a wide confidence interval.

Conclusions: Our analysis suggests a high prevalence of spin in the conclusion of abstracts and main texts of SRs of thoracic robotic surgery. Misleading reporting and interpretation are the most common categories of spin. Claims for equivalence or beneficial effect of robotic thoracic surgery did not take into account factors that lower the certainty in in the evidence, like bias, imprecision and inconsistency.

Patient or healthcare consumer involvement: Knowledge of how to identify inappropriate interpretation of the results of a SR (spin) is of importance to patients for making well-informed healthcare decisions.

The use of complementary checklists for data extraction and assessment of risk of bias and applicability of prediction model studies

Scheers H[1], Van Remoortel H[1], Vandekerckhove P[2], De Buck E[3]

[1] Centre for Evidence-Based Practice, Belgian Red Cross, Belgium; [2] Belgian Red Cross; Department of Public Health and Primary Care, Faculty of Medicine, KU Leuven, Belgium; [3] Centre for Evidence-Based Practice, Belgian Red Cross; Department of Public Health and Primary Care, Faculty of Medicine, KU Leuven; Cochrane First Aid, Belgium

Background: Prognostic prediction models require a specific approach for evaluation in a systematic review. The CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies (CHARMS) provides guidance for data extraction from prediction model studies. The Prediction model Risk Of Bias ASsessment Tool (PROBAST) is a complementary tool for in-depth assessment of risk of bias (RoB) and applicability of such studies.

Objectives: To systematically summarize results and assess RoB and applicability of studies that developed and/or validated a regression model predicting patient presentation rate (PPR) or transfer to hospital rate (TTHR) at mass gatherings, using the CHARMS and PROBAST checklists.

Methods: According to seven key items of the CHARMS checklist, we systematically searched and classified model development and model validation studies, and extracted predictors for PPR or TTHR from multivariable regression models. We used the PROBAST checklist to assess RoB and applicability in four domains (participants, predictors, outcome, and analysis). We implemented overall RoB and applicability judgement into the GRADE tool.

Results: We identified 13 prediction model development studies without validation and 3 external validation studies of existing models, comprising over 1700 mass gatherings. Main predictors of PPR and/or TTHR were accommodation (e.g. indoor vs outdoor), type of event (e.g. music concerts), and weather conditions (e.g. temperature). PROBAST domains that were most prone to bias were the method of analysis used and selection of participants (high RoB in 13 (81%) and 10 (62%) studies, respectively). Consequently, we judged overall RoB as 'high' for all included studies (Figure 1). Overall concerns for applicability were high in 12 studies (75%), mainly due to high concerns for selection of participants (9 studies, 56%) (Figure 2). The initial GRADE level of certainty in the body of evidence was set at 'high'. We downgraded with two levels due to the high overall RoB (-1) and concerns regarding applicability (GRADE domain 'indirectness', -1), ending up with 'low' certainty in the effect estimates.

Conclusions: The CHARMS and PROBAST checklists proved to be useful for data extraction and assessment of RoB and applicability in prediction modelling studies on medical usage at mass gatherings. As such, they are complementary with the GRADE evaluation of the body of evidence.

Patient or healthcare consumer involvement: This systematic review and the development and validation of a proper prediction model, is conducted in collaboration with the Relief Service at the Belgian Red Cross, which co-ordinates the preventive aid campaigns at mass gatherings in Flanders (Belgium). Regular meetings with central co-ordinators and representatives of local volunteers helped identifying strengths and weaknesses of the current databases and desired features of our own prediction model for medical usage rate at mass gatherings.

Additional files:

Treatment effect estimates for nontruncated trials need adjustment when conducting sensitivity analyses to assess risk of bias in early-stopped trials

Schou IM[1], Marschner I[2], Askie L[2]

[1] Janssen-Cilag, Sydney, Australia; [2] National Health and Medical Research Council (NHMRC) Clinical Trials Centre, University of Sydney, Australia

Background: Stopping studies early due to an apparent treatment benefit (truncated studies) may lead to overestimation of the treatment effect and thus a risk of bias. GRADE guidelines recommend sensitivity analyses in which truncated studies are omitted from meta-analyses, to assess whether early stopping has caused overestimation bias. However, GRADE recommendations do not address the assessment of studies that were subjected to interim monitoring but did not stop early (nontruncated studies). Such studies lead to underestimation which may balance the overestimation from truncated studies.

Objectives: To investigate how sensitivity analyses of nontruncated studies should be undertaken to adjust treatment effect estimates for the underestimation that results from statistical conditioning on nontruncation.

Methods: Simulation studies generated conditional and unconditional probability distributions of treatment effect estimates for randomized controlled trials (RCTs) that were monitored for early stopping due to benefit, with a maximum number of (equally spaced) analyses between two and five. Outcomes were assumed to have a normal distribution and a moderate effect size of 0.25. Analyses were based on 100 000 simulations of RCTs with 90% power and 5% significance level, corresponding to a sample size of approximately 350 per treatment group. For each simulated RCT, early stopping due to benefit occurred if the estimated treatment effect at an interim analysis was sufficiently large (using the O'Brien-Fleming rule) in the direction of benefit. We then conducted meta-analyses on collections of 4, 12, and 24 simulated studies. In each collection, a proportion of studies (from 25% to 75%) was subjected to interim monitoring for benefit, up to a maximum of three equally spaced interim analyses. For each scenario, we performed 1000 meta-analyses using each of four meta-analysis strategies: omitting nontruncated studies (crude), restricted to studies with no interim monitoring (restricted), adjusting nontruncated studies (adjusted) and including all studies (all-study), with both fixed- and random-effects models.

Results: Figure 1 illustrates that the crude strategy led to underestimation of treatment effects (red box plots). The other three strategies yielded meta-analysis estimates that were approximately unbiased. The all-study strategy (blue) yielded estimates that were the least variable. The adjusted strategy (green) exhibited less variation than the restricted strategy (yellow).

Conclusions: The primary meta-analysis in a systematic review should involve all studies, including those that stopped early for benefit. If a sensitivity analysis is conducted, treatment effect estimates from nontruncated studies subjected to interim analyses should first be statistically adjusted to ensure the meta-analysis is unbiased. Researchers should report all details required for statistical adjustment of treatment effect estimates when reporting studies that had interim monitoring.

Patient or healthcare consumer involvement:–Nil – methodological study.

Additional file: Figure 1

STATISTICAL METHODS

A survey of Cochrane editors revealed several problem areas related to time-to-event (meta-)analyses

Goldkuhle M[1], van Dalen EC[2], Macbeth F[3], Skoetz N[4]

[1] University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Internal Medicine, Center for Integrated Oncology Aachen, Bonn, Cologne, Dusseldorf; Cochrane Haematology, Cologne, Germany; [2] Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands; [3] Centre for Trials Research, Cardiff University, Cardiff, UK; [4] University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Internal Medicine, Center for Integrated Oncology Aachen, Bonn, Cologne, Dusseldorf; Cochrane Cancer, Cologne, Germany

Background: Previous work demonstrated flaws in the analysis and interpretation of time-to-event outcomes by authors of recent Cochrane Reviews. As part of a Cochrane Networks Innovation Fund project, we conducted a survey of Cochrane editors.

Objectives: To identify which training and informational resources around time-to-event analyses editors use and suggest to Cochrane Review authors. Furthermore, to identify difficulties and questions editors frequently encounter from review authors or face themselves.

Methods: Based on previous methodological work we developed a survey that included 27 quantitative and narrative items. A general part of the survey addressed questions on review production, methodological expertise as well as relevant training and informational resources. The specialized part consisted of questions that addressed particular difficulties with time-to-event analyses that could arise along the review development process. We distributed the survey among the editorial staff of all Review Groups within the Cochrane Cancer Network. Furthermore, we invited senior editors of other Networks to forward it to editors if they found it relevant to their work. We discussed and appraised the results of the survey.

Results: Overall, we received 12 responses (all except one from within the Cancer Network). The majority of respondents were statistical editors. Most respondents rated their knowledge on time-to-event analyses three on a scale of five ("familiar with the basic methods"). Review Groups frequently recommend specific time-to-event training materials to authors, of which the most prominent one is the instruction paper authored by Tierney and colleagues in 2007 (1). Difficulties and questions were identified in all areas addressed by the survey. The most problematic according to quantitative items were the underlying assumptions of analytic methods (67%; 8/12), the reconstruction of data from primary reports (58%; 7/ 12) and the interpretation of effects (58%; 7/12). Besides the reconstruction of survival data, narrative responses revealed the reconstruction of survival data, the proportional hazards assumption, the timing of randomization, competing events, the concept of censoring and absolute effects as areas with increased uncertainty.

Conclusions: Our findings suggest that Cochrane editors frequently encounter issues with the crucial concepts underlying time-to-event analyses. Several particularly problematic areas are not yet covered by training and information resources. However, in the course of our Cochrane Networks Innovation Fund project targeted resources are currently under development.

Patient or healthcare consumer involvement: None.

Reference:

1. Tierney JF, Stewart LA, Ghersi D, Burdett S, Sydes MR. Practical methods for incorporating summary time-to-event data into meta-analysis. *Trials* 2007;8:16.

Adjusting for exposure misclassification in an individual participant data meta-analysis of observational studies

De Jong VM[1], Campbell H[2], Jaenisch T[3], Gustafson P[2], Debray TP[1]

[1] Universitair Medisch Centrum (UMC), Utrecht, The Netherlands; [2] The University of British Columbia, Canada; [3] Universität Heidelberg, Germany

Background: A common problem in the analysis of multiple data sources, including individual participant data meta-analysis (IPD-MA), is the presence of misclassification of binary variables. Misclassification may bias estimates of parameters (including covariate effects), even when the misclassification is entirely random. Available methods for addressing misclassification in the analysis of exposure-outcome associations do not account for between-study heterogeneity in IPD-MA.

Objective: We aimed to develop statistical methods that facilitate unbiased estimation of adjusted and unadjusted exposure-outcome associations and between-study heterogeneity in IPD-MA where the extent and nature of exposure misclassification may vary across or within studies.

Methods: We present Bayesian methods that allow misclassification of binary exposure variables to depend on study- and participant-level characteristics. We illustrate this in an example of the differential diagnosis of dengue using two variables, where the gold standard measurement for the exposure variable is unavailable for some studies which only measured a surrogate prone to misclassification. We present a simulation study to assess bias, root mean square error (RMSE), coverage and power in estimating an exposure-outcome association.

Results: In the example, our methods yielded estimates with less error than analyses naive with regard to misclassification or based on gold standard measurements alone. In our simulations, the evaluated misclassification model yielded valid estimates of the true exposure-outcome association, with less RMSE, greater power and similar coverage compared to an analysis restricted to available gold standard measurements.

Conclusions: Our proposed framework can appropriately account for the presence of binary exposure misclassification in IPD-MA. It requires that 1) some studies supply IPD for the surrogate and gold standard exposure and 2) misclassification is exchangeable across studies conditional on observed covariates (and outcome). Further work is needed to address other types of misclassification.

Patient or healthcare consumer involvement: Not applicable; we developed new statistical methods for researchers.

Can rapid-learning health systems accommodate individual participant data? An example of the timelines involved in an IPD meta-analysis

Cruickshank M[1], Hudson J[1], Aucott L[1], Jayasena C[2], Bhattacharya S[1], Gillies K[1], Brazzelli M[1]

[1] University of Aberdeen, UK; [2] Imperial College, London, UK

Background: Rapid-learning health systems are underpinned by timely sharing of data and synthesis of existing evidence. Individual participant data (IPD) meta-analysis (MA) allows robust evaluation of treatment effects and offers numerous advantages both clinically and statistically over aggregate data MA. Acquiring IPD from triallists has become more time consuming since the introduction of the General Data Protection Regulation (GDPR) in 2018. IPD may be regarded as incompatible with the principles of rapid-learning health systems. **Objectives:** To present an example of the timeline required for collection of IPD for a two-year UK NIHR-funded evidence synthesis hosted by two academic institutions.

Methods: Systematic review and IPD MA of placebo-controlled randomized controlled trials of testosterone therapy in people with low testosterone levels. We contacted the authors of all eligible trials to request IPD. Those interested were sent a GDPR-compliant data sharing agreement to be signed by a representative of the respective institution. The initial length of the agreement (1939 words, 6 pages) was prohibitive so it was revised to a more manageable length (1423 words, 4 pages). Nevertheless, various international institutions requested further specific amendments to the agreement, each of which required perusal by the contract departments of the two host universities and added to the original timeline.

Results: We identified 35 eligible trials in the current literature. At 12 months since the start of the IPD study (September 2018), signed data sharing agreements (and datasets) from five collaborators had been received. At 18 months, six further signed agreements had been received, providing data from around one-third of the total number of participants enrolled (lower than the 60% recommended for robust MA of IPD). Two of the 11 received datasets did not have a corresponding signed agreement (subsequently obtained); one was not anonymized and was deleted and re-requested. At 18 months, negotiations regarding two further data sharing agreements were still ongoing. Another data sharing agreement was refused due to concerns about its legality. A further triallist, who had initially agreed to provide IPD, did not return the data sharing agreement or respond to follow-up emails.

Conclusions: Rapid-learning health systems may benefit from the advantages conferred by the robustness of IPD analyses. However, the process of obtaining data from international investigators by consenting to a data sharing agreement can be prohibitive in terms of time required for data collection and synthesis. Early recruitment of potential sources of IPD is recommended, including both eligible investigators and their respective contract department, in order to identify and address potential issues in an effective and timely manner.

Patient or healthcare consumer involvement: Opportunities may be missed for full utilization of patient data and the potential benefits to healthcare consumers and policy makers.

Does the inverse of Freeman-Tukey double arcsine transformation provide misleading results in meta-analysis of single proportions?

Bardach A[1], Ciapponi A[1], Glujovsky D[1], Gibbons L[2]

[1] Instituto de Efectividad Clínica y Sanitaria (IECS-CONICET), Argentina; [2] Institute for Clinical Effectiveness and Health Policy (IECS-CONICET), Argentina

Background: Standard generic inverse variance methods for the combination of single proportions are based on transformed proportions using the logit, arcsine, and Freeman-Tukey double arcsine transformation and more recently generalized linear mixed models. Irrespective of the approach, meta-analysis results are typically back-transformed to the original scale in order to facilitate interpretation. A recent study suggested that the Freeman-Tukey double arcsine transformation, one of the most frequently used approaches, could provide misleading results after the back-transformation (1).

Objectives: To evaluate the consistency of the back-transformed results obtained with the inverse of Freeman-Tukey double arcsine transformation with alternative methods of meta-analyses of single proportions.

Methods: We will analyze the results of five meta-analyses of single proportions obtained by the inverse of Freeman-Tukey double arcsine transformation (2) and we will compare with the results obtained by the logit transformations. Additionally, we will analyze if the sample size could be influential in the potential discrepancies

Results: The comparison of the results by method will be presented at the Colloquium.

References:

1. Schwarzer G, Chemaitelly H, Abu-Raddad LJ, Rücker G. Seriously misleading results using inverse of Freeman-Tukey double arcsine transformation in meta-analysis of single proportions. *Research Synthesis Methods* 2019;10(3):476-83.
2. Ciapponi A, Glujovsky D, Virgilio SA, Bardach AE. Conducting and disseminating epidemiological systematic reviews in Latin America and the Caribbean: pitfalls and lessons learned. *Value in Health Regional Issues* 2017;14:64-72.

Empirical evaluation of five statistical models in meta-analyses of treatments of chronic medical conditions: a meta-epidemiological study

Wang Z[1], Murad MH[1]

[1] Mayo Clinic, USA

Background: Meta-analysis has become an important tool to support decision-making. Different meta-analysis models have been used to address heterogeneity between studies; which may provide results with varying precision and subsequently different certainty in evidence. This can affect decisions made about the care of patients with chronic medical conditions. These models have been evaluated in simulation studies, not empirically. Chronic medical conditions are associated with important morbidity, mortality and public health burden.

Objectives: To empirically evaluate the precision of five meta-analysis methods. Fixed-effect methods included the inverse variance method (IV) and the t distribution (IVT). Random-effects methods included those by DerSimonian and Laird (DL), Hartung-Knapp-Sidik-Jonkman (HKSJ) and the profile likelihood (PL).

Methods: We identified meta-analyses about treatments of chronic medical conditions published in from 2007 to 2019 in the 10 medical journals with the highest impact factor. We included meta-analyses with at least five randomized controlled trials and chose one binary outcome deemed to be most important to patients (e.g. mortality, stroke, and myocardial infarction). Each meta-analysis was performed using the five methods. We defined discordance between methods when either boundary of 95% confidence interval of the relative risk reduction changed by more than 0.15 (an arbitrary threshold of clinical importance). We also evaluated changes of statistical significance (two-tailed p value < 0.05) across methods.

Results: We identified 88 meta-analyses with 1114 RCTs (average 12.60 RCTs per meta-analysis. The average I^2 was 26% (interquartile range: 0% to 46%). The PL method failed to converge in 3% of meta-analyses. Discordance between the IV and IVT method occurred in 18.18% of meta-analyses, 32.95% between the DL and HKSJ method, and 18.82% between the DL and PL method. Changes from statistical significance to non-significance occurred in 6.82% of meta-analyses (from IV to IVT method), 10.23% (from DL to HKSJ), and 4.55% (from DL to PL). When analysis was limited to the earliest five RCTs within a meta-analysis, we found increased discordance and changes of statistical significance (Table 1, Figure 1).

Conclusions: The precision of pooled estimates is a key component of the construct of overall certainty in evidence. Empirical evaluation of studies of chronic medical conditions showed that precision frequently changed when different pooling methods were used, particularly when the number of studies within a meta-analysis was small. The variation in trial population and settings provides rationale for using the random-effects model. Therefore, sensitivity analyses using more than one random-effects method are highly recommended. If results were not robust in terms of precision, certainty in the evidence should be lowered.

Additional file: <u>Figure 1</u>; <u>Table 1</u>

Exclusion of trials with no events in both arms from meta-analyses can change the conclusions

Xu C[1], Li L[1], Lin L[2], Chu H[3], Lehana T[4], Zou K[5], Sun X[1]

[1] Chinese Evidence-Based Medicine Center and Cochrane China Center, West China Hospital, Sichuan University, Chengdu, China; [2] Department of Statistics, Florida State University, Tallahassee, USA; [3] Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, USA; [4] Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton; Biostatistics Unit, Father Sean O'Sullivan Research Centre, St Joseph's Healthcare, Canada; [5] Chinese Evidence-based Medicine Center and Cochrane China Center, West China Hospital, Sichuan University, Chengdu, China

Objectives: Classical meta-analyses routinely treat trials with no events in both arms as non-informative and exclude them from analyses. This study assessed whether such trials are statistically meaningful and have any influence on conclusion of meta-analyses.

Design and setting: We collected meta-analyses of binary outcomes with at least one trial having no events in both arms from Cochrane systematic reviews published between 2003 and 2018. We used the generalized linear mixed model to reanalyze these meta-analyses by two approaches: one including studies with no events in both arms and one excluding such studies. We compared the magnitude and direction of odds ratio (OR), P value, and the width of 95% confidence interval (CI). We conducted a simulation study to examine the robustness of results.

Results: We identified 442 meta-analyses. In comparing the meta-analyses that included studies with no events in both arms versus those that did not, the flipping of direction occurred in eight (1.80%) comparisons; 41 (9.28%) altered conclusions on statistical significance. Substantial changes in P value occurred (55.66% increased, 44.12% decreased) and the width of 95% CI (50.68% inflated, 49.32% declined) when excluding studies with no events. The simulation study confirmed these findings.

Conclusions: Studies with no events in both arms are not necessarily uninformative. Excluding such studies may alter conclusions.

Exploring the empirical distribution of tau from IQWiG reports for the application in Bayesian (network) meta-analyses

Bender R[1], Sturtz S[1], Kiefer C[1]

[1] Institute for Quality and Efficiency in Health Care (IQWiG), Germany

Background: Meta-analysis and network meta-analyses are the methods of choice in systematic reviews to summarize the effect estimates of the included studies. Frequently, random-effects meta-analyses and network meta-analyses are applied, which require the estimation of the heterogeneity parameter tau. However, in the case of very few studies, the heterogeneity parameter cannot be reliably estimated leading to broad confidence intervals (Bender et al., 2018). In such situations, the application of Bayesian methods with informative prior distributions is an option (Friede et al., 2017; Bender et al., 2018). Different choices for prior distributions for tau are possible according to several proposals given in the literature (e.g. Turner et al., 2015; Friede et al., 2017; Rhodes et al., 2018).

Objectives: To explore the empirical distribution of tau from Institute for Quality and Efficiency in Health Care (IQWiG) reports in order to inform future Bayesian (network) meta-analysis in the case of few studies.

Methods: We collected all published meta-analyses from IQWiG reports for the years 2005 to 2019 and recalculated the estimates of tau by applying random-effects meta-analyses and the Paule-Mandel method. Sensitivity and subgroup analyses were not taken into account. We used the effect measures as used in the original meta-analysis. In the case of binary data, we calculated the risk ratio and the odds ratio for the same data. We summarized the empirical distributions of tau in various settings (comparison, endpoint category, effect measure) and compared these distributions with the proposals for prior distributions in the literature.

Results: Different empirical distributions of tau can be derived from IQWiG reports in various settings. Descriptive analyses of the various distributions will be reported at the Colloquium.

Conclusions: It should be discussed in which situations prior distributions for Bayesian meta-analyses and network meta-analyses in the framework of health technology assessment can be derived from the empirical distributions of tau from IQWiG reports in various settings.

Patient or healthcare consumer involvement: Not applicable

References:

1. Bender R. Friede T, Koch A, Kuss O, Schlattmann P, Schwarzer G, et al. Methods for evidence synthesis in the case of very few studies. *Research Synthesis Methods* 2018:9,382-92.
2. Friede T. Röver C, Wandel S, Neuenschwander B. Meta-analysis of few small studies in orphan diseases. *Research Synthesis Methods* 2017;8:79-91.
3. Rhodes KM, Turner RM, Savović J, Jones HE, Mawdsley D, Higgins JPT. Between-trial heterogeneity in meta-analyses may be partially explained by reported design characteristics. *Journal of Clinical Epidemiology* 2018;95:45-54.
4. Turner RM, Jackson D, Wei Y, Thompson SG, Higgins JPT. Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis. *Statistics in Medicine* 2015;34:984-98.

How does meta-analysis handle studies with zero events in both arms? A systematic survey of Cochrane systematic reviews

Xu C[1], Li L[1], Sun X[1]

[1] Chinese Evidence-Based Medicine Center and Cochrane China Center, West China Hospital, Sichuan University, Chengdu, China

Background: Meta-analysis of rare events represents a challenging issue, particularly in the presence of studies having no events in both arms. However, it remains unclear how such studies were handled in the published meta-analyses; a systematic understanding may improve future practice.

Methods: We searched Cochrane systematic reviews published from January 2003 to May 2018 for meta-analyses that had at least one study with zero events in both arms. We extracted original data (in .rm5 format) from each eligible meta-analysis, including significance of the pooled effect, between study heterogeneity, effect measures, and how studies with zero events in both arms were handled. All these data were used only for academic research purposes.

Results: We identified 831 meta-analyses that had one or more studies with zero events in both arms. Of these, 206 (24.79%) were statistically significant ($P < 0.05$) in the pooled estimates and 49 (5.9%) trended to be significant ($0.05 < P < 0.1$); 75 (9.03%) chose odds ratio as an effect measure, 410 (49.34%) risk ratio, 343 (41.28%) risk difference, and three (0.36%) hazard ratio. Of the 831 meta-analyses, 458 (55.11%) removed studies with zero events in both arms, and 343 (41.28%) used the continuity correction to deal with such studies. No other methods (e.g. Bayesian, GLMM) were used.

Conclusion: The most common approaches to handling studies with no events in Cochrane systematic reviews were the exclusion of such studies or continuity correction. Other advanced methods were not used.

I² statistic in meta-analysis of prevalence: worthwhile or worthless?

Borges Migliavaca C[1], Stein C[2], Colpani V[2], Hugh Barker T[3], Munn Z[3], Falavigna M[2]

[1] National Institute for Health Technology Assessment, Post-Graduate Program in Epidemiology, Federal University of Rio Grande do Sul, Brazil; [2] Hospital Moinhos de Vento, Brazil; [3] Joanna Briggs Institute, University of Adelaide, Australia

Background: Prevalence estimates are critical for health decision making, with systematic reviews (SR) and meta-analyses being useful to generate a mean estimate of prevalence. Heterogeneity is an important aspect in meta-analysis, and one way it is usually assessed is with the I² statistic.

Objectives: To describe and evaluate the use of the I² statistic to assess heterogeneity in meta-analysis of prevalence.

Methods: This is a work from the Prevalence Estimates Reviews – Systematic Review (PERSyst) Methodology Group. We searched PubMed using the terms 'prevalence' and 'systematic review' in the title, from February 2017 to February 2018. We included SR on the prevalence of any clinical conditions published in English. If the SR conducted a meta-analysis, we extracted data regarding the assessment of heterogeneity. For the analysis, we classified the I² as high (> 50%) or low (≤ 50%). We then used the Mann-Whitney test to assess the association between the I² result and the number of studies included in each meta-analysis.

Results: We included 235 SRs; 152 performed meta-analysis, and 144 assessed heterogeneity through I², according to the description of their methodology. However, only 134 presented the I² result for their main meta-analysis. The median I² was 96.9% (interquartile range (IQR) 90.5 to 98.7). Seven meta-analyses (5%) presented I² ≤ 50%; 3 (2%) presented I² from 50% to 70%; and 124 (93%) presented I² > 70%. Of note, 102 meta-analyses (76%) presented I² higher than 90%. There was an association between the number of studies included in the meta-analysis and the level of I²: meta-analyses with I² > 50% included more studies (median 19, IQR 10 to 28) than meta-analyses with I² ≤ 50% (median 9, IQR 6.5 to 9.5; P = 0.004). All meta-analyses with more than 21 included studies presented I² > 50% (Table 1). Despite the high inconsistency observed, only 3 (2%) SRs reported prediction intervals.

Table 1. I² assessment according to the number of studies included in the meta-analysis

Studies included in the meta-analysis	Number of SRs	I² (median, IQR)	I² ≤50% n (%)	I² >50% n (%)
02 to 10	41	92,2 (82,4 to 97,8)	6 (14.6%)	35 (85.4%)
11 to 20	35	97,7 (93,6 to 98,8)	1 (2.9%)	34 (97.1%)
21 to 30	31	97,7 (95,1 to 99,1)	0	31 (100.0%)
31 or more	27	97,3 (95,7 to 98,8)	0	27 (100.0%)

SR: systematic reviews; I²: I-square; IQR: interquartile range.

Conclusions: Overall, meta-analyses of prevalence commonly present high inconsistency. This can be due to the nature of proportional data, where due to large datasets precise estimates are often provided, and small variance is observed even in studies with small sample size; this leads to minimal overlap of confidence intervals in these types of meta-analysis. Moreover, true heterogeneity is expected in prevalence estimates due to differences in the time and place where included studies were conducted. I² statistics may not be discriminative and should be interpreted with caution in this case. Prediction intervals are a more conservative way to incorporate uncertainty in the analysis when true heterogeneity is expected; however, it is still underused in meta-analysis of prevalence. Whilst our study was limited to the evaluation of SRs of prevalence, we expect similar conclusions for reviews of other proportions (such as incidence). **Patient or healthcare consumer involvement:** Prevalence estimates play a key role in supporting healthcare decision making. Understanding the underlying heterogeneity of this data is critical to decision making.

Identifying common statistical errors in Cochrane Reviews of interventions

Anglemyer A[1], Moore T[2], Dwan K[2]

[1] Cochrane Methods Support Unit, New Zealand; [2] Cochrane Methods Support Unit, UK

Background: Meta-analysis (MA) in Cochrane systematic reviews (SR) of interventions is used to synthesise the effects of interventions from multiple studies to answer healthcare research questions. They can improve precision of effect estimates, investigate factors that can improve or reduce the treatment effect, and allow researchers to answer questions that they might be unable to answer from individual studies. However, these benefits are at risk if the methods of MA are not applied correctly.

Objectives: The aims of this study were 1) to assess the application and interpretation of MA methods in newly published Cochrane Reviews in 2019 to inform Cochrane guidance and training and 2) to compare our findings to similar work carried out in 2014 to see if reporting had improved.

Methods: We evaluated all Cochrane SRs published in 2019. Two authors independently extracted data on primary outcome, statistical methods and methods used to assess reporting bias. We adapted methods employed previously by Page and colleagues comparing Cochrane and non-Cochrane SRs (Page 2018).

Results: We have evaluated 35 SRs (March 2020). SRs included a median of 8 studies (interquartile range (IQR) 4 to 13), and 77% (n = 27) of SRs performed a MA. The median number of included participants was 1591 (IQR 192 to 1329); the median number of MAs performed was 6 (IQR 2.5 to 14). In SRs with a MA, 63% (n = 17) of the MAs of the primary outcome was evaluated with a random-effects model, but only justified by the authors in 35% (n = 6) of MA. A funnel plot was constructed in 19% of SRs with a MA; of these, the recommended minimum number of studies (n = 10) was seen in all funnel plots. The proportion of SRs with an MA and at least one subgroup planned or performed was 74% (n = 20), while the median number of subgroups actually performed 0 (IQR 0 to 2).

Conclusions: In these early results, we found that Cochrane Reviews have largely remained unchanged in their reporting of various methods over the past 5 years. As previously found in 2014, approximately 60% of MAs of primary outcomes used a random-effects model, and 65% of those MAs did not state a rationale for its use. Fewer funnel plots are being produced than in 2014 (19% vs 41%), though a larger proportion of those produced have the recommended minimum number of studies included (100% vs 24%). Development of Cochrane guidance and encouragement for review teams to involve statisticians at the planning and drafting stages should continue and is a focal point of the Cochrane Methods Support Unit.

Patient or healthcare consumer involvement: Not applicable

The application and progress of E-value for sensitivity analysis in observational studies

Fang H[1], Zhang H[1], Liu J[2], Zhang Y[2]

[1] Beijing University of Chinese Medicine, School of Clinical Medicine, China; [2] Center for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine., China

Background: In observational studies, the key limitation is the potential for confounding bias because exposures are not random. In the past, researchers used to control the confounding bias by matching, stratification and regression analysis. But when facing unmeasurable confounding bias, it is found that both statistical method bias analysis and additional assumptions are affected by researchers subjectively, as well as too simplified premise assumptions, which makes the unmeasurable confounding bias exert an inestimable impact on the conclusion.

Objectives: Through systematic investigation, to describe the progress of E-value application in observational research, and explore the advantages and limitations of E-value application in observational research.

Methods: A systematic search was conducted in CNKI and PubMed (from 2018 to 2019), including medical-related observational studies using E-values.

Results: a total of 180 articles were retrieved and 48 research reports were selected. In these reports, 26 cohort studies, 10 case-control studies, 7 cross-sectional studies, 1 clinical randomized trial and 4 meta-analyses were included. It is worth noting that in all reports, for mixed bias, most of them use 2-3 sensitivity analysis methods with E-value in class, such as stratified analysis, tendency score analysis, instrumental variable analysis, multivariate analysis linear regression or logistic regression or Cox regression. However, E-value is often used as a sensitivity analysis method for unmeasurable confounding bias. Researchers try to further enhance the reliability and robustness of the results with E-value. We found that the most commonly used E-values were about cardiovascular and cerebrovascular diseases, up to 13 studies, followed by tumor diseases and drug use evaluation, both of which had 8 reports. Of course, the application of E-value has become a common sensitivity analysis method in recent years. The investigation shows that the maximum value of E is 38.9, which was mentioned by Carl Michael Baraveli et al. In porphyria cutanea tarda increases risk of the patriotic carcinoma and prediction death: a national short study. Among the research exposures involved, disease and drug are the main indicators of exposure, 17 reports mentioned in the former and 12 in the latter. In addition, there are 5 reports using social problems such as psychological problems, policies and regulations as indicators of exposure.

Conclusions: The application of E-value in sensitivity analysis, especially in observational research, due to the influence of many factors, can not completely and accurately evaluate the correlation of confounding factors with the results, but E-value can make the causal inference of observational research more credible.

Patient or healthcare consumer involvement:

The comparison of generalized linear mixed models with the generic two-stage methods for meta-analysis of rare events: a simulation study

Xu C[1], Li L[1], Sun X[1]

[1] Chinese Evidence-Based Medicine Center and Cochrane China Center, West China Hospital, Sichuan University, Chengdu, China

Background: In systematic reviews, handling studies with rare events is a challenging issue. The generalized linear mixed models (GLMMs) based on a one-stage framework may have good performance in dealing with studies with rare events, especially when no events occur in a single trial. In this study, we compared the performance of two GLMMs with the generic two-stage methods.

Methods: We used simulations to generate trials of grouped data with the risk of control group as 0.01. We set the studies for each meta-analysis as 10, and the sample size in both group as a uniform distribution from 15 to 58. We compared the statistical properties of random-intercept GLMM (Method 1) and the random-intercept and random-coefficient GLMM (Method 2) with the fixed-effect inverse variance (Method 3), the random-effect inverse variance (Method 4), the fixed-effect MH (Method 5), the random-effects MH (Method 6), and the Peto method (Method 7). We used the inverse variance method through continuity correction. We used the percentage bias, mean square error (MSE), and coverage probability as performance indicators. We set 25 scenarios and each scenario generated 3000 loops of meta-analyses by simulation. We compared the statistical properties of these methods under different effect sizes (odds ratio (OR) = 1, 2, 3, 4, 5) and heterogeneity (Tau as 0.2, 0.4, 0.6, 0.8. and 1.0 for mild to large heterogeneity).

Results: Our simulation suggested that when the heterogeneity was mild (Tau < = 0.6), all seven methods had low percentage bias, and the two GLMMs (Method 1 and 2) had the lowest bias. The inverse variance methods (Methods 3 and 4) and MH (Methods 5 and 6) had the lowest MSE. The coverage was generally good and similar (> 95%) across these methods, except that the Peto method showed a poor coverage (< 90%) at an odds ratio of 1. In the presence of substantial heterogeneity (Tau > = 0.8), the performance of these methods declined, especially the Peto method. The two GLMMs continued to show lowest bias and good coverage, although inverse variance methods (Method 3 and 4) and MH (Method 5 and 6) had lower MSE.

Conclusions: The generalized linear mixed model may be preferred over the generic two-stage methods when handling studies with no events. Empirical studies are warranted to confirm this finding.

SYNTHESIS OF OTHER STUDY DESIGNS AND DATA

A systematic review of splicing analysis and prediction tools (SAPT) using SWiM (synthesis without meta-analysis) guidance

Jones B[1]

[1] Faculty of Medicine, University of Southampton, UK

Background: Genetic variants affecting splicing play a fundamental role in disease pathogenicity. Prediction of whether a genetic variant will affect splicing is difficult; many in silico tools exist which require adjustment for accurate splice prediction. Best practice guidelines often do not exist and different tools can provide confounding results. New high-throughput next-generation sequencing has increased biological target capture of potential splice sites. Experimental validation is required to characterise any variants in the splice region. The volume of this data, however, is vast; validation is slow, costly and non-viable at scale. Computational tools offer a method to filter results to an actionable quota suitable for experimental follow-up. Prediction of whether variation will affect splicing is challenging; successful tools accelerate diagnosis and aid prioritization of variants of unknown significance with high accuracy and reliability.

Objectives: To determine the effectiveness of eligible splicing analysis and prediction tools (SAPT) and, where possible, rank them alongside providing best practice in their use whilst accounting for quality during the appraisal of eligible tools.

Methods: This study systematically reviewed the literature ranging from 1 January 1980 to 21 October 2019 on SAPT. Statistical measures of specificity, sensitivity and/or accuracy were extracted to provide a hierarchical ranking of tools efficacy and recommendations for best use to aid researchers and clinicians to prioritise experimental follow-up. 'synthesis without meta-analysis' (SWiM) PRISMA-DTA guidance shaped the review framework. Manual Pearl Gathering and PRISMA methods were followed for database searching. The CHARMS checklist provided qualitative assessment rigour. Quantitative analysis of eligible papers weighted SAPT in order preference. Idea Webbing and Triangulation were applied to complete analysis.

Results: Across the subgroups core SAPT: MES, HSF, NNS and SSF-like had high-performance > 85% accuracy. Combination tools emerged with superior performance with four exceeding > 95% accuracy: SPiCE, HSF+SSF-like, HSF+SSF-like+MES, SPIDEX. Established SAPT: dbscSNA, PSSM and CADD alongside SpliceAI reported high performance. Innovative study design within MMS and IntSplice reported adequate performance 70% to 85% accuracy standalone.

Conclusions: Evidence was robust with minimal bias across the studies. Improvements are required in the literature when reporting the delineation of thresholds. Common themes extracted: Effective tools performed best on large curated datasets with separation of candidate predictors, determined in statistical manner without human selection, using both positive and negative datasets. Highly targeted, small window < 100 nucleotide or whole genome, excluding invariant positions, returned results with established veracity. This study successfully developed a hierarchical list of SAPT effectiveness with recommendations on optimal use.

Patient or healthcare consumer involvement: None.

Association between alcohol intake, mild cognitive impairment and progression to dementia: a dose–response meta-analysis

Lao Y[1], Hou L[1], Li J[1], Hui X[1], Yan P[1], Yang K[1]

[1] Lanzhou University, China

Background: Previous study has found a dose–response relationship between alcohol intake and risk of dementia. Mild cognitive impairment (MCI) is a cognitive state falling between normal aging and dementia. However, the relationship between alcohol intake and risk of MCI as well as progression to dementia (PDM) in people with MCI remains unclear.

Objectives: To synthesize available evidence and clarify the relationship between alcohol intake and risk of MCI as well as PDM.

Methods: We searched PubMed, Embase, the Cochrane Library, and China Biology Medicine disc (CBM)) from inception to 1 October 2019. We included only prospective studies that reported at least three levels of alcohol exposure. We used categorical meta-analysis for quantitative synthesis of the relationship between light, moderate and heavy alcohol intake with risk of MCI and PDM. We used restricted cubic spline and fixed-effects dose–response models for dose–response analysis.

Results: We included six prospective cohort studies including 4244 individuals. We observed an unstable linear relationship between alcohol intake (drinks/week) and risk of MCI (P-linear = 0.0396). It suggested that a one-drink increment per week of alcohol intake was associated with an increased risk of 3.8% for MCI (risk ratio (RR) 1.038; 95% confidence interval (CI) 1.002 to 1.075). Heavy alcohol intake (> 14 drinks/week) was associated with high risk of PDM (RR 1.76; 95% CI 1.10 to 2.82). And we found a nonlinear relationship between alcohol intake and risk of PDM. Drinking more than 16 drinks/week (P-nonlinear = 0.0038, hazard ratio (HR) 1.42; 95% CI 1.00 to 2.02), or 27.5 grams/day (P-nonlinear = 0.0047, HR 1.46; 95% CI 1.00 to 2.11) would increase the risk of PDM.

Conclusions: There was a nonlinear dose–response relationship between alcohol intake and risk of PDM. Excessive alcohol intake (> 16 drinks/week, or 27.5 grams/day) was associated with higher risk of PDM. There was an unstable linear dose–response relationship between alcohol intake and risk of MCI.

Patient or healthcare consumer involvement: None.

Examining the effect of nutrition interventions to reduce hyperphosphatemia in chronic kidney disease: is including non-randomized trials a waste of time?

Rozga M[1], Cheng F[1], Moloney L[1], Handu D[1]

[1] Academy of Nutrition and Dietetics, USA

Background: Randomized controlled trials (RCTs) are the gold standard for examining the efficacy of healthcare interventions. However, RCTs of nutrition interventions may not always be feasible due to the long periods of time that nutrition may require to affect health outcomes, lack of generalizability, and lack of funding for nutrition interventions.

Objectives: To examine if there is a difference in the effect size and certainty of evidence from RCTs only versus from RCTs and non-randomized controlled trials (NRCTs) combined in a nutrition-focused intervention. This systematic review included trials examining the efficacy of phosphorus-specific nutrition counseling provided by a registered dietitian nutritionist (RDN), compared to usual care or an alternative intervention, on serum phosphorus levels in dialyzed individuals with chronic kidney disease (CKD).

Methods: We searched the literature using MEDLINE, Embase, CINAHL, Web of Science and other databases from 1 January 2000 to 23 November 2019. We assessed RCTs or NRCTs for risk of bias using the ROB 2.0 and ROBINS-I tools and we determined the certainty of evidence using the GRADE method. We included the results in meta-analysis when possible and compared results between RCTs alone and RCTs in combination with NRCTs.

Results: Eleven RCTs and three NRCTs met the inclusion criteria. Serum phosphorus levels were a primary outcome in all 14 studies. Nutrition interventions resulted in a significant decrease in phosphorus levels in the ten RCTs included in meta-analysis (Mean (95% confidence interval) -0.80 (-1.18 to -0.43) mg/dl; I^2 = 65.5%). In the RCTs, overall risk of bias was high in two studies and there were some concerns in the remaining nine studies, primarily due to randomization and deviations from the intended interventions. Two of the three NRCTs were included in pooled analysis and, in combination with RCTs, results describe a reduction of -0.83 (-1.14 to -0.51) mg/dl (I^2 = 59.7%) in serum phosphorus levels in the intervention groups. In the three NRCTs, two studies resulted in an overall judgement of serious risk of bias and one in moderate risk of bias, primarily due to confounding and measurement of outcomes. Certainty of evidence was not affected whether or not NRCTs were included in the body of evidence; either way, certainty of evidence was "LOW", suggesting phosphorus-focused nutrition therapy from an RDN or equivalent reduces serum phosphorus levels.

Conclusions: When examining the effect of phosphorus-focused nutrition therapy on phosphorus levels in individuals with CKD on dialysis, overall effect size and certainty of evidence was not notably affected by including or excluding NRCTs. Therefore, when RCTs are identified a priori, it may save time and effort to consider RCTs without NRCTs. Further research is needed to determine if inclusion of long-term cohort studies improves understanding of the long-term feasibility and effects of nutrition interventions on patient-centered outcomes in more generalizable populations when assessed in tandem with RCTs.

How low should you go? The impact of including cross-sectional study designs on guideline recommendations

De Brier N[1], Borra V[2], Scheers H[1], Vandekerckhove P[3], De Buck E[4]

[1] Centre for Evidence-Based Practice, Belgian Red Cross, Belgium; [2] Centre for Evidence-Based Practice, Belgian Red Cross; Cochrane First Aid, Belgium; [3] Belgian Red Cross; Department of Public Health and Primary Care, Faculty of Medicine, KU Leuven, Belgium; [4] Centre for Evidence-Based Practice, Belgian Red Cross; Department of Public Health and Primary Care, Faculty of Medicine, KU Leuven; Cochrane First Aid, Belgium

Background: It is not always easy to set selection criteria for study design when conducting a systematic review, especially if only limited evidence is available concerning the topic of interest. An example of such a topic is providing mental health first aid (MHFA) to someone who has experienced a traumatic event.

Objectives: To compare the evidence base when only including experimental studies versus also including observational studies when reviewing the effectiveness of MHFA for assisting people in the aftermath of a traumatic event. The review was performed as part of the development of evidence-based MHFA guidelines of the Belgian Red Cross.

Methods: We conducted systematic literature searches in MEDLINE, PsycINFO and Embase. Previous systematic evaluations (Dieltjens 2014, Fox 2012) on this topic did not identify any evidence when only including controlled study designs. As a result, we also reviewed the body of evidence in studies with a cross-sectional approach. We appraised the certainty of evidence according to the GRADE methodology. The scientific evidence was evaluated by a multidisciplinary panel of mental health experts and patient representatives to formulate evidence-based recommendations.

Results: Out of 1724 articles, we were not able to identify experimental studies fulfilling the predefined selection criteria. Conducting controlled research activities during the aftermath of a traumatic event is difficult to perform since the timing and context are unpredictable and they may hamper the assistance to victims. When study designs of lower quality were screened for eligibility, nine cross-sectional studies were included. This highlights the importance to search for study designs comprising the best available evidence for addressing the research question. However, cross-sectional studies suffer from coexisting methodological limitations. No causal relationships could be inferred from the results since the relevant exposures and outcomes were collected simultaneously. The analyses of associations in the included studies were also subject to selection bias and confounding. Overall, the certainty of the evidence was very low because of the study type, risk of bias and imprecision. Despite these methodological issues, cross-sectional studies currently provide the best possible evidence for developing MHFA guidelines. The expert panel took the limitations of the body of evidence into account in formulating weak evidence-based recommendations.

Conclusions: Evidence-based recommendations can be formulated based on statistical associations inferred from cross-sectional studies when these associations were critically interpreted by the expert panel. When the search was limited to experimental studies, the guideline would only consist of good practice points based on expert opinion.

Patient or healthcare consumer involvement: Involving both mental health experts and patient representatives within an expert panel was critical to develop a MHFA manual based on very low certainty evidence.

Meta-analysis methods used in systematic reviews of interrupted time series studies

Korevaar E[1], Karahalios A[1], Forbes AB[1], Turner SL[1], McDonald S[1], Taljaard M[2], Grimshaw JM[3], Cheng AC[4], Bero L[5], McKenzie JE[1]

[1] School of Public Health and Preventive Medicine, Monash University, Australia; [2] Clinical Epidemiology Program, Ottawa Hospital Research Institute and School of Epidemiology, Public Health and Preventive Medicine, University of Ottawa, Canada; [3] Clinical Epidemiology Program, Ottawa Hospital Research Institute, School of Epidemiology, Public Health and Preventive Medicine, and Department of Medicine, University of Ottawa, Canada; [4] School of Public Health and Preventive Medicine, Monash University and Infection Prevention and Healthcare Epidemiology Unit, Alfred Health, Australia; [5] Faculty of Medicine and Health and Charles Perkins Centre, The University of Sydney, Australia

Background: Many systematic reviews appropriately restrict their inclusion to randomized trials. However, for systematic reviews that aim to synthesise the effects of organizational, policy change, or public health interventions or exposures, non-randomized studies may provide the only available evidence, or important additional evidence to that of randomized trials. The interrupted time series (ITS) design is a type of non-randomized study where a series of measurements are collected at regular intervals before and after an interruption. The period before the interruption can be used to estimate the underlying time trend. This trend can then be projected into the post-interruption period to provide a counterfactual for what would have occurred in the absence of the interruption, allowing the calculation of different effect measures that quantify the immediate and long-term effects of the interruption. Meta-analysis can be used to combine effect estimates across ITS studies. However, the ITS design presents challenges for meta-analysts. To date there have been no reviews examining the approaches and methods used to meta-analyse effect estimates from ITS designs.

Objectives: In this review, we aim to:

1) investigate whether review authors re-analyse primary ITS studies included in reviews, and if so, what re-analysis methods are used;
2) examine the meta-analysis method(s) used;
3) describe the effect measures reported and the completeness of their reporting; and
4) explore the tools and domains that are used to assess the risks of bias and/or methodological quality of the included ITS studies.

Methods: We searched eight electronic databases from a range of disciplines (e.g. public health, economics), between 2000 and 2019 to identify reviews that included a meta-analysis of at least two ITS studies. At least two authors of the review team independently selected studies. From eligible reviews, two authors will independently extract details at the review level: including discipline, and type of interruption; and at the meta-analytic level: type of outcome, effect measure(s), meta-analytic method(s), and any methods used to re-analyse the primary ITS studies. The characteristics of the reviews will be summarized with descriptive statistics.

Results: The search retrieved 4213 citations. After removing duplicates, we screened 2677 titles and abstracts published between 2000 and 2019. We excluded 2346 from the title and abstract screening. Full-text screening of 331 articles yielded approximately 60 reviews that include a meta-analysis of ITS studies. The results will be presented at the Colloquium.

Conclusions: Findings from this review will be used to inform future research examining how different meta-analysis methods for combining results from ITS studies perform, with a view to developing guidance for systematic review authors.

Patient or healthcare consumer involvement: No patients/consumers were involved in the design/reporting of this study.

The benefits of including non-randomized studies in a systematic review of epilepsy surgery

Wilbe Ramsay K[1], Bjellvi J[2], Olsson I[2], Malmgren K[2]

[1] Swedish Agency for Health Technology Assessment and Assessment of Social Services, Sweden; [2] Sahlgrenska Academy, University of Gothenburg, Sweden

Background: Real-world data derived from patient records or databases can be used to investigate treatment effects in fields where randomized controlled trials (RCTs) are difficult to conduct for ethical or other reasons. Here we present an example from an HTA report on epilepsy surgery where real-world data from observational studies were used to strengthen the certainty of evidence.

Objectives: To investigate the effect of resective epilepsy surgery compared to pharmacological treatment only through a systematic review.

Methods: We conducted a systematic review following the PRISMA guidelines. We searched for RCTs and observational studies with a comparison group and assessed them for eligibility and risks of bias. The population was patients with drug-resistant focal epilepsy who had been clinically assessed and selected as candidates for resective epilepsy surgery. The primary outcome was seizure freedom 1 to 2 years after surgery. We performed meta-analyses separately for RCTs and observational studies, but the certainty of evidence was assessed on the basis of all studies.

Results: We included three RCTs and five observational studies after assessment of eligibility and risks of bias. The observational data were derived from patient records, national registries, or local databases. The results from the observational studies supported the results from the RCTs but generated a somewhat larger estimate of the effect. The certainty of evidence was assessed as 'high' for a nonnull effect favoring resective surgery in comparison to pharmacological treatment only (GRADE ++++). Downrating in the imprecision domain was prevented by the larger study population that was provided through the observational studies.

Conclusions: Real-world data strengthened the certainty of evidence and improved the generalizability to clinical practice in this example from the field of epilepsy surgery. Despite challenges relating to controlling for confounding factors, observational studies may constitute an important input to the evidence base in areas where RCTs are scarce.

Patient or healthcare consumer involvement: This systematic review was part of a larger project in which national guidelines were developed for the diagnosis and treatment of epilepsy. Patient representatives were involved at a later stage in the guideline process.

The development and implementation of learning health systems: a comparative case study of personalized medicine models within different health conditions

Gul H[1], Best S[1], Long J[1], Ellis LA[1], Vedovi A[1], Smith J[1], Mahmoud Z[1], Smith K[1], Zurynski Y[1], Rapport F[1], Braithwaite J[1]

[1] Australian Institute of Health Innovation, Macquarie University, Australia

Background: Learning health systems (LHSs) and personalized medicine have fundamental synergy. Both advocate for the same underlying processes: to drive health system and practice improvement. These processes include the use of real-time data for clinical and patient decision making, cycles of iterative learning by communities of practice, and continuous development of the evidence base via a dynamic flow of knowledge between research and practice. We have studied three personalized medicine models within various health conditions. These models are at different stages of development and implementation which has provided us with a unique opportunity to: 1) investigate how potential LHSs emerge and evolve, and 2) identify the different stages of LHS development.

Methods: We collected qualitative interview data on the implementation of three models of precision medicine: intellectual disability (ID) genomics, neurofibromatosis (NF) genetic integrated care, and renal genomics. Participants included: clinical geneticists, genetic counselors, disease specialists, consumer representatives, medical education specialists, and health system management staff (n = 52). 1) Inductive thematic analysis was used to identify factors that give rise to the development of LHSs and then the results interpreted using complex systems theory to explore LHS emergence and evolution. 2) To assess which LHS features were detected within each model, and which features were absent or in development, we developed a qualitative interview coding guide to code for observable features of LHSs as defined by the Institute of Medicine (IoM 2013).

Results: We classified the ID genomics model as being in the exploration stage – here the need for a learning community with access to real-time knowledge was identified. However, the model requires both the design and implementation of a digital data-sharing platform connecting across the community. We assessed the NF integrated care model as being in the preparation stage, where the human and digital infrastructure has been established, and the future focus is on the improvement of implementation and scale-up of the LHS. The renal genomics model was assessed as the most advanced LHS model. Here the human and digital infrastructure has been implemented and scaled to a national network consisting of basic laboratory research teams, multidisciplinary clinical research, which are directly integrated with clinical care. Patient data is being used in real-time to lead clinical decisions. The future focus for the renal genomics model is sustainability and to increase the use of patient experience data. A comparison of the 3 models shows that each LHSs emerged from frontline clinicians who recognized the need for a continuous learning community and accompanying digital infrastructure to support real-time access to knowledge.

Conclusions: At a given time LHSs exhibit varying features and evolve at different rates, however they all begin with the human infrastructure as the foundation.

Patient or healthcare consumer involvement: Consumer representatives included in all 3 studies.

OVERVIEWS AND OTHER TYPES OF EVIDENCE SYNTHESIS

A bibliometric analysis and visualization of knowledge structure of frailty

Zhang W[1], Cao L[1], Li Y[1], Ma B[2], Yang K[2]

[1] Evidence-Based Social Science Research Center, School of Public Health, Lanzhou University, China; [2] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, China

Background: With the development of the aging process, frailty research has drawn an increasing amount of attention over the past decade. However, there is rare research on frailty from the perspectives of bibliometrics and visualization.

Objective: To analyze the knowledge structure and evolution of frailty research, and explore the hotspots and frontiers from the past decade.

Methods: We retrieved data on publications from 2010 to 2019 from the Web of Science Core Collection database. We used CiteSpace V to analyze literature information, including countries/regions, institutions, authors, journals, knowledge base, research hotspots, and research frontiers.

Results: We identified 12,768 publications. There was a steady growth trend for the quantity of publications over the past decade. The United States was the leading country for publications, and the leading institution was the Johns Hopkins University. Kenneth ranked first in the number of papers published. The Journal of the American Geriatrics Society was the most cited journal. Frailty, mortality, older adult, health, and risk were some of the high frequency keywords in co-occurrence analysis of keywords. Burst detection analysis of top keywords showed that frailty model, cohort study, meta-analysis, and cognitive impairment were the new research foci. Co-cited reference cluster analysis revealed the clustered network was divided into 13 clusters. The top three clusters were #0 Whitehall II, #1 aortic valve replacement, and #2 kegalle district.

Conclusion: The development prospects of frailty research could be expected. Researchers should pay more attention to the detection and interventions of frailty. There is a relationship between cognitive impairment and frailty as well as frailty problems of the elderly in low-income and lower middle-income countries.

Patient or healthcare consumer involvement: No

Assessing the impact of climate change on Eastern equine encephalitis virus in Canada

Corrin T[1], Ackford R[1], Mascarenhas M[1], Waddell LA[1]

[1] Public Health Agency of Canada, Canada

Background: Eastern equine encephalitis virus (EEEV) is a mosquito-borne virus that is primarily found in North America and the Caribbean. Over the past decade there has been an increase in virus activity, including large outbreaks in human and horse populations. Predicted climate change is expected to affect the range of mosquitos including vectors of EEEV, which may alter disease risk, posing a public health concern.

Methods: A scoping review (ScR) was prioritized to identify and characterize the global evidence on EEEV. We conducted a thorough search in relevant bibliographic databases and government websites. Two review authors screened titles and abstracts for relevance and the extracted the characteristics of relevant papers using a uniformly implemented data collection form. The study protocol was developed a priori and reporting follows the PRISMA-ScR guidelines.

Results: The ScR included 718 relevant research articles. The majority of the articles originated from North America (97%) between 1933 and 2019. EEEV has been identified in 35 species of mosquitos, over 200 species of birds, various domestic animals, wild mammals, reptiles, and amphibians. Articles identified in this ScR primarily covered three topic areas: epidemiology of hosts and vectors (344 articles) including surveillance (138), pathogenesis of EEEV in hosts (193), and in vitro studies characterizing EEEV (111). Fewer articles evaluated the accuracy of diagnostic tests (63), the efficacy of mitigation strategies (62), transmission dynamics (56), treatment of EEEV in hosts (10), societal knowledge, attitudes, and perceptions (4), and economic burden (2). No predictive models or research on climate change impacts on EEEV were identified.

Conclusions: Despite the lack of direct evidence or predictive models on EEEV, the projected impact of climate change on mosquito populations could result in the spread of EEEV into previously unaffected areas and could change the burden of EEEV in currently affected areas. Public health should be prepared to respond and synthesise research to provide a foundation for evidence-based decision-making.

A scoping review of prospective meta-analyses in health research

Cheyne S[1], Seidler AL[1], Hunter KE[1], Ghersi D[2], Berlin JA[3], Askie LM[1]

[1] National Health and Medical Research Council (NHMRC) Clinical Trials Centre, University of Sydney, Australia; [2] National Health and Medical Research Council (NHMRC), Australia; [3] Johnson and Johnson, US

Background: Prospective meta-analyses (PMA) may reduce many of the issues that can occur in traditional (retrospective) meta-analyses by reducing biases in publication and selective reporting. A summary of PMA literature is needed to gain greater clarity and inform future reporting of PMA.

Objectives: To identify and describe the key features, methods and reporting characteristics of PMA in health research.

Methods: We searched for studies using search terms derived from previously identified PMA, and by consulting topic experts. We systematically searched PubMed, Embase, the Cochrane Database of Systematic Reviews, PROSPERO and performed grey literature searches. One review author screened the search results, with a sample screened by a second review author, to identify any potential PMA or methods articles on PMA. One review author then extracted data, which were checked by a second review author. We used these data to create a survey, which was sent to authors of potential PMA.

Results: Title and abstract screening identified 1109 articles, including two from additional sources. We screened the full text of 289 articles. Of these, 117 articles were identified, including 51 potential and 15 confirmed PMA. We contacted the authors of 51 potential PMA with a survey to determine if they were PMA. We received a 76% response rate to the survey. Of the respondents, 10 were confirmed as not being a PMA, and three did not go ahead due to lack of funding, 12 were confirmed as PMA and another 12 remained uncertain. Of the 27 PMA identified, some reported not having a published protocol and had varied reporting on details such as committees involved, or methods used to identify studies for inclusion. PMA was undertaken in several health areas, with the most prominent being in cancer, cardiovascular disease and reproductive and child health. The identified PMA methods papers were mostly outdated and inconsistent and 63% included only a short section on PMA.

Conclusions: Identification of PMA was difficult. There is insufficient information reported and when details were reported, this was done with much variation. The prevailing uncertainty after contacting authors highlighted the lack of understanding of the definition of PMA. Therefore, a standardized way of reporting PMA is greatly needed. The PMA methods group is currently working on developing such standardized reporting tools.

Patient or healthcare consumer involvement: PMA have the potential to help patients in areas that may not have been possible before. For example, questions can be answered for subgroups and rarer outcomes that may have never before been possible to address.

Addressing harms of screening: a review of outcomes in Cochrane Reviews and suggestions for next steps

Johansson M[1], Borys F[2], Bilamour G[2], Peterson H[2], Bruschettini M[2]

[1] Cochrane Sustainable Healthcare; Cochrane Sweden, Sweden; [2] Cochrane Sweden, Sweden

Background: To facilitate evidence-informed decisions for patients, clinicians and policy makers, systematic reviews of screening should address both their benefits and harms.

Objective: To investigate if Cochrane Reviews and protocols that assess screening interventions address their major harms.

Method: We performed a systematic search for Cochrane Reviews and protocols that assess screening interventions and investigated whether these addressed their major harms. Two authors independently screened titles and abstracts, assessed full-texts, and extracted data from included reviews and protocols. For each review or protocol, two authors judged whether each pre-defined harm was relevant. When the harm was judged as of questionable relevance, the review or protocol was excluded from the denominator in our calculations.

Results: We included 41 reviews and five protocols. Overdiagnosis was addressed in 6 of 39 (15%), overtreatment in 7 of 42 (17%) and psychosocial consequences in 29 of 46 (63%) of reviews and protocols where this was judged relevant. When data on harms were reported, they were generally not treated with the same methodological rigour as the benefits; they lacked assessment of risk of bias or certainty of the evidence. Further, about half of the abstracts, Plain Language Summaries, and 'Summary of findings' tables did not include any harms.

Conclusion: The underreporting of harms of screening interventions in Cochrane Reviews and protocols likely reflect primary research and is problematic from an individual and organizational perspective. We call for broad collaboration to develop reporting guidelines and core outcome sets for primary studies and systematic reviews of screening interventions. To not risk that such guidelines will constitute a bureaucratic barrier to timely publication of up-to-date evidence, editors and authors of Cochrane Reviews should have an active role in this work to make sure that the practical challenges they will face are addressed. Further, we believe it is of great importance that such reporting guidelines are derived through a broad consensus process with representation of the public, clinicians, policy makers, methodologists, medical ethicists, as well as people from outside of medicine (for example social sciences and economics). These should all be free from conflicts of interests in relation to screening interventions.

Patient or healthcare consumer involvement: No patient or healthcare consumer was involved in this work.

Canadian practitioners' opioid-related treatment practices: a scoping review

Harding S[1], Tanzini E[1], Duench S[1], Selby P[1]

[1] Nicotine Dependence Services, Centre for Addiction and Mental Health, Canada

Background: Harms from the use of opioids continue to increase in Canada. While many novel tools and guidelines have been implemented to support safe opioid prescribing, the impact that these tools have on the treatment practices of Canadian practitioners remains unclear.

Objectives: To identify, characterize, and summarize the research related to Canadian practitioners' opioid-related treatment practices and to identify the gaps in the literature.

Methods: We conducted a scoping review using a comprehensive protocol, developed a priori, and a pre-tested search strategy. Two independent review authors screened titles and abstracts for relevance and characterized full-text articles using a pilot tested data screening and extraction form. Any conflicts were resolved by consensus or a third review author, if needed. We exported all data to Excel for cleaning. We performed descriptive analysis of the data using standardized methods.

Results: Preliminary results indicate that 866 relevant articles were captured from the search strategy; half of which (n = 433) were conducted in Ontario. The majority of studies were observational by design (n = 643, 74%), and many (n = 357, 41%) obtained treatment practice data from chart reviews or other existing databases (e.g. Ontario Drug Benefit (ODB) database). The included studies describe a wide range of pharmacological and non-pharmacological treatment practices for patients using opioids or with an opioid use disorder. The most commonly reported treatment setting was a hospital (n = 503, 58%), and opioids were most commonly used for acute pain either before, during, or after surgery (n = 375, 43%). Treatment for opioid use disorder (n = 140, 16%) and chronic pain (n = 119, 14%) were also reasons for using opioids. Opioids were commonly used in combination with other non-opioid medications (n = 542, 63%). Studies reported on the number (n = 181, 21%), dose (n = 447, 52%), and duration (n = 173, 20%) of opioid prescriptions. Many studies however, failed to report the formulation used (n = 492, 57%), type of practitioner involved (n = 478, 55%), or whether opioids were used as first line treatment (n = 559, 65%). Only a small number of studies described practitioner characteristics such as years of training or size of practice (n = 117, 14%), or their perceptions and knowledge (n = 167, 19%) regarding opioid-related treatment. Statistical associations between a treatment practice and a patient characteristic or outcome were reported in less than half of the included studies (n = 347, 40%).

Conclusions: There appears to be a lack of standardized treatment for Canadians who are using opioids or diagnosed with an opioid use disorder. Further research is necessary to elucidate which interventions work, have limited benefit, or lack sufficient evidence. This research will help to inform the continued development of tools which aim to support practitioners' application of evidence-based treatment.

Challenges with integrating quantitative and qualitative evidence through mixed methods systematic reviews

Stern C[1], Lizarondo L[1], Rieger K[2], Godfrey C[3], Apostolo J[4], Carrier J[5], Borges Dos Santos K[6], Kirkpatrick P[7], Loveday H[8]

[1] Joanna Briggs Institute (JBI), University of Adelaide, Australia; [2] University of Manitoba, Canada; [3] Queens University, Canada; [4] Escola Superior de Enfermagem de Coimbra, Portugal; [5] Cardiff University, UK; [6] Universidade de São Paulo, Brazil; [7] Robert Gordon University, UK; [8] University of West London, UK

Background: Mixed methods systematic reviews (MMSR) represent an important development for those involved in evidence synthesis by combining quantitative and qualitative evidence to inform policy and practice. Although MMSR are becoming increasingly popular, there is a degree of complexity and guidance regarding the methodology of combining quantitative and qualitative data is limited.

Objectives: To provide an account of the challenges associated with integrating quantitative and qualitative evidence following the JBI methodology for conducting MMSR.

Methods: We conducted two systematic reviews: one following the JBI convergent integrated approach to synthesis and integration and the other following the convergent segregated approach. We present a descriptive account of the main issues that we encountered during the data synthesis and integration stages of the reviews and recommendations regarding future enhancements to the methodological guidance.

Results: While undertaking synthesis and integration of both reviews a number of challenges arose. Using the integrated approach, issues relating to data transformation (including the qualitization of quantitative data) and integration (including assembling and pooling the data) emerged. In the review following the segregated approach, issues around the available evidence and heterogeneity of included studies impacted on the ability of the two types of research being able to 'speak to each other' whilst keeping their epistemological foundations. This resulted in limitations to the configured analysis of the qualitative and quantitative evidence.

Conclusions: MMSR provide an innovative approach for addressing important questions in health care as they allow for a richer understanding of a topic/phenomena. However, the methods for synthesis and integration in MMSR are complex. Further methodological work is required to shed light to this emerging but rapidly evolving methodology.

Patient or healthcare consumer involvement: Nil involvement associated with the development of this protocol.

Cochrane Rehabilitation collaborates with the World Health Organization to establish a package of rehabilitation interventions based on the best available evidence in stroke

Patrini M[1], Arienti C[1], Lazzarini SG[1], Gimigliano F[2], Kiekens C[3], Negrini S[4]

[1] IRCCS Fondazione Don Carlo Gnocchi, Milan, Italy; [2] Department of Mental and Physical Health and Preventive Medicine, University of Campania "Luigi Vanvitelli", Naples, Italy; [3] Spinal Unit, Montecatone Rehabilitation Institute, Imola, Italy; Department of Physical and Rehabilitation Medicine, University Hospitals Leuven, KU Leuven, Leuven, Belgium; [4] Department of Biomedical, Surgical and Dental Sciences, University of Milan "La Statale", Milan, Italy; IRCCS Istituto Ortopedico Galeazzi, Milan, Italy

Background: The World Health Organization (WHO) sees Universal Health Coverage (UHC) as a strategic priority to meet the United Nations Sustainable Development Goal 3: "Ensure healthy lives and promote well-being for all at all ages". To address UHC in terms of rehabilitation, due to the growing ageing population and the increase in chronic and non-communicable diseases, WHO launched "Rehabilitation 2030". Following the initiative, WHO Department of Non-communicable Diseases started developing a package of rehabilitation interventions (PRI) which provides a set of prioritized evidence-based interventions and resource requirements for delivery to inform countries' healthcare planning.

Objectives: To present the methodology developed to answer to the requests of WHO in the development of PRI in stroke rehabilitation.

Methods: The main health condition studied has been stroke. The literature search used the "tagging process" of Cochrane Rehabilitation as reported by Levack et al (Archives of Physical Medicine and Rehabilitation, 2019). We limited the search to the last 10 years. The reviewers collected for each CSR author, publication year, title, date of search, number of included studies (number of participants), population, setting, intervention, control, outcome, corresponding risk (95% confidence interval (CI)), relative effect (95% CI), quality of evidence (GRADE), statistical method, heterogeneity, upgrade or downgrade motivation. Where not available in the original Cochrane systematic review (CSR), we prepared a 'Summary of findings' table using GRADE evaluation.

Results: Of all the tagged CSRs on the 20 health conditions, we selected the 245 published in the past 10 years. Of these, 158 reported GRADE levels of evidence, 87 did not. After the screening, we analyzed 171 CSRs and prepared 'Summary of findings' tables with GRADE assessments for 46 of the 62 reviews lacking. The remaining 16 either were empty, gave a descriptive summary, focused on secondary outcomes or were overviews. For each health condition, CSRs evidence was summarized in a single spreadsheet. Table 2 reports the extracted data set. The information from CPGs and CSRs is being summarized according to a codification based on the International Classification of Functioning, Disability and Health (ICF).

Conclusions: This work gives a solid basis to the decisional process of the panels in providing the PRI. Evidence from the CSRs helps defining the strength of the recommendations, ruling over conflicting recommendations from different CPGs and revealing areas where primary research is missing.

Patient or healthcare consumer involvement. Not applicable.

Conducting high quality scoping reviews: challenges and solutions

Munn Z[1], Pollock D[1], Peters M[2], Alexander L[3], Khalil H[4], McInerney P[5], Godfrey C[6], Tricco A[7]

[1] Joanna Briggs Institute (JBI), University of Adelaide, Australia; [2] University of South Australia, Australia; [3] Robert Gordon University, UK; [4] School of Psychology and Public Health, La Trobe University, Australia; [5] Wits University, South Africa; [6] Faculty of Health Sciences, Queen's University, Canada; [7] Li Ka Shing Knowledge Institute, St Michael's, Canada

Background: Scoping reviews (ScRs) are a type of knowledge synthesis that use a systematic process to identify and map evidence on a topic and identify main concepts, theories, sources, and knowledge gaps.

Objectives: The objective of this oral is to provide guidance to ScR authors regarding some of the challenges they may face when conducting a ScR and how these can be overcome.

Methods: The JBI Scoping Reviews Methodological Group was established in 2013 to develop guidance regarding the conduct of scoping reviews. This guidance was updated in 2017 and 2020. A survey of scoping review authors and an evaluation of published scoping reviews using JBI methods identified some challenges faced by authors. As such, the group has provided suggested solutions to these challenges.

Results: Challenges identified by the group included misconceptions from editors and peer reviewers, lack of training, difficulty determining when a ScR is appropriate, and issues when presenting results for ScR. Solutions included following methodological guidance and reporting guidelines, establishing formal training programs, ensuring ScRs are protocol driven, and using novel approaches for data visualization and presentation.

Conclusions: ScRs can make an important contribution to science and are a versatile knowledge synthesis approach when conducted and reported appropriately. By overcoming known challenges when conducting ScRs authors can ensure that ScRs are better placed to achieve their aims and objectives.

Patient or healthcare consumer involvement: ScRs are increasingly used by knowledge users including healthcare consumers and other stakeholders to determine the range and breadth of evidence on a topic and establish research and policy priorities.

Coronavirus disease 2019 (COVID-19): a scoping review

Lv M[1], Luo X[1], Liu Y[1], Ren M[1], Zhang X[1], Wang L[1], Wang X[1], Chen Y[2]

[1] School of Public Health, Lanzhou University, Lanzhou, China; [2] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China

Background: In early 2020, COVID-19, a pneumonia caused by a novel coronavirus (SARS-CoV-2) emerged in Wuhan, China, and rapidly spread to 23 countries around the world.

Objectives: To understand the research gaps related to COVID-19 and propose recommendations for future research.

Methods: We conducted a scoping review. We comprehensively searched databases and other sources to identify literature on COVID-19 between 1 December 2019 and 6 February 2020. We analyzed the sources, publication date, study design and research topic of the retrieved studies, and guidelines on COVID-19.

Results: We included 249 articles in this scoping review. More than half of the studies (59.0%) were conducted in China. The most common type was guidelines and consensuses statements (22.6%). Most (192; 77.1%) articles were published in peer-reviewed journals, 35 (14.1%) on preprint servers and 22 (8.8%) posted online. Ten genetic studies (4.0%) focused on the origin of SARS-CoV-2; the topics of molecular studies varied. Nine (40.9%) out of the 22 epidemiological studies focused on the estimation of the basic reproduction number of COVID-19 infection (R0). Among all identified guidelines, only ten (28.6%) were evidence-based, the rest were interim guidelines. The number of articles published per day increased exponentially until the end of January.

Conclusions: The number of articles on COVID-19 steadily increased before 6 February 2020. However, they lack diversity and are almost non-existent in some study fields, such as clinical research. The findings suggest that evidence for the development of clinical practice guidelines and public health policies will be improved when more results from clinical research becomes available.

Patient or healthcare consumer involvement: None

Coronavirus disease 2019 (COVID-19): overview of systematic review

Lu S[1], Luo X[2], Liu X[2], Liu Y[2], Ren M[2], Chen Y[3]

[1] School of Medicine, University of Electronic Science and Technology of China, Chengdu, China; [2] School of Public Health, Lanzhou University, Lanzhou, China; [3] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China

Background: In December 2019, pneumonia caused by severe acute respiratory syndrome coronavirus2 (SARS-CoV-2) broke out in Wuhan, China. The World Health Organization declared COVID-19 to have pandemic characteristics on 11 March. There is some controversy in the academic community about how to deal with COVID-19. Systematic reviews of COVID-19 have provided recommendations and evidence for clinical trials. However, the quality of these studies is uncertain.

Objectives: The aim of our study was to assess the quality of studies published as systematic reviews (SR) or meta-analyses (MA) or rapid reviews (RR) in the field of COVID-19.

Methods: We comprehensively searched electronic databases (MEDLINE, Embase, the Cochrane Database of Systematic Reviews (CDSR), CKIND, Web of Science, WANFANG) for systematic reviews of COVID-19, and two authors independently reviewed all titles and abstracts, assessed the full text of potentially eligible studies and assessed the quality of included studies, resolving any discrepancies by discussion and with help from the third review author. We scored the quality of each SR and MA using the Assessing the Methodological Quality of Systematic Review (AMSTAR 2) checklist and risk of bias in systematic reviews (ROBIS) tool.

Results: Our search retrieved 24 studies. We report the quality of eligible systematic reviews and summarize the results of those with an AMSTAR score ≥ 32. We will continue to monitor the publication of COVID-19 systematic reviews and rapid reviews and present the results at the conference.

Conclusions: We find evidence for drugs treatment and other interventions to deconstruct some controversial clinical issues. We hope to guide the generation of better systematic evaluations in the future.

Patient or healthcare consumer involvement: None.

Development and regular update of a meta-analytic database of randomized controlled trials: the case of psychological treatments for anxiety disorders

Papola D[1], Karyotaki E[2], Sijbrandij M[2], Cuijpers P[2], Barbui C[1]

[1] WHO Collaborating Centre for Research and Training in Mental Health and Service Evaluation, Department of Neuroscience, Biomedicine and Movement Science, Section of Psychiatry, University of Verona, Verona, Italy; [2] Department of Clinical, Neuro and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, The Netherlands

Background: The field of anxiety disorders is characterized by the production of many randomized controlled trials. Most of these studies include small samples of participants suffering from a vast array of anxiety conditions and assess the efficacy of a number of heterogeneous psychological interventions. It becomes particularly interesting, therefore, to develop and maintain a well-organized database of such a continuously growing body of evidence, to be used to plan and conduct meta-analytic re-analyses able to address specific research questions in a timely and comprehensive way.

Objectives: To develop and maintain a database of randomized controlled trials results of psychological treatments for anxiety disorders.

Methods: Four bibliographical databases, PubMed, Embase, PsycINFO, and the Cochrane Register of Controlled Trials (CENTRAL) were firstly searched in April 2019, and last updated in January 2020. The database will be updated every year. We included all randomized trials in which a psychotherapy condition was compared with any other condition. That could be another psychotherapy, pharmacotherapy, a control condition (such as waiting list, and care-as-usual). We also included studies comparing combined treatment of psychotherapy and pharmacotherapy with either of these alone. We are currently collecting data on the participant characteristics, therapies, general characteristics of the studies, and the effect sizes for each of these RCTs. We are using the 'Risk of bias 2' assessment tool to assess the validity of included studies.

Results: To date, the database includes 527 randomized controlled trials, that are categorized as follows: 126 panic disorder and agoraphobia trials; 82 generalized anxiety disorder; 148 social anxiety disorder/social phobia; 106 specific phobias/fears; 65 any anxiety disorders. The publication time span ranges from 1968 to 2020. This database has several purposes. First, it can give other researchers access to the studies we collected to facilitate replications and independent analyses of selections of studies. Second, it will provide background information about already published meta-analyses. Third, as many studies have already been conducted in the field, this database will help researchers planning new trials to focus on gaps in existing knowledge without wasting resources and time.

Conclusions: Psychological treatment for anxiety disorders is one of the best examined fields of psychotherapy for any mental health problem. To date, 527 trials on psychotherapy for anxiety disorders in adults are available. Because of this huge body of knowledge, it is important that the results of these studies are summarized and integrated using meta-analytic techniques, to keep a good, up-to-date overview of this field. It will also save considerable time and effort for researchers who want to conduct a specific meta-analysis. Furthermore, it will guide future research as knowledge gaps will be easily identified.

Patient or healthcare consumer involvement: Not required for this type of project.

Additional files: Flowchart; Figure

Does tobacco smoking have some potential health benefits? An overview of systematic reviews

Wang L[1], Luo X[1], Estill J[2], Zhang X[3], Wang L[3], Li P[3], Lv M[1], Liu Y[1], Ren M[1], Lu S[4], Wang Q[5], Chen Y[6]

[1] School of Public Health, Lanzhou University, Lanzhou, China; [2] Institute of Global Health, University of Geneva, Geneva, Switzerland; [3] The First School of Clinical Medicine, Lanzhou University, Lanzhou, China; [4] Sichuan Provincial People's Hospital, Chengdu, Sichuan, China; [5] McMaster Health Forum, McMaster University, Hamilton, Canada; [6] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou, China

Background: Smoking is a prevalent public health problem worldwide. Most of the research on smoking and human health have focused on the harmful effect of smoking. However, the phenomenon that smoking has a protective effect on some diseases has attracted the attention of some researchers, and many systematic reviews have been carried out.

Objective: To review and analyze the systematic reviews of the protective effect of tobacco smoking.

Methods: We searched PubMed, Embase, the Cochrane Library and Web of Science from the establishment of each database to September 2019 to collect published systematic reviews of the protective effect of smoking on any condition. We used AMSTAR 2 (A MeaSurement Tool to Assess systematic Reviews) to assess methodological quality of included studies, and the GRADE method to assess the quality of evidence of main outcome indicators. We reported the effect sizes of the incidence of the studied condition between smokers and non-smokers with 95% confidence interval (CI) and heterogeneity.

Results: We included thirteen systematic reviews. Five reviews assessed Parkinson's disease: the risk ratio (RR) of the disease comparing smokers with non-smokers across ranged from 0.35 (95% CI 0.26 to 0.47) to 0.80 (95% CI 0.69 to 0.93). Two reviews were about preeclampsia, reporting effect sizes of RR 0.70 (95% CI 0.67 to 0.73 for cohort studies)/odds ratio (OR) 0.68 (95% CI 0.57 to 0.81 for case-control studies) and RR 0.67 (95% CI 0.60 to 0.75). In addition, we found reviews on Alzheimer's disease (OR 0.78, 95% CI 0.62 to 0.98), skin cancer (RR 0.72, 95% CI 0.58 to 0.86 for melanoma in males, RR 0.94, 95% CI 0.90 to 0.98, for basal cell carcinoma in males and females), endometrial cancer (RR 0.81, 95% CI 0.74 to 0.88 for cohort studies, OR 0.72, 95% CI 0.66 to 0.79, for case-control studies), pterygium (RR 0.82, 95% CI 0.69 to 0.97), celiac disease (OR 0.52, 95% CI 0.32 to 0.84), and human acute cognitive response (six indicators, Hedges' g ranging from 0.16, 95% CI 0.02 to 0.31, to 0.44, 95% CI 0.17 to 0.71). Five studies were of low methodological quality and eight were of very low methodological quality. For all 19 main outcome indicators, the quality of evidence was very low.

Conclusion: The available evidence suggests that smoking may have a protective effect against Parkinson's disease, preeclampsia, Alzheimer's disease, skin cancer, endometrial cancer, pterygium, celiac disease and acute cognitive response in humans. However, the overall quality of these studies is relatively low. There is no doubt that smoking does more harm than good. But figuring out the relevant protective mechanisms may be helpful in the treatment of these diseases. The exact mechanisms of the protective analysis are currently poorly understood and need further assessment.

Key words: Smoking; Systematic review; Meta-analysis

Patient or healthcare consumer involvement: None.

Dihydroartemisinin-piperaquine versus sulfadoxine-pyrimethamine for malaria during pregnancy: a systematic review and meta-analysis of randomized controlled trials

Chu X[1], Li M[1], Yan P[2], Feng L[1], Li J[1], Liu X[1], Yang K[1]

[1] Lanzhou University, China; [2] Institute of Clinical Research and Evidence Based Medicine, Gansu Provincial Hospital, China

Background: Malaria in pregnancy is one of the serious global problems of our time. There were wide concerns about dihydroartemisinin-piperaquine versus sulfadoxine-pyrimethamine for prevention of malaria during pregnancy.

Objectives: To assess the current latest evidence on the efficacy and safety of dihydroartemisinin-piperaquine versus sulfadoxine-pyrimethamine for malaria in pregnancy.

Methods: We searched the Cochrane Library, Embase, PubMed and Web of Science from the earliest publication date available to 4 July 2019. We included randomized controlled trials comparing dihydroartemisinin-piperaquine with sulfadoxine-pyrimethamine for malaria in pregnancy. We analysed outcomes using Risk ratios (RR) and 95% confidence intervals (CI). We did subgroup analysis about different intervals, including 4 to 6 weeks or 8 weeks.

Results: We included five studies with 4660 HIV-uninfected pregnant women in area of high malaria-transmission intensity in final synthesis. Meta-analysis showed dihydroartemisinin-piperaquine for intermittent preventive treatment resulted in lower rates of placental malaria (RR = 0.50; 95% CI, 0.43 to 0.59) and infection with malaria parasites at delivery (RR = 0.05; 95% CI, 0.01 to 0.24). In the subgroup analysis, dihydroartemisinin-piperaquine for intermittent preventive treatment at 4 to 6 weeks of administration was associated with a better effect for infection with malaria parasites at delivery.

Conclusions: Dihydroartemisinin-piperaquine was a promising alternative drug to sulfadoxine-pyrimethamine for intermittent preventive treatment in settings with high sulfadoxine-pyrimethamine resistance, especially at 4 to 6 weeks of administration. Based on real-world and other epidemiological settings, more data will be needed to identify the risk of adverse effects.

Patient or healthcare consumer involvement: None.

Diversity of scale and context of implemented learning healthcare systems: a narrative literature review

Zurynski Y[1], Smith K[1], Vedovi A[1], Ellis L[2], Gul H[2], Braithwaite J[1]

[1] Partnership Centre for Health System Sustainability, Australian Institute of Health Innovation, Macquarie University, Australia; [2] Centre for Healthcare Resilience and Implementation Science, Australian Institute of Health Innovation, Macquarie University, Australia

Background: The concept of a learning healthcare system (LHS) has been gaining traction for over a decade. In 2013, the US National Academy of Medicine identified four main areas necessary to establish an LHS: science and informatics, patient-client relationship, incentives, and culture. As healthcare institutions increasingly look to develop learning cultures, improve clinical processes, and exploit the untapped potential of electronic health records in the age of artificial intelligence, a growing literature has emerged on the theory and, crucially, the emerging practical implementation of LHSs.

Objectives: To explore a sample of healthcare systems that self-identified as emerging or implemented LHSs to describe the current field.

Methods: We carried out a narrative/scoping review (January 2016 to May 2019), using search terms ("learning health system" and "learning health care system") in PubMed and Scopus. We identified 276 publications; 79 (28.6%) were excluded as they mentioned LHSs only in passing (e.g. among the keywords, abstract, and/or conclusion). There were 197 remaining publications with an explicit discussion of LHSs, and 67 of these presented case studies or profiles of self-identifying emergent or established LHSs.

Results: The 67 articles described 51 LHS systems: 33 in the United States, 12 in other countries, and six multi-country LHS networks. Twenty-two described data networks or IT architecture platforms at local, national, or international levels to support collection, storage, sharing, and/or analysis of health data; 11 were practice-based research networks (PBRNs) or other 'learning communities', typically focused on particular conditions or disciplines. Five were combinations of a PBRN/learning community and a data network/platform. There were five LHSs operating at the level of single medical centre or clinical unit. We identified eight large-scale systems: four were private corporations and four were public systems at national community or specific population levels.

Conclusions: The LHS concept is a journey rather than a destination, and it has been implemented to some extent in several different contexts and at varying scales. Almost half of the systems centred on data usage and interoperability to provide fundamental infrastructure to underpin LHSs. PBRNs and PBRNs supported by data networks are also emerging; however, large-scale, systemwide LHSs are scarce. The use of the term "LHS" was applied to different systems regardless of their scope or scale. This is not surprising in this emerging and evolving field. Our review included LHSs described in the peer-reviewed literature; the grey literature may reveal additional examples.

Patient or healthcare consumer involvement: As this is a scoping review of the literature, consumers were not directly involved. The Australian Institute of Health Innovation and the Partnership Centre for Health System Sustainability collaborate with the Consumers Health Forum of Australia.

Economic evaluations of interventions for urinary incontinence and mapping the evidence gaps

Haston S[1], Wallace S[2], Sobiesuo P[3], Vale L[4]

[1] c/o Cochrane Incontinence, Health Economics Group, Population Health Sciences Institute, Faculty of Medical Sciences, Newcastle University, UK; [2] Cochrane Incontinence, Evidence Synthesis Group, Population Health Sciences Institute, Faculty of Medical Sciences, Newcastle University, UK; [3] Campbell and Cochrane Economics Methods Group, Health Economics Group, Population Health Sciences Institute, Faculty of Medical Sciences, Newcastle University, UK; [4] Cochrane Incontinence, Health Economics Group, Population Health Sciences Institute, Faculty of Medical Sciences, Newcastle University, UK

Background: Urinary incontinence (UI) is an involuntary loss of urine and it is a common problem, with a reported prevalence between 25% and 45%. Prevalence is higher among women with the risks of UI increasing with age. Though the condition is not life threatening, it is life limiting.

Objectives: The aim is to develop an evidence gap map of economic evaluations of UI interventions to facilitate the incorporation of economic evidence into Cochrane Incontinence reviews. It is unclear how extensive the economic evaluations are; the types of interventions compared in terms of costs and effects, and the context and settings of these studies. An evidence gap map will give a broad picture of the "breadth, depth and methodology" of the economic evaluations of the management of UI.

Methods: We performed systematic searches for full economic evaluations related to UI in NHS EED (inception to 31 December 2014), MEDLINE (1946 to June Week 2 2019) and Embase (1974 to 2019 Week 23) on 14 June 2019. No language restrictions were applied. No restrictions were made with respect to population and type of intervention. Two review authors independently screened studies and any conflicts were resolved. We extracted data on type of: UI; intervention; economic evaluation; outcome measures as well as funding and data sources. A second review author double-checked the data. We constructed a dot plot mapping the four main intervention categories, health outcomes and the funding source.

Results: The literature search produced 2193 unique records. After two rounds of title and abstract screening, first based on our economic evaluation definition (1770 excluded) and then on our UI definition (220 excluded), 203 records remained for full-text screening of which 110 were included. The majority of the studies came from Europe or North America (n = 91) or were unclear (n = 12); none included children, 52 included only women and 5 included only men. The evidence gap map suggests that more economic evaluations are available for drug and surgical interventions (over 50% of these studies were commercially funded) than conservative measures (Figure 1). Only one study compared the costs and effectiveness of the three main types of interventions.

Conclusions: For urinary incontinence, the evidence base on cost-effectiveness upon which Cochrane Reviews can draw is limited. The evidence gap map highlights important gaps for future rigorous studies; it is also a data source, which can streamline searching for economic evidence.

Patient or healthcare consumer involvement: There was no direct patient involvement. However, the identified evidence and gaps could lead to better targeting and uses of limited health resources that will benefit patients. Rigorous economic evaluations based on reliable data sources is crucial in improving both physical and subjective welling of patients.

Additional file: [Figure 1](#)

Efficacy and safety of stem cell transplantation in patients with myocardial infarction: an overview of systematic reviews and meta-analyses

Zhang W[1], Hou L[1], Zhou Y[2], Song W[3], Yang X[4], Chen S[3], Diao M[4], Gao X[3], Liu P[4], Liu X[4], Ma B[5], Yang K[5]

[1] Evidence-Based Social Science Research Center, School of Public Health, Lanzhou University, China; [2] The Second Clinical Medical School, Lanzhou University, China; [3] School of Stomatology, Lanzhou University, China; [4] The First Clinical Medical School, Lanzhou University, China; [5] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, China

Background: Stem cell transplantation has provided a new therapeutic strategy for myocardial infarction (MI). However, its efficacy and safety were still conflicting and uncertain. Moreover, the quality of existing evidence was unknown.

Objectives: To summarize and critically evaluate the quality of the evidence from systematic reviews (SRs) and meta-analyses assessing the efficacy and safety of stem cell therapy for MI.

Methods: We searched Web of Science, PubMed, the Cochrane Library, Embase, China National Knowledge Infrastructure, China Science and Technology Journal Database, Wanfang Database and SinoMed from inception to June 2019. We used the Assessment of Multiple Systematic Reviews (AMSTAR) and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) to assess the methodological and reporting quality of the included reviews, respectively. We used GRADE to rate the quality of the evidence.

Results: We included 37 SRs. The results suggest that stem cell transplantation has beneficial effects for patients with MI, especially in improving left ventricular ejection fraction (LVEF) and left ventricular end-systolic volumes (LVESV). And it is safe without increasing the risk of major adverse clinical events. However, there was inconclusive evidence that stem cell therapy had significant effects on left ventricular end-diastolic volumes (LVEDV) and infarct size. According to AMSTAR, only two of these 37 studies satisfied all the items and the coincidence of "Yes" for each of the 11 items ranged from 5.41% to 100%. Among the 27 items of PRISMA, none of the included SRs fulfilled all 27 items, and the score of reporting quality ranged from 12.5 to 26, with an average of 21.5. Most of the evidence was rated as "moderate quality" or "low quality" by GRADE, mainly because of poor quality of primary trials, publication bias and inconsistency.

Conclusions: Stem cell transplantation generally appears to be effective and safe for patients with MI. However, the reporting and methodological quality of systematic reviews and meta-analyses were suboptimal, which demands improvement. More high-quality evidence is needed to further determine the efficacy and safety of stem cell therapy.

Patient or healthcare consumer involvement: None.

Egg consumption and health outcomes: a global evidence mapping based on an overview of systematic reviews

Zhang X[1], Lv M[2], Luo X[2], Ren M[2], Liu Y[2], Wang L[2], Wang X[2], Chen Y[3]

[1] The First School of Clinical Medicine, Lanzhou University, Lanzhou; Cochrane China Network for GRADE and Guideline Working Group, Lanzhou, China; [2] School of Public Health, Lanzhou University, China; [3] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou; Cochrane China Network for GRADE and Guideline Working Group, Lanzhou, China

Background: The evidence regarding the impact of egg consumption on human health is controversial. An increasing number of systematic reviews and meta-analyses have assessed the association between egg consumption and human health.

Objectives: To combine global evidence-based research to explore the relationship between egg consumption and health outcomes.

Methods: We conducted a systematic literature search to identify systematic reviews and meta-analyses that assessed the association between egg consumption and any type of health outcome. We used AMSTAR (A MeaSurement Tool to Assess systematic Reviews) to evaluate the methodological quality of the reviews, and GRADE to determine the quality of evidence. We visualized the results using a human anatomy diagram and evidence mapping.

Results: Our search revealed 29 systematic reviews and meta-analyses. According to the AMSTAR scores, eight studies were of high quality, fifteen studies of medium quality, and four studies of low quality (Figure 1). We identified 34 primary outcomes from the included 29 reviews, covering a total of 22 different health outcomes. According to our assessment using the GRADE approach, two of the primary outcomes were based on high-quality evidence, 18 on moderate-quality evidence, and 14 on low-quality evidence. Among the 22 health outcomes, egg consumption was found protective against two diseases, and harmful for six health outcomes. No significant association was found for ten outcomes, and the results regarding four outcomes were controversial (Figure 2).

Conclusions: The association between egg consumption and the incidence of cancer, diabetes, cardiovascular diseases and other related diseases has been studied in several meta-analyses. The results were often controversial between studies on the same topic, which can be confusing for making dietary choices.

Patient or healthcare consumer involvement: None.

Additional files: Evidence mapping; Localization

Evaluation of clinical practice guidelines recommendations on the screening and management of frailty

Pan B[1], Wang Y[2], Lai H[2], Ge L[2]

[1] Gansu Provincial Hospital, China; [2] Lanzhou University, China

Background: The population is growing older rapidly throughout the world in recent centuries. Healthy aging has become an extremely challenging issue. Frailty, a geriatric syndrome resulting from the decline of multiple physiological systems, characterized by malnutrition, exercise intolerance, dependence, longer bed rest, lower gait speed, weakness, weight loss, anorexia, hip fracture, risk of falling, delirium, dementia, and staying indoors, has become one of the biggest challenges in facilitating healthy aging. Several organizations have developed clinical practice guidelines (CPGs) to guide patients and healthcare staff to screen and manage frailty in older adults. However, these CPG recommendations may be inconsistent, and their quality is still unknown.

Objectives: To systematically review the consistency of globally available clinical practice guideline recommendations on the screening and management of frailty, and to assess their methodological and reporting quality.

Methods: We searched PubMed, Embase, Web of science, CNKI, WanFang data, Guidelines International Network (GIN), National Institute for Health and Clinical Excellence (NICE), National Guideline Clearinghouse (NGC), Scottish Intercollegiate Guidelines Network, British Geriatrics Society (BGS), and American Geriatrics Society (AGS) to identify related CPGs. We used Rayyan online literature management software to manage the literature search records. Two review authors independently screened the title and abstract of all retrieved bibliographic records according to our eligibility criteria. Any conflict will be resolved by a third review author. The following information was extracted independently: basic characteristic (title, population, development organization or individual, year of publication, country, and funding source), source of evidence (whether it was systematically retrieved evidence, databases that it was searched, whether report the detail search strategy, beginning and end years of search, method of evidence grading), detail recommendation information (grading systems, recommendation for frailty screen, recommendation for frailty assessment, recommendation for frailty management, the basis of recommendations forming, strength of recommendation). Four researchers independently assessed their methodological and reporting quality using the AGREE II (Appraisal of Guidelines for Research and Evaluation) instrument and the RIGHT (Essential Reporting Items for Practice Guidelines in Healthcare) checklist.

Results: This study is ongoing and results will be presented at Colloquium as available.

Conclusions: This study is ongoing and results will be presented at Colloquium as available.

Patient or healthcare consumer involvement: Not Applicable.

Evidence mapping: interventions to lose weight for obese population

Chen N[1], Li X[1], Yang K[1]

[1] Evidence-Based Social Science Center, Lanzhou University, Lanzhou, China

Background: Interventions have been proposed for weight loss, but it is difficult to understand the breadth and depth of evidence, as various obese conditions may respond differently to therapy.

Objectives: To provide a visual overview of the evidence distribution for obesity therapy, as well as an accompanying narrative that will help stakeholders interpret the state of evidence to inform policy and clinical decision-making.

Methods: We searched PubMed, Embase, and the Cochrane Library for systematic reviews reporting outcomes for interventions. We assessed the quality of each review using the Assessing the Methodological Quality of Systematic Reviews (AMSTAR) criteria. We used a bubble plot to depict the number of included articles, the effect of interventions for obesity, and the strength of findings for each included systematic review.

Results: Data extraction is in progress.

Conclusions: Prior reviews have conclusions of low strength of evidence because few primary studies of large samples with rigorous methods have been conducted, leaving evidence gaps about specific intervention types for obesity. Primary studies often do not provide adequate details of the intervention provided, limiting the extent to which reviews can draw conclusions about characteristics such as provider type.

Patient or healthcare consumer involvement: None.

Guidance for overviews of reviews continues to accumulate, but important challenges remain: A scoping review update

Gates M[1], Gates A[1], Guitard S[1], Pollock M[2], Hartling L[1]

[1] Alberta Research Centre for Health Evidence, Department of Pediatrics, University of Alberta, Canada; [2]Institute of Health Economics, Canada

Background: Overviews of reviews (overviews) provide an invaluable resource for clinical decision-makers by combining large volumes of systematic review (SR) data into a single synthesis. The production of high-quality overviews hinges on the availability of practical evidence-based guidance for conduct and reporting.

Objectives: Within the broad purpose of informing the development of a reporting guideline for overviews, we aimed to provide an up-to-date map of existing guidance related to the conduct of overviews, and to identify common challenges that authors face when undertaking overviews.

Methods: We updated our previous scoping review (published 2016) using the search methods that had produced the highest yield: ongoing reference tracking (retrospectively from 2014 to present in PubMed, Scopus, and Google Scholar), handsearching conference proceedings and websites, and contacting authors of published overviews. Using a qualitative meta-summary approach, one review author extracted, organized, and summarized the guidance and challenges presented within the included documents. A second review author verified the data and synthesis.

Results: We located 22 new guidance documents, for a total 74 documents produced by 30 research groups. The newly available guidance helps to resolve some existing challenges in the production of overviews. Important developments include the availability of a decision tool for selecting SRs for inclusion in overviews (e.g. in the event of overlapping and/or discordant SRs) and strengthened guidance on handling primary study overlap at the analysis stage. Despite marked progress, several areas continue to be hampered by inconsistent or lacking guidance. For example, there is ongoing debate about whether, when, and how supplemental primary studies should be included in overviews. Consensus is lacking on the preferred tool for assessing risk of bias or methodological quality of included SRs, and how these tools might best be applied in overviews. Guidance remains scant on how to extract and use appraisals of quality of the primary studies within the included SRs and how to adapt GRADE methodology to overviews. The challenges that overview authors face are often related to the above-described steps in the process where evidence-based guidance is lacking or conflicting.

Conclusion: The rising popularity of overviews has been accompanied by a steady accumulation of new, and sometimes conflicting, guidance. While recent guidance has helped to address some of the challenges that overview authors face, areas of uncertainty remain. These findings are being used to inform the development of a reporting guideline for overviews, which aims to support the high quality and clarity of reporting that is needed to substantiate overviews as a robust source of evidence for healthcare decision making.

Patient or consumer involvement: Patients and consumers were not directly involved, but are expected to benefit from the improved conduct and reporting of overviews that is supported by this scoping review.

Keeping up and keeping pace with the evidence, COVID-19 lessons learned

Waddell L[1]

[1] Public Health Agency of Canada, Canada

Background: Since the world was alerted to the SARS-COV-2 outbreak in Wuhan China there has been a scramble to monitor and publish on the epidemic, predict its spread, study the virus, transmission patterns, clinical presentation and disease. The challenge for public health organizations is keeping up with the rapidly evolving literature.

Objectives: To identify, categorize and briefly summarize the literature on SARS-COV-2 / COVID-19 on a 24-hour cycle to facilitate accurate and up-to-date information flow to decision-makers.

Methods: We designed a protocol for the systematic identification, classification, dissemination and maintenance of information on SARS-COV-2 to run on a 24-hour cycle, 5 days per week. We initiated a daily literature search and pulled new citations from 10 bibliographic databases, preprint databases and coronavirus dashboards set-up by prominent publishers. We added new citations to a citation management software, and duplicates were removed or updated as citations moved from preprint to accepted peer-reviewed publications. We assigned reviewers to read, classify and provide a brief synopsis of key outcomes in each manuscript, which were compiled into a daily report. We compiled daily reports into a filterable and searchable running list of daily reports. Read/export access to the citation database was also made available to end-users for personal referencing or exploration.

Results: Seven people were trained to work on this project part-time to facilitate quick daily dissemination of emerging literature. Standard topic categories were developed and used consistently throughout the project and each citation was tagged into one or more categories. Each day there were 20 to 150 new citations to review and summarize. By 3 March 2020 there were over 900 citations on SARS-COV-2 captured through this process. The topic areas with the most literature include clinical data (n = 203), epidemiology (177), predictive models (160) and coronavirology (116). There was also a substantial number of articles on transmission (n = 51), diagnostics (73), therapeutics (64), and vaccines (13). To manage on-going information needs, we developed data extraction tools for key epidemiological and clinical parameters, and topic specialists maintained up-to-date summaries on public health intervention research, healthcare intervention research, vaccines and therapeutics. Other outcomes were summarized on an as needed basis. We developed additional topic-based summaries from the SARS-COV-2 literature as required.

Conclusions: Linking together synthesis expertise with varying backgrounds across our organization prevented duplication of efforts. This improved the speed and efficiency of accurate responses to decision-makers and lead to persons with appropriate expertise responding to inquiries. Our experience highlights the use of synthesis research principles during an epidemic to support rapidly evolving evidence-based decision-making in a timely manner.

Patient or healthcare consumer involvement: None

Masks for preventing COVID-19 infection: a scoping review

Yang N[1], Feng X[2], Zhang J[3], Wang J[3], Ma Y[1], Wang M[4], Chen Y[1]

[1] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Cochrane China Network; Chinese GRADE Center, China; [2] Chengdu Medical College, China; [3] School of Public Health, Lanzhou University; Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Cochrane China Network; Chinese GRADE Center, China; [4] Department of Radiology, the First Hospital of Lanzhou University, China

Background: Masks play an important role in preventing infection with high risk at COVID-19 pandemic. However, the recommendations of masks are conflicting and evidence on effectiveness against COVID-19 are limited.

Objectives: To find the research gaps related to masks for preventing COVID-19 infection and propose recommendations for future research.

Methods: Considering of the few search results on COVID-19 based on strategy preliminary search, the scoping review will include any types studies on masks for Severe Acute Respiratory Syndrome (SARS), Middle East Respiratory Syndrome (MERS) and influenza. The search strategy includes the following terms "mask", "respirator", "COVID-19", "novel coronavirus", "2019-novel coronavirus", "Novel CoV", "SARS-CoV-2", "2019-CoV", "Middle East Respiratory Syndrome", "MERS", "Severe Acute Respiratory Syndrome", "SARS". Two review authors will search independently in the following electronic databases: the Cochrane Library, MEDLINE (via PubMed), Embase, Web of Science, China Biology Medicine disc (CBM), China National Knowledge Infrastructure (CNKI), and Wanfang Data. We searched all databases from their inception until 31 April 2020. Two review authors will also search the following websites for relevant publications: World Health Organization (WHO), the National Health Commission of the People's Republic of China, Google Scholar, BioRxiv, SSRN, and MedRxiv. In addition, we will scan published online articles on COVID-19 in selected major medical journals (Journal of the American Medical Association, The Lancet, New England Journal of Medicine and their sister journals). Two researchers will independently screen the records and extracted data, disagreements will be resolved through discussion or by a third party.

Results and conclusions: This study is ongoing and results will be presented at the Cochrane Colloquium as available.

Meta-research evaluation of clinical trials of breastmilk substitutes

Helfer B[1], Bero L[2], Boyle RJ[1]

[1] Department of Medicine, Imperial College London, UK; [2] Charles Perkins Centre, The University of Sydney, Australia

Background: Breastmilk substitutes (BMS, also known as infant formula), are widely consumed by infants at an early stage of development and in very high volumes. Clinical trials of BMS products evaluate their safety and influence on health and are essential for making decisions about infant nutrition. BMS trials form a core part of a global industry, but seem to be sensitive to commercial pressures, have a light level of regulation, and are thought to have issues with specific forms of bias. Meta-research investigates research practices to understand how to reduce biases in science and thereby deliver more reliable conclusions from scientific studies.

Objectives: A meta-research evaluation of recently published BMS trials to establish whether trial data were correctly and completely analysed, reported and interpreted and to describe whether financial conflicts of interest of study authors influenced study results.

Methods: We will include peer-reviewed publications of clinical trials with at least one BMS head-to-head comparison published in the last 10 years. A stepwise data-reduction strategy will identify the main trial publication and the main trial outcome for each trial. Analysis will include evaluation of biases (Cochrane Risk of Bias 2 tool), risk of undermining breastfeeding within the trial, registration and reporting practices, ethical standards, conflicts of interest, spin as well as ethical and marketing standards.

Results: Recently, several groups of academics and regulators identified flaws in published BMS trials. Specific issues included biases related to attrition and selective outcome reporting, a lack of independent trial oversight and unclear and variable regulatory requirements compared with drug trials. In some trials, study procedures appeared to directly contravene the International Code of Marketing of Breast-milk Substitutes, for example by providing free BMS products to trial participants. This meta-research evaluation will provide detailed evidence and evaluation of these issues.

Conclusions: This meta-research evaluation will generate knowledge on how to better design, conduct, report and regulate BMS trials, leading to better decisions regarding infant nutrition.

Patient or healthcare consumer involvement: We consulted former BMS research participants, who were surprised at the possibility that trial results may be subject to bias or commercial interests and offered insightful solutions on how to disincentivise switching to formula feeding within BMS trials.

Methodological and reporting quality of network meta-analyses in the field of nutrition

Zhang Q[1], Hou L[2], Guo Z[3], Chen M[3], Zhan Q[3], Ge L[2]

[1] School of Nursing, Lanzhou University, China; [2] Evidence-Based Social Science Research Center, School of Public Health, Lanzhou University, Lanzhou, China; [3] Department of ICU, Second Hospital of Lanzhou University, Lanzhou, China

Background: There have been few recent reports on the methodological quality of network meta-analysis, despite the enormous number of studies using network meta-analyses in the field of nutrition. The purpose of this study was to evaluate the quality of network meta-analyses about nutrition according to the Assessment of Multiple Systematic Reviews (AMSTAR) and Preferred Reporting Items for Systematic Reviews and Network Meta-Analyses (PRISMA-NMA) guidelines. **Objectives:** To evaluate the quality of conduct and reporting of published network meta-analyses in nutrition.

Methods: We searched PubMed, Embase, and the Cochrane Library from inception to March 2020 to identify all nutritional network meta-analyses ever published. Two review authors independently screened studies, extracted data, and assessed the methodological quality and reporting quality of included network meta-analysis applying the AMSTAR and PRISMA checklists. We also explore the factors that could influence the quality of network meta-analysis with meta-regression method. We will use STATA 15.1 software for all statistical analysis .

Results: This study is ongoing and results will be presented at the Cochrane Colloquium as available.

Conclusions: This study is ongoing, and results will be presented at the Cochrane Colloquium as available.

Patient or healthcare consumer involvement: Not applicable.

Methodological quality of systematic reviews on treatments for osteoporosis: a cross-sectional study

Tsoi AKN[1], Ho LTF[2], Wu IYX[3], Wong CHL[1], Ho RST[1], Lim JYY[1], Mao C[4], Lee EKP[1], Chung VCH[1]

[1] The Jockey Club School of Public Health and Primary Care, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong; [2] School of Chinese Medicine, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong; [3] Xiangya School of Public Health, Central South University, China; [4] Department of Epidemiology, School of Public Health, Southern Medical University, China

Purpose: Systematic reviews (SRs) provide the best evidence on the effectiveness of treatment strategies for osteoporosis. Carefully conducted SRs provide high-quality evidence for supporting decision-making, but the trustworthiness of conclusions can be hampered by limitation in methodological rigour. We aimed to appraise the methodological quality of a representative sample of SRs on osteoporosis treatments in a cross-sectional study.

Methods: We searched the Cochrane Database of Systematic Reviews, Embase, MEDLINE, and PsycINFO for SRs on osteoporotic treatments. AMSTAR 2 (A MeaSurement Tool to Assess systematic Reviews) was used to evaluate the methodological quality of SRs. We explored associations between bibliographical characteristics and methodological quality ratings using multivariate regression analyses.

Results: We appraised 101 SRs. Overall, one (1.0%) was rated "high quality", three (3.0%) were rated "moderate quality", 11 (10.9%) were rated "low quality", and 86 (85.1%) were rated "critically low quality" (Figure 1). Ninety-nine (98.0%) did not explain study design selection, 85 (84.2%) did not provide a list of excluded studies (84.2%), and 85 (84.2%) did not report funding sources of included studies (Table 1). SRs published in 2018 or after were associated with higher overall quality (adjusted odds ratio (AOR): 5.48; 95% confidence interval (CI) 1.12 to 26.89), while SRs focused on pharmacological interventions were associated with lower overall quality (AOR: 0.24; 95% CI 0.06 to 0.96) (Table 2).

Conclusion: The methodological quality of the included SRs is far from satisfactory. Future review authors must strengthen the methodological rigour of SRs by improving literature search comprehensiveness, registering and publishing a priori protocols, and optimizing study selection and data extraction. Better transparency in reporting conflicts of interest among review authors, as well as sources of funding among included primary studies, is also needed.

Additional file: [Figure and Tables](#)

Physical activity/exercise for health outcomes: an overview of Cochrane systematic reviews

Posadzki P[1], Pieper D[2], Bajpai R[3], Makaruk H[4], Könsgen N[5], Lena A[5], Semwal M[6]

[1] Kleijnen Systematic Reviews Ltd, UK; [2] Witten/Herdecke University, Germany; [3] Keele University, UK; [4] Józef Pilsudski Academy of Physical Education, Poland; [5] Witten/Herdecke University, Germany; [6] Nanyang Technological University, Singapore

Background: Physical activity/exercise is increasingly being recommended and offered in various healthcare systems and for a variety of health conditions/outcomes.

Objectives: To:

- find all the available evidence from Cochrane systematic reviews (CSRs) of the effectiveness of physical activity/ exercise for various health outcomes;
- evaluate the strength and quality of the existing evidence; and
- create recommendations for future researchers, patients, clinicians.

Methods: We adhered to guidance in the Cochrane Handbook for Systematic Reviews of Interventions and Preferred Reporting Items for Overviews of Reviews while writing and reporting this overview. We included Cochrane systematic reviews of randomized controlled trials (RCTs) involving both healthy individuals and medically compromised patients of any age and gender. We included only reviews assessing physical activity/exercise as a stand-alone intervention. Reviews evaluating any type of health-related outcome measures; and any types of controls were deemed eligible. We excluded complex interventions assessing exercise or physical activity and diet, lifestyle/behavioural changes. The methodological quality of the CSRs was independently evaluated by two review authors using the AMSTAR 2 tool (A MeaSurement Tool to Assess systematic Reviews).

Results: A total of 150 CSRs met the inclusion criteria. There were 54 different conditions. Most CSRs were of high methodological quality. One hundred and thirty CSRs employed meta-analytic techniques and 20 did not. Limitations for studies were the most common reasons for downgrading the quality of the evidence. Based on 10 CSRs and 187 RCTs with 27,671 participants, there was a 13% reduction in mortality rates risk ratio (RR) 0.87, 95% confidence interval (CI) 0.78 to 0.96); I^2 = 26.6%, prediction interval (PI) 0.70 to 1.07), median effect size (MES) 0.93 [interquartile range (IQR) 0.81 to 1.00]. Data from 15 CSRs and 408 RCTs with 32,984 participants showed a small improvement in quality of life (QOL) standardized mean difference (SMD) 0.18, 95% CI 0.08 to 0.28; I^2 = 74.3%; PI -0.18 to 0.53], MES 0.20 [IQR 0.07 to 0.39]. Subgroup analyses by the type of condition showed that the magnitude of effect size was the largest among patients with mental health conditions.

Conclusions: There is a plethora of CSRs evaluating the effectiveness of physical activity/exercise. The evidence suggests that physical activity/exercise reduces mortality rates and improves QOL with minimal or no safety concerns.

Patient or healthcare consumer involvement: None.

Probiotics in the overweight or obese pregnant women: systematic review and meta-analysis of randomized controlled trial

Chu X[1], Zhang N[1], Feng L[1], Yan P[2], Yang K[1]

[1] Lanzhou University, China; [2] Institute of Clinical Research and Evidence-Based Medicine, Gansu Provincial Hospital, China

Background: Gestational diabetes mellitus (GDM) is associated with increased risks of complications during pregnancy and delivery. Overweight and obese pregnant women are at higher risk of GDM. The approach of probiotics to prevent GDM among overweight and obese pregnant women has not reach a unified conclusion.

Objectives: To assess the current latest evidence on probiotics for overweight and obese pregnant women as comprehensively as possible.

Methods: We searched the Cochrane Library, Embase, PubMed and Web of Science from the earliest publication date available to 23 September 23 2019. We included randomized controlled trials comparing probiotics with other interventions in overweight and obese pregnant women. We also screened the reference lists of relevant reviews and meta-analysis. No language restrictions were applied. Two authors independently included studies, extracted data, and assessed risk of bias. We used the weighted mean difference (WMD) as the effect size for the continuous variables. We presented summary risk ratios (RR) with 95% confidence intervals (CI) if the results were binary variables.

Results: We included five studies with 1048 overweight or obese pregnant women in this meta-analysis: 519 in probiotics intervention groups, 529 in placebo control groups. Meta-analysis showed the occurrence of GDM did not differ significantly between the probiotics and control groups (RR = 1.03; 95% CI 0.81 to 1.30). On the contrary, the fixed effects analysis showed a significant reduction on the incidence of preeclampsia in control groups (RR = 1.91; 95% CI 1.03 to 3.55).

Conclusions: Probiotics may not be a promising diet supplementation to prevent GDM. More data about the effects of probiotics supplementation on primary, maternal and birth outcomes are needed.

Patient or healthcare consumer involvement: No

Progress toward a reporting guideline for overviews of reviews of healthcare interventions: Preferred Reporting Items for Overviews of Reviews (PRIOR)

Gates M[1], Gates A[1], Pollock M[2], Fernandes RM[3], Pieper D[4], Tricco AC[5], Hartling L[1]

[1] Alberta Research Centre for Health Evidence, Department of Pediatrics, University of Alberta, Canada; [2] Institute of Health Economics, Canada; [3] Clinical Pharmacology Unit, Instituto de Medicina Molecular, University of Lisbon; Department of Pediatrics, Santa Maria Hospital, Portugal; [4] Institut für Forschung in der Operativen Medizin, Department für Humanmedizin, Universität Witten/Herdecke, Germany; [5] Li Ka Shing Knowledge Institute of St. Michael's Hospital; Dalla Lana School of Public Health and Institute of Health Policy, Management and Evaluation, University of Toronto, Canada

Background: Overviews of reviews of healthcare interventions (overviews) compile information from multiple systematic reviews to provide a single synthesis of relevant evidence. There are currently no systematically developed reporting guidelines for overviews. As a result, published overviews are often incomplete and lacking in transparency.

Objectives: We are using explicit, systematic, and transparent methods to develop an evidence-based and agreement-based reporting guideline for overviews of healthcare interventions: Preferred Reporting Items for Overviews of Reviews (PRIOR).

Methods: We are developing the PRIOR reporting guideline in four stages, using established methods. We are developing an original stand-alone guideline that will allow us to focus on the challenges that are unique to the reporting of overviews and facilitate future guideline extensions. First, in December 2018 we launched the project and established an international and multidisciplinary expert advisory board who have been overseeing the conduct of the project and providing methodological support. Second, we conducted comprehensive literature reviews (scoping review of methods guidance for overviews; descriptive review of the reporting characteristics of published overviews) that were used to inform a list of prospective checklist items. Third, we are currently using a 3-round modified Delphi exercise to achieve a high level of expert agreement on the list of items to be included in the PRIOR reporting guideline. We aimed to recruit 100 international experts (editors, authors, peer reviewers, and end-users of published overviews such as guideline developers, policymakers, patients and consumers) to participate in the Delphi process. The first two rounds occurred via online survey. The third round will occur during a smaller (~8 to 10 participants) in-person meeting following the 2020 Cochrane Colloquium. Fourth, after reaching agreement on the included items, we will produce and disseminate the PRIOR reporting guideline.

Discussion: This poster will present preliminary findings from the first two Delphi rounds, comment on progress, and build awareness about the forthcoming PRIOR reporting guideline for overviews. By holding overviews to a minimum standard of reporting, we expect PRIOR to enhance the translation of otherwise overwhelming volumes of literature into accurate, complete, and transparent syntheses for use by healthcare decision-makers.

Patient or consumer involvement: Patients and consumers will be involved in the development of PRIOR as participants on the Delphi panel.

Quality of guideline recommendations about physical activity and exercise to reduce cardiometabolic risk factors: a systematic review and critical appraisal

Siedler M[1], Lamadrid P[1], Humphries M[1], Mustafa RA[2], Falck-Ytter Y[3], Sultan S[4], Dahm P[4], Murad MH[5], Morgan RL[6]

[1] University of South Florida, USA; [2] University of Kansas Medical Center, USA; [3] Case Western Reserve University, USA; [4] University of Minnesota, USA; [5] Mayo Clinic, USA; [6] McMaster University, Canada

Background: Physical activity can reduce the risk of a number of diseases including heart disease, diabetes, and stroke, which are among the top ten leading causes of death worldwide. Numerous guidelines for the prescription of physical activity are published each year, but the quality and practicability of these guidelines is currently unknown.

Objectives: This systematic review and critical appraisal of physical activity guidelines summarizes the current quality of these guidelines and provides suggestions to improve their development.

Methods: We searched Ovid MEDLINE, Embase, Scopus, and the Cochrane Central Register of Controlled Trials (CENTRAL) from January 2000 through October 2019 for physical activity guidelines that included recommendations about physical activity for the prevention of cardiometabolic disease. We systematically reviewed and critically appraised 95 guidance documents using three tools: AGREE II, the National Academy of Medicine (NAM) Standards for Trustworthy Clinical Practice Guidelines, and the Frequency, Intensity, Time and Type (FITT) score.

Results: For the average guideline, AGREE II domain scores ranged from 38% to 84% (median: 47%) while the portion of criteria fulfilled within each NAM domain ranged from 7% to 39% (median: 28%). Of the AGREE II domains, guidelines scored highest on Scope and Purpose and Clarity of Presentation and lowest on Applicability and Editorial Independence (Figure 1). Of the NAM criteria, guidelines scored highest on Articulation of Recommendations and lowest on External Review (Figure 2). The average FITT score for all recommendations was 2.48 of 4. Recommendations for aerobic activity were more likely to have a higher FITT score (3.01) than recommendations for resistance training (2.21), balance/flexibility (1.89), unstructured play (2.4), or activities of unspecified modality (1.94). While guidelines improved according to both AGREE II and NAM standards over time, their FITT scores did not improve. Guidelines produced by governmental or other non-profit agencies or using the GRADE approach were of higher quality.

Conclusions: Organizations producing guidelines about physical activity should improve processes for public representation, external review, and conflict of interest management. Future recommendations about physical activity should also be more specific and include strategies for implementation.

Patient or healthcare consumer involvement: No

Additional files: Figure 1; Figure 2

Research trends and hotspots analysis of multiple sclerosis diagnosis from 2010 to 2019: a bibliometric analysis

Cao L[1], Li X[2], Li M[1], Li Y[1], Yang K[1]

[1] School of Public Health, Lanzhou University, China; [2] School of Public Health, Evidence-Based Social Science Research Center, Lanzhou University, China

Background: Multiple sclerosis (MS) is a chronic immune-mediated disease of the central nervous system (CNS) that can cause substantial disability. It causes a major socioeconomic burden, both for the individual and for society. It can seriously impact on people's health-related quality of life (HR-QOL) and is associated with high economic costs for patients, their families and society.

Objectives: This study aimed to comprehensively analyze the global scientific output of multiple sclerosis diagnosis research and explore the hotspots and frontiers from 2010 to 2019.

Methods: We downloaded publications data from the Web of Science Core Collection database .We used Microsoft Excel 2016 to detect the trend of annual numbers of publications, and used Citespace V and VOSviewer 1.6.8 software as the bibliometric method to analyze the authors, journals, countries, institutions, keywords, citation reports, research hotspots, and research frontiers.

Results: Until 30 December 2019, a total of 19,042 papers in multiple sclerosis diagnosis research were identified as published between 2010 and 2019. We identified that the number of publications on multiple sclerosis diagnosis is increasing over time. The USA was the leading country for publications, and the leading institution was the Mayo Clinic. Co-cited reference analysis revealed the top landmark articles in the field. Multiple sclerosis, diagnosis, biomarker, identification, and disease are some of the high frequency keywords in co-occurrence cluster analysis and co-cited reference cluster analysis, indicating that biomarker and related diseases remain the hotspots in multiple sclerosis research.

Conclusions: This study revealed that our understanding of the link between multiple sclerosis and associated diagnosis has evolved dramatically over time. The emerging new diagnosis methods in multiple sclerosis would be the focus of future research.

Patient or healthcare consumer involvement: Not applicable.

Scoping review on methodological approaches for clinical trials in "personalized medicine"

Gerardi C[1], Banzi R[1], Hulstaert F[2], San Miguel L[2], Imaz I[3], Garcia P[4], Demotes-Mainard J[4], Superchi C[5], Porcher R[5]

[1] Istituto di Ricerche Farmacologiche "Mario Negri" IRCCS, Milan, Italy; [2] Belgian Health Care Knowledge Centre (KCE), Brussels, Belgium; [3] Health Technology Assessment Agency (AETS), Carlos III Institute for Health, Madrid, Spain; [4] European Clinical Research Infrastructure Network (ECRIN), Paris, France; [5] Centre of Research in Epidemiology and Statistics, Université de Paris (CRESS-UMR1153), Paris, France

Background: "Personalized, precision, stratified medicine" is understood as a medical approach in which patients are stratified based on their disease subtype, risk, prognosis, or treatment response, to base medical decisions on individual patient characteristics. The concept of personalized medicine (PM) impacts how treatments are discovered and developed. The current optimism (often hype) on PM is affecting how patients are diagnosed and treated, how healthcare systems allocate their resources and research funding. The European Commission launched the International Consortium for Personalized Medicine and a series of supporting research projects, like the Personalized Medicine Trials (PERMIT https://permit-eu.org/) aimed to investigate various aspects of the methodology used in PM research. Rigorous methodologies applied to clinical trials in PM are critical to correctly select participants and treatments to be tested. However, there are several open questions on the advantages and challenges of different innovative and complex trial designs.

Objectives: To map the methods for clinical trials in PM and identify standards and gaps in methods.

Methods: Scoping reviews have great utility for synthesizing research evidence and mapping existing literature about the nature, features, and volume of a given field. We are conducting a scoping review following the guidance suggested by the Joanna Briggs Institute on the theme of clinical trials in PM. Briefly, we firstly defined the scope and research questions. We will retrieve relevant articles and reports through rounds of formal literature searches on relevant databases (i.e. MEDLINE, Embase, the Cochrane Library) and informal checks of reference lists and institutional websites. We will include methodological reports and guidance issued by regulatory authorities or agencies for health technology assessment and examples of published or ongoing trials in personalized medicine. The collected evidence will be assembled, summarized and reported to address the research questions. We will discuss results in dedicated consultations and workshops with field experts considering gaps in methodology and implications for policy, practice and research. Review protocol under publication (Zenodo https://about.zenodo.org/).

Results: At the time of this submission, the review team defined the scope of the review, turning the key concepts into research questions. The search strategy and screening process are ongoing. The scoping review will be finalized by September 2020.

Conclusions: The results of this scoping review may inform all parties involved in planning, conducting, funding, overseeing, reviewing and publishing PM research.

Patient or healthcare consumer involvement: The European Patients' Forum, member of the PERMIT consortium, participated in topic and scope definition, and were involved in the definition of the scoping review protocol. They will also be involved in the discussion of the scoping review findings and consultation exercise.

Scoping reviews on climate-health topics: recommendations for methods development

Thomson D[1], Hayes K[2], Dennett L[3], Newton AS[1], Ebi KL[4]

[1] Department of Pediatrics, University of Alberta, Canada; [2] Climate Change and Innovation Bureau, Health Canada, Canada; [3] John W Scott Health Sciences Library, University of Alberta, Canada; [4] Department of Global Health, University of Washington, USA

Background: The multiple health impacts of climate change include the effects of increased heat, changes in the geographic range or seasonality of infectious diseases, or stress on health system weaknesses, with disproportionate impacts on marginalized and vulnerable populations. Due to the complexity of this topic, relevant evidence is complex and heterogenous. Scoping reviews can map the literature relevant to a research question, and determine patterns and trends. Recently, several scoping reviews have been published on the health impacts of climate change and extreme weather events. There are limitations to available methodological guidelines for conducting scoping reviews in climate-health topics; specific recommendations would improve the rigour and utility of reviews.

Objective: To consider challenges for scoping reviews of climate-health topics and present recommendations for future methods.

Methods: We conducted a scoping review of scoping reviews on climate-health topics, extracting the research question/objective, methods guidance cited, and definition (if any) of climate change that was used. We combined the findings with the experience of team authors.

Results: Areas where further guidance would be useful include: Establishing definitions of key terms that are workable for evidence synthesis purposes. Climate change, and climate change adaptation, are not simple or straightforward topics. Many scoping reviews equate evidence on the health impacts of extreme weather events with evidence of the impacts of climate change itself; however, this may not account for the dynamic nature of climate change-related impacts in the future. Authors should also develop familiarity with the statistical method of detection and attribution, which can determine the extent to which an extreme event was due to climate change.

Sources of evidence: Existing research may not always be sufficient to answer particular questions, and author teams may therefore need to find additional sources, such as expert knowledge, Delphi and other approaches, input from relevant populations and/or indigenous knowledge. Methods for incorporating this evidence into scoping reviews are needed. Balancing breadth and feasibility of what is included in the scoping review. The health impacts of climate change are, in many cases, not new, but rather differences in existing patterns of disease, exacerbations in pressures on healthcare systems, etc. Studies that are relevant to the scoping review question may not be framed from a climate change perspective, and the review must be designed to encompass literature from a wide range of disciplines. Review teams have to consider how to deal with large quantities of citations for screening, and data for inclusion.

Conclusions: Modifications to existing methodological guidelines for conducting scoping reviews are needed to improve the rigour and utility of reviews of climate-health topics.

Patient or healthcare consumer involvement: None.

Systematic review and meta-analysis of radiation therapy for various head and neck cancers

Cao L[1], Yao L[2], Li M[1], Yang K[3]

[1] School of Public Health, Lanzhou University, China; [2] Health Research Methodology, Department of Health Research Methods, Evidence and Impact, McMaster University, Canada; [3] Evidence- Based Social Science Research Center, School of Public Health, Lanzhou University, China

Background: Head and neck cancer covers a heterogeneous group of cancers, which includes paranasal and sinonasal cancer and cancer of the salivary gland, lip, oral cavity, pharynx and larynx. Treatment options for head and neck cancer consist of radiotherapy, surgery, chemotherapy or a combination of these modalities. There is a growing body of evidence suggesting that more aggressive treatment regimens, such as altered fractionation schedules for radiotherapy or (concomitant) chemoradiation improve tumor control and survival.

Objectives: To assess the overall survival rates, local control rates and progression-free survival of radiation therapy for head and neck cancers and identify the difference in LC and toxicity between adjuvant, salvage, and primary therapy using radiation.

Methods: We conducted a systematic literature review using PubMed Web of science and Embase databases to assess information available regarding the overall survival rates, local control rates, progression-free survival and recurrence for carbon-ion and proton radiation therapy for head and neck cancers. Two review authors independently extracted the data. We used the case series evaluation tool to assess the quality of included studies. We used the Stata 12.0 software to perform meta-analysis.

Results: We included 10 articles, involving 701 participants. The quality of the included studies ranged from moderate to high. The results of the meta-analysis showed that five-year overall survival was significantly higher after carbon-ion therapy compared to conventional protons therapy (52% versus 44%, P = 0.007). Also, five-year local control after carbon-ion therapy was significantly higher for head and neck cancer compared to protons therapy (92% versus 74%, P = 0.045). The meta-analysis showed two- and five-year progression-free survival after carbon-ion therapy is 42% and 45%. In addition, significant reductions were observed in recurrence rates (28%, I^2 = 89%).

Conclusions: Compared with the proton therapy, carbon-ion radiation therapy may prolong the overall survival of patients with head and neck cancer at two and five years. In addition, carbon-ion radiotherapy may have significant effects on local control rates, progression-free survival rates, adverse reactions, and reduction of recurrence rates in patients with head and neck cancer. However, since the overall quantity and quality of data regarding carbon-ion and proton therapy is poor and there is high heterogeneity across trials, these results need to be interpreted with caution. Well-designed and rigorous studies will be required in the future.

Patient or healthcare consumer involvement: Not applicable.

Systematic review of mobile health: an overview

Wang Q[1], Lai H[2], Guo T[2], Xiao Y[2], Ge L[1]

[1] Evidence-Based Social Science Research Center, School of Public Health, Lanzhou University, China; [2] Department of Social Science and Health Management, School of Public Health, Lanzhou University, China

Background: With the rapid development of wireless communication technology and the popularity of mobile phones, mobile medicine provides a powerful platform for patients to receive personalized medical services and real-time and convenient communication in recent years. Mobile medicine has been widely used in our daily life. Using mobile medicine technology to help people's health has become common. The number of systematic reviews in this area has increased gradually, however, the quality of different systematic reviews is unclear, some results are even contradictory, which makes it difficult for people to make effective decisions in the face of complex evidence. It is necessary to evaluate the subject through a re-evaluation of the systematic review, which aims to provide more concentrated and high-quality evidence for practical work and decision-making.

Objectives: To conduct an overview of systematic review in mobile medicine.

Methods: We searched PubMed, Embase, MEDLINE, Chinese National Knowledge Infrastructure (CNKI) and the Cochrane Library to collect systematic reviews and meta-analysis about mobile medicine as fully as possible. We included "mobile medicine", "systematic review", "meta-analysis" and extracted the first author, year of publication, name of mobile medicine, number of studies, funding sources and so on. Four review authors independently screened literature and extracted data using EndNote software. Any conflicts in the results are resolved after careful discussion. We used the AMSTAR scale (A MeaSurement Tool to Assess systematic Reviews) to analyze the methodology quality of the included researches.

Results: This study is ongoing and all results will be presented at the Cochrane Colloquium as available.

Conclusions: This study is ongoing and all results will be presented at the Cochrane Colloquium as available.

Patient or healthcare consumer involvement: Not applicable.

Systematic reviews addressing questions of prevalence: the past, present and future

Munn Z[1], Barker T[1], Borges Migliavaca C[2], Stein C[3], Colpani V[3], Falavigna M[4]

[1] Joanna Briggs Institute (JBI), University of Adelaide, Australia; [2] Programa de Pós-Graduação em Epidemiologia, Universidade Federal do Rio Grande do Sul, Brazil; [3] Hospital Moinhos de Vento, Brazil; [4] Programa de Pós-Graduação em Epidemiologia, Brazil

Background: Prevalence systematic reviews and meta-analysis is an emerging methodology in the field of evidence synthesis. The number of published reviews has increased more than tenfold over the last ten years. These reviews can provide useful information for healthcare professionals and policymakers on the burden of disease, show changes and trends over time in disease, and inform geographical distributions of disease and conditions. The estimates also can be used to inform the absolute impact on health outcomes, from association measures from clinical studies (e.g. relative risk) and be used for estimating costs through the development of economic models.

Objectives: To report on the work of the Prevalence Estimates Reviews – Systematic Review (PERSyst) Methodology Group, including an overview of the current state of prevalence reviews and developments planned for the future.

Methods: We formed a methodological working group to create guidance for conducting systematic reviews of studies reporting prevalence. As part of the group's work, we performed a methodological cross-sectional study evaluating how systematic reviews of prevalence have been conducted.

Results: The evaluation of how prevalence systematic reviews are conducted displayed substantial variability in methods for searching, study selection, risk of bias, analysis and reporting. This work has led to the group drafting a program of work and ways forward to advance the state of systematic reviews of prevalence.

Conclusions: Prevalence reviews are increasingly being performed to inform policy and decision making. To improve the conduct of such reviews further work is required to develop reporting guidelines, establish the certainty in estimates of prevalence, for risk of bias and for analysis of prevalence information.

Patient or healthcare consumer involvement: Systematic reviews of prevalence are important to inform priority of problems and baseline risk, which is important for shared decision making.

The development of search strategies for climate change and health in MEDLINE and Embase

Lampard P[1], Golder S[1]

[1] University of York, UK

Background: Search strategies for systematic reviews have to manage a trade-off between sensitivity and precision. Search filters are useful for conducting searches in Cochrane Reviews and the Cochrane RCT filter is commonly used. The increasing urgency for the retrieval of studies on the health impact of climate change (across all areas of health and levels of health system) requires a search strategy which maximises sensitivity as well as precision. Previous searches have attempted to retrieve the totality of research in this area (as required for any review), but none have reported on the sensitivity or precision of their search strategies.

Objectives: To create search filters in the area of climate change and health in MEDLINE and Embase and to improve on the methodology of creating search filters.

Methods: We used objective methods (such as online word frequency software (www.writewords.org)) and subjective methods (using terms from previously employed strategies and terms already known to the searchers) to obtain the climate change and health terms. These terms, using systematically randomized samples, were then subject to univariate analysis, to establish relevance by counting relevant articles in the samples, and multivariate analysis, to explore the relationship between terms when searched for in the same search string. Once the strategies' search terms were finalized, we screened the full results for inclusion in Endnote. Precision was established by dividing the number of relevant articles by the total number of articles retrieved. An estimate of sensitivity was based on running samples (N = 1000) for each database using only the climate change terms. Articles relevant to health were counted and, by working out the percentage captured by our search filter, we established an estimate of sensitivity. We aimed for an acceptable precision of ≥ 50% and an estimate of sensitivity ≥ 90%.

Results: Following deduplication, we retrieved 7236 articles from MEDLINE, of which 4924 were counted as true positives (68.04% precision). 10,602 articles were retrieved from Embase, of which 6377 were counted as true positives (60.14% precision). The samples used for estimating sensitivity found 155 articles out of 1000 relevant to health for MEDLINE, of which 145 were identified by our search (93.5% estimated sensitivity). In Embase this number is 164 out of 1000, of which 15 were identified by our search (90.9% estimated sensitivity).

Conclusions: The performance measures in these search strategies, for both MEDLINE and Embase, are both higher than those generally accepted for systematic reviews. The methods we used to create these search strategies may be generalizable to other filters useful for Cochrane Reviews.

Patient or healthcare consumer involvement: Due to the nature of this study, patients or healthcare consumers were not involved.

The effectiveness of ginger for the treatment of dysmenorrhea: a systematic review and meta-analysis

Cao L[1], Li X[1], Li M[1], Yang K[1]

[1] School of Public Health, Lanzhou University, China

Background: Dysmenorrhoea is painful uterine cramps of menstrual origin and is the most common gynaecological complaint in women. It is highly prevalent in women of reproductive age and is responsible for a decrease in quality of life, absence from work or school, reduced participation in sport and social activities. Ginger (Zingiber officinale Roscoe) is a monocotyledonous in the family Zingiberaceae originating from southeast Asia, with a complex mixture of pharmacological compounds containing several hundred known constituents, including gingerols, beta-carotene, capsaicin, caffeic acid, curcumin, and salicylate.

Objectives: To evaluate the current evidence for the effectiveness of ginger for treating dysmenorrhea.

Methods: We searched five electronic databases (PubMed, Embase, the Cochrane Library, Web of Science and Chinese Biomedical Databases (CBM)) from their inception to January 2020. We included publications in English of randomized controlled trials comparing ginger against placebo or active treatment in people with dysmenorrhea. We assessed the quality of the studies using the risk-of-bias tool recommended by the Cochrane Handbook version 5.1.0. We performed meta-analysis with RevMan version 5.3. Outcomes of interest included pain severity, severity of dysmenorrhea.

Results: We included eight articles, involving 903 participants. The quality of the included studies ranged from moderate to high. The results of the meta-analysis showed that compared with the placebo group, significant effects of ginger were observed on reducing pain severity (mean difference (MD) -1.73, 95% confidence interval (CI) -2.51 to -0.94, P < 0.05). No significant difference was found between ginger and non-steroidal drugs in the severity of dysmenorrhea (P > 0.05).

Conclusions: Available evidence suggests that ginger could be an effective treatment for menstrual pain in dysmenorrhea. However, we should interpret the findings of these reviews with caution, considering the small number of studies and the overall limited methodological quality.

Patient or healthcare consumer involvement: Not applicable

The effects of interventions for ADHD in children and adolescents: an updated review of systematic reviews

Elvsashagen M[1], Dahlgren A[1], Eidet LM[1], Øgrim G[2]

[1] Regional Centre for Child and Adolescent Mental Health (RBUP) Eastern and Southern Norway, Norway; [2] Østfold Hospital Trust, Norway

Background: Children and adolescents with ADHD are vulnerable and in need of evidence-based treatments. Research on interventions that treat their condition is indispensable. In this review we present results from systematic reviews concerning the treatment of ADHD in children and adolescents.

Objectives: To review the evidence on the effect of treatments for children (< 18 years) with ADHD.

Methods: We conducted an update of a review of systematic reviews. We searched for reviews in the database IN SUM, covering systematic reviews of the effects of treatments relevant to child and youth mental health. We also handsearched the websites of The National board of health and welfare (SE), The Health and Medicines Authority (DK), and NICE (UK). We identified 29 systematic reviews, of which we included 20 evaluating interventions for ADHD. Exclusion reasons for the nine reviews were overlap with the already included reviews or overlap with the other newly found reviews. We assessed the quality of the 20 included reviews using AMSTAR (A MeaSurement Tool to Assess systematic Reviews). We used the GRADE criteria to assess the certainty of the evidence and made judgements about the overall effectiveness for each treatment, weighing the benefits and harms.

Results: This update includes 39 systematic reviews from the original review and 20 new systematic reviews published between March 2018 and October 2019. Nearly half of the newly identified reviews (eight reviews) concern pharmacological treatments, and one review evaluates cognitive behavioral therapy. Other reviews cover interventions that have not previously been evaluated. These are massage therapy, vitamin D, peer support interventions, and combination treatments. There is evidence for the effectiveness of training and educational interventions, including parent training, social skills training and cognitive training with digital tools. Furthermore, cognitive behavioral therapy appears to be effective. Physical activity programs and school-based measures can help alleviate symptoms. The benefits and harms of customized diet or supplements are unknown, but supplements of omega 3 and omega 6 may be helpful. Massage therapy has promising effects on some outcomes, while peer inclusion interventions may have a small effect. Vitamin D supplement may reduce ADHD symptoms, and pharmacological treatment in combination with non-pharmacological treatment seem to have a positive effect.

Conclusions: Several treatments for ADHD show promising effects, however many of the commonly used treatments for children and adolescents are associated with large uncertainties. In particular, there is a need for research on electrophysiological treatments, psychological therapies, mindfulness interventions and acupuncture and dietary supplements. The high number of systematic reviews published in one year reveals large activity in this field of research and demonstrates the need for regular updates of the evidence.

Patient or healthcare consumer involvement: This review aims to improve the public and healthcare providers' access to reliable evidence on treatments for ADHD.

The heterogeneous methods used to conduct systematic reviews of prevalence studies in adult populations: a cross-sectional study

Buitrago-Garcia D[1], Robles Rodriguez WG[2], Eslava-Schmalbach J[3], Salanti G[4], Low N[4]

[1] Institute of Social and Preventive Medicine, Graduate School of Health Sciences, University of Bern, Switzerland; [2] Maestría en Epidemiología Clínica, Fundación Universitaria de Ciencias de la Salud (FUCS), Colombia; [3] Facultad de Medicina, Universidad Nacional de Colombia, Colombia; [4] Institute of Social and Preventive Medicine, University of Bern, Switzerland

Background: Systematic reviews (SRs) of prevalence can provide information about the prevalence of a condition or disease, identify differences in prevalence between populations, inform stakeholders and clinicians about the burden of a disease, and assess the impact of interventions on the prevalence over time and space. However, few resources have been invested so far into understanding and enhancing the methods underpinning systematic reviews of prevalence studies. In particular, empirical evidence about systematic reviews of prevalence studies is scarce.

Objectives: To examine the methodological characteristics of published systematic reviews of prevalence studies in adult populations.

Methods: We conducted a systematic search in the following databases: MEDLINE-Ovid, Embase-Ovid, CINAHL and LILACS. We included systematic reviews of prevalence studies related to medical conditions in adult populations from January 2010 to September 2019. We collected data using a 30-item form. We performed a descriptive analysis using Stata16.

Results: We included 870 systematic reviews. There was an eightfold increment in the number of publications over the years (2% in 2010 to 17% in 2019). Most of them aimed to answer questions related to psychiatric conditions (15%), cardiovascular conditions (10%), respiratory conditions (5%), diabetes (5%), and neurological conditions (5%). The median of included studies was 23 (interquartile range (IQR) 13 to 42). The quality of the included studies was assessed in 61% of SRs; we identified 38 different tools used to perform this step. It is noteworthy that 15% of authors chose to design their own tool or to adapt an existing tool for their systematic review. Meta-analysis was done in 61% of the included SRs, and in 45% of them, the authors did not mention in detail the method used. In the cases where it was mentioned, the most frequently used transformation was the Double Arcsine, followed by DerSimonian and Laird's random-effects synthesis model. Heterogeneity was assessed in 93% of meta-analyses, mainly statistical heterogeneity using the I^2 statistic.

Conclusions: The increase in the publications of SRs of prevalence makes evident the importance of these studies, which are relevant for a variety of populations and conditions. Although there are methodological proposals for developing SRs of prevalence, there is a need to establish a consensus in order to remove biases and support the publication of high-quality evidence. Many decisions made in public health and clinical practice are based on the evidence gathered from systematic reviews. In many cases, the only data available come from prevalence studies. Due to the heterogeneity in the development of these studies, there is a potential risk of introducing bias in the reported evidence.

The nature and quality of systematic reviews of prognostic studies

Fayter D[1], Wolff R[1], Kleijnen J[1]

[1] Kleijnen Systematic Reviews Ltd, UK

Background: Understanding and improving the prognosis of a disease or health condition has been identified as a priority in clinical research and practice. The number of systematic reviews (SR) of prognostic studies is increasing and guidance exists for their conduct.

Objectives: To investigate the nature and quality of SRs of prognostic studies using a random sample from KSR Evidence (a database including all SRs and meta-analyses in healthcare published since 2015).

Methods: We generated a 2016 random sample of SR publications and screened records to identify reviews classified as 'prognostic'. We extracted data on review type (according to the PROGRESS framework) and study characteristics (including country and disease type classified according to ICD-10). We appraised all studies using an adapted version of the ROBIS (Risk Of Bias In Systematic Reviews) tool. ROBIS considers four domains: study eligibility criteria; identification and selection of studies; data collection and study appraisal; and synthesis and findings. From these four domains an overall summary of the risk of bias (ROB) is generated.

Results: From a random sample of 516 SRs, we identified 87 (17%) prognostic reviews. Most were SRs of one or more prognostic factors of disease (79, 91%). Two (2%) were classified as fundamental prognostic research, four (5%) considered the development, validation or impact of prognostic models. None considered the use of prognostic information to tailor treatment decisions and two (2%) covered multiple categories. Prognostic SRs were identified in thirteen disease areas with the majority relating to cancer (26, 30%) followed by mental health (10, 11%) and circulatory diseases (10, 11%). Twenty-two countries were represented in the sample with China contributing the most SRs (21, 24%) followed by the USA (14, 16%) and the UK (11, 13%). Generally, SRs were at high ROB across all four domains with domain 2 (identification and selection of studies) being the weakest (85% at high or unclear ROB). Overall, just seven SRs (8.5%) were at low ROB. The main areas of concern were not reporting a comprehensive search strategy, restrictions on sources of information and an inadequate assessment of the quality of the included studies.

Conclusions: SRs of prognostic studies are increasing in number but according to our random sample, despite extensive guidance available, the majority are at high ROB. The literature is dominated by cancer and most SRs investigated prognostic factors rather than models. We aim to use this research as a starting point for further exploration of trends in prognostic SRs.

Patient or healthcare consumer involvement: Whilst no healthcare consumers were involved, gaining a better understanding of the methodology of prognostic research should ultimately result in better healthcare outcomes for patients.

The reliability, usability, and utility of tools to appraise quality and risk of bias in systematic reviews: a prospective evaluation of AMSTAR, AMSTAR 2 and ROBIS

Gates M[1], Gates A[1], Duarte G[2], Cary M[3], Becker M[4], Prediger B[4], Vandermeer B[1], Fernandes RM[5], Pieper D[4], Hartling L[1]

[1] Alberta Research Centre for Health Evidence, Department of Pediatrics, University of Alberta, Canada; [2] Clinical Pharmacology Unit, Instituto de Medicina Molecular, University of Lisbon, Portugal; [3] Centre for Health Evaluation and Research (CEFAR), National Association of Pharmacies, Portugal; [4] Institut für Forschung in der Operativen Medizin, Department für Humanmedizin, Universität Witten/Herdecke, Germany; [5] Clinical Pharmacology Unit, Instituto de Medicina Molecular, University of Lisbon; Department of Pediatrics, Santa Maria Hospital, Portugal

Background: Readers of systematic reviews (SRs) and overview authors require valid, reliable, and practical means to evaluate the methodological quality and risk of bias of SRs. Evidence of the comparative reliability, usability, and utility of common tools will inform their use and interpretation.

Objective: To evaluate and compare the inter-rater and inter-centre reliability, usability, and utility (how the tool may be used to inform the inclusion of SRs in overviews) of three available tools for appraising the quality or risk of bias of SRs: AMSTAR, AMSTAR 2, and ROBIS.

Methods: Using a sample of 30 SRs of randomized trials, two reviewers at each of three centres (Canada, Germany, and Portugal) independently appraised the methodological quality or risk of bias of each SR using AMSTAR, AMSTAR 2, and ROBIS in a random sequence and reached consensus. To test for inter-rater reliability between pairs of reviewers and consensus decisions between centres, we used Gwet's AC1 statistic. To estimate usability, we calculated the median (interquartile range (IQR)) time to complete the appraisal and reach consensus for each tool. To inform utility in informing the inclusion of SRs in overviews, we tested for associations between methodological quality or risk of bias and the results and conclusions of the SRs.

Results: Reviewers completed AMSTAR, AMSTAR 2, and ROBIS in median (IQR) 15.7 (11.3), 19.7 (12.1), and 28.7 (17.4) minutes, and reached consensus in 2.6 (3.2), 4.6 (5.3), and 10.9 (10.8) minutes, respectively. Across all centres, inter-rater reliability was substantial to almost perfect (AC1 0.61 to 0.99) for 8/11 (73%) AMSTAR, 9/16 (56%) AMSTAR 2, and 12/24 ROBIS (50%) items. Inter-centre reliability was substantial to almost perfect for 6/11 (55%) AMSTAR, 12/16 (75%) AMSTAR 2, and 7/24 (62.5%) ROBIS items. Inter-centre reliability for confidence in the results of the review or overall risk of bias was moderate (AC1 0.58, 95% confidence interval (CI 0.30 to 0.85) to substantial (AC1 0.74, 95% CI 0.30 to 0.85) for AMSTAR 2 and poor (AC1 -0.21, 95% CI -0.55 to 0.13) to moderate (AC1 0.56, 95% CI 0.30 to 0.83) for ROBIS. There was no clear relationship between centre-specific appraisals and the results or conclusions of the SRs.

Conclusions: Compared to AMSTAR 2 and ROBIS, reviewers completed AMSTAR appraisals more quickly and with better agreement. Inter-centre reliability was highest for AMSTAR 2, but ratings on the overall confidence in the results was variable. Both inter-rater and inter-centre reliability were highly variable for ROBIS. Low levels of inter-centre reliability, particularly on overall ratings of confidence or risk of bias, may limit readers' ability to interpret the ratings applied by various review groups. It is not clear whether reviewers' appraisals could be used to inform the inclusion or exclusion of SRs in overviews without altering the overview's results or conclusions.

Patient or consumer involvement: Patients and consumers were not directly involved, but the findings will assist consumers in interpreting appraisals reported in overviews.

The role of health consumers in learning health systems: scoping the literature

Zurynski Y[1], Smith K[1], Vedovi A[1], Ellis L[2], Gul H[2], Braithwaite J[1]

[1] Partnership Centre for Health System Sustainability, Australian Institute of Health Innovation, Macquarie University, Australia; [2] Centre for Healthcare Resilience and Implementation Science, Australian Institute of Health Innovation, Macquarie University, Australia

Background: The US Institute of Medicine identified four areas necessary to establish a Learning Health System (LHS): science/informatics, patient-clinician relationships, incentives, and culture. Consumer engagement in the development, implementation, evaluation and quality improvement of LHSs is considered essential for the establishment of a functioning and sustainable LHS. As research activity and translational research is integral to LHSs, tensions between the roles of health consumers as "patients" and as "research participants" are inevitable. This is further complicated by the conceptual separation between clinical care and research, ethical approval processes and privacy legislation.

Objectives: To explore to what extent consumer engagement is reported in peer-reviewed literature and what it means to patients to be cared for under an LHS framework.

Methods: A narrative/scoping review (January 2016 to May 2019), using terms ("learning health system" and "learning health care system") in PubMed and Scopus; 197 papers had an explicit discussion of LHSs. Six of these focussed on at least one aspect of LHS consumer engagement.

Results: Thirty-five papers reported challenges and barriers around health consumer engagement limiting LHS progress. Several papers reported the difficulties of gaining multiple permissions to access patient data for clinical use or research. Fifty-nine papers referred to health consumer participation or accessing consumer data. Only six papers, including one literature review, focussed on the role of health consumers in an LHS. Three concentrated on accessing patient health records or specific data from patient reported outcome or experience measures. The other three papers discussed the expectation to participate in research while accessing clinical care in LHS facilities. Patients were likely to participate in research if they perceived benefits and had developed trusting relationships with their clinicians. One study found that patients feel obliged to take part in research studies as it is an expectation in an LHS. Two papers raised the issue of the appropriateness of seeking informed consent when undertaking research about clinical care delivery as this type of research crosses the boundaries between clinical care, research and quality improvement.

Conclusions: Data and knowledge about patients is a significant driver of LHSs; however, we found few studies about LHSs that explored the role of patients or health consumers. Many studies spoke in general terms about patient engagement and reported the lack of engagement as a barrier to LHS implementation. More research is needed to explore the best ways of engaging with consumers when developing and implementing LHSs.

Patient or healthcare consumer involvement: Consumers were not directly involved in this scoping review. The Partnership Centre for Health System Sustainability collaborates with the Consumers Health Forum of Australia.

The top cited systematic reviews and meta-analyses: a bibliometric analysis

Hou L[1], Wang Q[1], Zhang Q[2], Zhang W[1], Ge L[1]

[1] Evidence-Based Social Science Research Center, School of Public Health, Lanzhou University, China; [2] School of Nursing, Lanzhou University, Lanzhou, China

Background: The last few decades have witnessed the establishment of evidence synthesis, particularly systematic reviews and meta-analyses, as a key component of evidence-based medicine (EBM). Meta-analyses of randomized trials have become more widely accepted by clinicians, researchers and policy makers as a useful tool to critically assess the totality of evidence in a research question. However, little work has been done to identify the great scientific output in this field. Citation analysis has been regarded as a useful method to evaluate the impact of articles.

Objectives: To identify and analyze the most highly cited systematic reviews and meta-analyses. To analyse the main features of the 100 most-cited articles in the field, excluding methodology studies.

Methods: We searched the literature on 25 November 2019 using Clarivate Analytics 'Web of Science Core Collection (WoSCC)'. The search subjects were 'systematic review*', 'meta-analysis' and 'meta-analyses', 'meta analysis', 'meta analyses', 'metaanalysis', 'metaanalyses', 'pooled analysis' and 'pooled review' in the title section. There was no restriction on the publication year of the article. Using the Clarivate Analytics 'Web of Science Core Collection (WoSCC)', we ranked the selected articles in descending order on the basis of their citation counts. Two review authors independently read the abstract of each article on the list. We excluded methodological studies. Finally, we reached a unanimous decision on the list of the top 100 most-cited publications from the included systematic reviews and meta-analyses. We used VOSviewer (Centre for Science and Technology Studies, Leiden University, Leiden, The Netherlands) and CiteSpace (Chaomei Chen, China) to make visualization mapping in this paper.

Results: Our initial search identified a total of 207,673 papers, with 60,248 published as 'article' and 97,129 classified as 'review'. Of those, 13,177 articles were cited more than 100 times. After excluding methodological studies about systematic reviews and/or meta-analyses, we included the top 100 most-cited publications. More information about main features of 100 most-cited articles are in progress.

Conclusions: Our study is ongoing.

Patient or healthcare consumer involvement: Not applicable

Therapeutic interventions for childhood cancer: an umbrella review of randomized evidence

Lavasidis G[1], Markozannes G[1], Panagiotou O[2], Trikalinos NA[3], Petridou ET[4], Voorhies K[5], Ntzani EE[1]

[1] Evidence-based Medicine Unit, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Greece; [2] Department of Health Services, Policy & Practice, Brown University School of Public Health, USA; [3] Division of Oncology, Department of Medicine, Washington University in St Louis, USA; [4] Department of Hygiene, Epidemiology and Medical Statistics, School of Medicine, National and Kapodistrian University of Athens, Greece; [5] Department of Biostatistics, Brown University School of Public Health, USA

Background: Treatment advances in pediatric cancer have substantially improved prognosis of the second most common cause of child mortality in developed countries. However, the strength of the supporting evidence has not yet been thoroughly evaluated.

Objectives: To critically appraise the evidence in the field, we performed an umbrella review of meta-analyses (MAs) of randomized controlled trials (RCTs) examining the efficacy and safety of therapeutic interventions for pediatric malignancies.

Methods: We searched PubMed from inception to July 2019 for relevant MAs. For each study, we re-estimated the summary effect size using random-effect and fixed-effects models, as well as 95% confidence and prediction intervals. We further estimated the between-study heterogeneity using the I^2 metric, assessed evidence of small-study effects and excess significance bias and evaluated the replicability, consistency and quality (using AMSTAR 2, A MeaSurement Tool to Assess systematic Reviews) of the evidence.

Results: Eleven MAs assessing 54 comparisons and 240 individual study estimates were eligible. The median number of RCTs/MA was 3 (range 2 to 16), with a median of 841 participants/MA (range 275 to 8873). Acute lymphoblastic leukemia (ALL), notably the commonest leukemia type, was the most frequently investigated cancer (29 MAs), followed by sarcomas (12 MAs). The most commonly assessed outcomes were event-free survival and overall survival (22 and 8 MAs, respectively). The summary random effects were statistically significant (P < 0.05) in 30% of the comparisons but the 95% prediction intervals excluded the null value in only 2 meta-analyses. We detected moderate to high heterogeneity (I^2 = 50% to 75%) in seven MAs, while only one showed very high heterogeneity (I^2 > 75%). There were limited indications for small-study effects (n = 1) and excess significance bias (n = 1). There was evidence of replicability/consistency of the treatment effect in seven MAs. Based on previously applied evidence grading criteria (P value, I^2, largest study, prediction intervals, small-study effects, excess-significance), the most concrete evidence pertained to the use of methotrexate and vincristine plus prednisone pulses for ALL, both improving event-free survival. The evidence pertaining to other cancer types was relatively weak. Finally, the quality of the included studies was generally low.

Conclusions: We found few small MAs of randomized evidence of varying consistency assessing therapeutic interventions for pediatric cancer focusing mainly on ALL and usually assessing event-free survival as an outcome. Randomized evidence stemming from adult populations seems to serve as a valuable indirect evidence backup feeding informed clinical decisions. More and better-quality RCTs as well as MAs of individual patient data are needed to increase certainty and precision in the care of pediatric cancer patients.

Patient or healthcare consumer involvement: None.

Understanding patients, physicians and caregivers' perspectives on Artificial Intelligence: a targeted review of qualitative studies

Chu X[1], Zhang N[1], Feng L[1], Lu C[1], Yang K[1]

[1] Lanzhou University, China

Background: Artificial intelligence (AI) refers to the development of computer algorithms to accomplish tasks traditionally associated with human intelligence, such as the ability to learn and solve problems. AI has been applied in some medical fields. Potential benefits of AI may be increased diagnostic certainty, faster turnaround, and better quality of work life for physicians and caregivers. Although the role of AI is anticipated to increase soon, this technology should also be embraced by patients, physicians and caregivers, who are important but still neglected stakeholders. At present, it is still unknown how these groups view the developments of AI in medical fields in terms of awareness of this topic, uncertainties, and expectations.

Objectives: To better understand the underlying patients, physicians and caregivers' perspectives on AI, qualitative research may be ideally suited to discover subtleties and nuances, which often cannot be gleaned from quantitative research alone. Thus, we conducted a targeted literature review of the qualitative evidence on patients, physicians and caregivers's perspectives on AI.

Methods: We conducted systematic and reproducible literature searches in PubMed, Embase, Web of Science and the Cochrane Library for articles published from the earliest publication date available. We also screened the reference lists of relevant reviews and meta-analysis. No language restrictions were applied. Two authors independently included studies and extracted data. In addition, two independent researchers evaluated the articles using the National Institute for Health and Care Excellence (NICE) quality appraisal checklist qualitative studies, according to which articles were scored on the basis of their theoretical approach (whether the qualitative design was appropriate), study design, data collection method, trustworthiness, analysis, and ethics. Based on the checklist, we rated articles as good, fair, or mixed.

Results: Final results will be available by the time of the Cochrane Colloquium.

Conclusions: The findings of this project will provide information about current patients, physicians and caregivers' perspectives on AI. The relevant researchers should focus on the level of knowledge of AI in these groups and identify domains related to the use of AI in medical fields.

Patient or healthcare consumer involvement: None.

RAPID REVIEWS

A rapid horizon scanning review: digital interventions to reduce alcohol-related harm

Campbell F[1]

[1] University of Sheffield, UK

Background: The published data for some interventions is already out of date. The rapidly evolving field of digital interventions is one such example. In order to address this challenge and to inform the future funding priorities of our commissioners, we adopted a horizon scanning approach which, due to limited time frames, had to be undertaken rapidly. This case study illustrates the approach we used to meet these objectives.

Objectives:

1) To identify and describe innovations and newly emerging digital interventions that can be used to reduce alcohol related harm.
2) To consider how these compare with existing digital interventions.
3) To identify potential gaps in the types of digital interventions that are being developed or subjected to rigorous evaluation.

Methods: We conducted a rapid horizon scanning review of the published and grey literature on digital interventions to reduce alcohol-related harm covering the period January 2017 to April 2019. We used and updated an existing search strategy from a relevant Cochrane Review. The 'grey literature' search was informed by the horizon scanning methods by HTA organizations (CADTH, AHRQ) and our search included trial registries, grants awarded, patents, new alerts and app stores. We used a narrative approach to data analysis.

Results: We included over 150 published studies, 32 registered trials, five patents, and 276 apps. Populations that are most at risk of alcohol-related harm have been understudied relative to community-dwelling and student populations. There were notable recent trends for digital interventions to be increasingly delivered via the internet and smartphone apps, sometimes in combination with biosensors, and for interventions for delivery in clinical and primary prevention settings. Alcohol reduction apps available on the app stores may increasingly be using interactive features rather than simply providing information.

Conclusions: This is an active and rapidly evolving field of research and technological development. It is important to develop and evaluate digital interventions for populations that are most at risk of alcohol-related harm and use a range of research methods to evaluate newly emerging digital technologies. It is also important to ensure that the features of effective digital interventions are identified, and consider how these might differ depending on the population that is receiving the intervention. Given the speed of development in this field there is a need to keep the evidence base regularly updated and a 'living systematic review' may be a way forward providing a high quality, online summary of health research which is updated as new research becomes available.

Patient or healthcare consumer involvement: This project was funded by an independent charity; healthcare consumers were involved in designing the protocol and in the dissemination of the results of the report.

Accelerating integration of emerging evidence into healthcare delivery: rapid reviews for learning health systems

Hager M[1], Jungbauer R[1]

[1] Oregon Health & Science University, USA

Background: Ideally, clinical care is based on real-time, best available evidence. In reality, research findings take on average 17 years to translate into policy and practice. Developing and implementing health system guidelines using a rapid review process can reduce the time it takes to integrate emerging evidence into care. Recently, Oregon Health & Science University (OHSU) created a learning health system with local community hospitals, focused on delivery of consistent, evidence-based care in the region. As part of that system, OHSU created the Office of Clinical Integration and Evidence-based Practice (OCI-EBP), a team dedicated to reviewing, evaluating, and applying best available healthcare evidence.

Objectives: The objectives of the OCI-EPB are to enable the learning health system to rapidly adapt as best evidence and healthcare delivery evolves, as well as relieve clinicians from time-consuming evidence review and summation, to improve the experience and outcomes for clinicians and patients.

Methods: Guideline topics are nominated by an oversight committee and are developed in partnership with multidisciplinary content expert teams, engaging representatives from each hospital, and patient advocates. The OCI-EBP uses GRADE methodology to appraise and summarize research evidence. Content expert teams use their clinical expertise to inform development of clinical questions, practice recommendations, and consensus statements. The OCI-EPB creates clinical decision support tools to support the guideline's implementation. Multidisciplinary, clinical teams formally implement each guideline, design workflows that promote the delivery of consistent care, and use validated quality metrics to evaluate and apply continuous improvement efforts.

Results: To date, the OCI-EPB has engaged more than 150 learning health system members to develop nine clinical guidelines, including opioid prescribing, colorectal cancer screening, and others. Current guidelines under development include obesity management and pediatric pneumonia. Post-implementation data from the guidelines have shown reductions in length of stay, readmission rates, and opioid use. Using our rapid review process and established partnerships, the OCI-EPB has contributed to Coronavirus care by conducting rapid reviews of emerging research to inform predictive modeling, development of triage risk assessment tools, and management of confirmed cases.

Conclusions: The OCI-EBP has been an innovative model for driving evidence-based change management and decision-making at OHSU. The team has worked closely with clinicians and health system members to critically analyze emerging evidence and rapidly build consensus; this framework and level of co-ordination is necessary for supporting meaningful integration of the best evidence available to improve patient care and quality.

Patient or healthcare consumer involvement: Patient advocates are included throughout process to provide their perspective during the appraisal and adaption of evidence into recommendations.

An international survey reveals current worldwide practice for rapid reviews of diagnostic tests

Arevalo-Rodriguez I[1], Steingart K[2], Tricco A[3], Nussbaumer-Streit B[4], Kaunelis D[5], Alonso-Coello P[6], Baxter S[7], Bossuyt P[8], Emparanza JI[9], Zamora J[1]

[1] Clinical Biostatistics Unit, Hospital Universitario Ramón y Cajal, Instituto Ramón y Cajal de Investigación Sanitaria (IRYCIS), CIBER of Epidemiology and Public Health, Spain; [2] Department of Clinical Sciences, Liverpool School of Tropical Medicine, UK; [3] Li Ka Shing Knowledge Institute, St. Michael's Hospital, Unity Health Toronto; Epidemiology Division, Dalla Lana School of Public Health and Institute of Health Policy, Management and Evaluation, University of Toronto, Canada; [4] Cochrane Austria, Danube University Krems, Austria; [5] Canadian Agency for Drugs and Technologies in Health (CADTH), Canada; [6] Iberoamerican Cochrane Center-Servicio de Epidemiología Clínica y Salud Pública, Biomedical Research Institute (IIB-Sant Pau), Spain; [7] School of Health and Related Research (ScHARR), University of Sheffield, UK; [8] Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Amsterdam University Medical Centres, University of Amsterdam, The Netherlands; [9] Clinical Epidemiology Unit, Hospital Universitario Donostia, BioDonostia, CIBER of Epidemiology and Public Health, Spain

Background: Rapid reviews (RRs) have emerged as an efficient alternative to meet the demand for accelerated evidence synthesis for healthcare decision-making. In a previous scoping review, we examined the characteristics of RRs of diagnostic tests by scrutinizing repositories of Health Technology Assessment (HTA) agencies and papers published in indexed journals. However, due to an incomplete description of the methods used, as well as inclusion of only published reports, we could not provide a detailed account of the current practice for the development of RRs of diagnostic tests.

Objective: To perform an international survey to better understand how RR methods are currently used to synthesize diagnostic evidence.

Methods: We invited representatives from institutions that perform evidence synthesis from all over the world to participate in this closed survey, including members of the International Network of Agencies for HTA, the World Health Organization collaborating centers on HTA, the HTA Network of the Americas, and the HTA International Network (non-profit members). We collected data from April to July 2019 and de-identified and anonymized the survey responses for all analyses. We performed all descriptive analysis using STATA 15.0.

Results: We contacted 74 institutions by email, and 25 of them indicated that they performed RRs of diagnostic tests. All these institutions reported the implementation of one or more methods to define the scope of the RR, e.g. limiting the number of index tests (76%) and limiting the intended applications of the test (80%). However, only one strategy (defining a structured question) was used by ≥ 90% of participants. All participants used at least one methodological shortcut, including the use of a previous review as a starting point (92%) and the use of limits on the search (96%). Parallelization and automation of review tasks were not extensively used (48% and 20%, respectively).

Conclusion: To the best of our knowledge, this is the first international survey assessing the current practice of methods for diagnostic test RRs. Our survey indicates greater use of shortcuts and limits for conducting diagnostic test RRs versus the results of our previous scoping review analyzing published RRs. However, only two strategies (i.e. defining a structured question and the use of a previous review as starting point) were used for ≥ 90% of participants. Several shortcuts are used without knowing how their implementation affects the results of the evidence synthesis in the setting of diagnostic test reviews. Thus, a structured evaluation of the challenges and implications of the adoption of these RR methods is warranted.

Patient or healthcare consumer involvement: No patients were involved in this research.

Co-ordination and characterization of rapid reviews at WHO in response to COVID-19

Norris SL[1]

[1] WHO, Switzerland

Background: The World Health Organization (WHO) uses the best available evidence to inform its policies and technical guidance, and to guide program implementation. The WHO Incident Management Team (IMT) for COVID-19 required rapid identification of research studies to inform daily strategic decisions as well as its normative guidance.

Objectives: To describe the approach for prioritizing and co-ordinating rapid reviews, and the characteristics of the reviews performed in response to the COVID-19 outbreak.

Methods: The WHO Rapid Review Team (RRT) emerged in an organic manner, with one technical member of staff soliciting topics for reviews starting early February 2020. This evolved into a small team of experienced review authors who responded to ad hoc requests for answers to specific questions from WHO staff in the IMT and other technical units. In addition, the RRT commissioned rapid reviews from external review teams, using a list of prioritized questions from IMT.

Results: At the time of writing, 16 rapid reviews have been commissioned from external teams and three were performed by RRT members. An additional four topics were reviewed in brief by RRT members. Topic scope ranged from very broad (e.g. quarantine) to very focused (e.g. single drugs such as remdesivir). The 16 external reviews took 5 to 10 days to produce by teams ranging in size from 5 to 30 members; reports ranged in length from 8 to 150 pages. Chinese-language literature was reviewed in most reviews; pre-prints were rarely included. Fifteen of the 16 external reviews used indirect evidence from diseases caused by other coronaviruses or from influenza. The quality of the reviews was uniformly high as assessed with AMSTAR 2 (A MeaSurement Tool to Assess systematic Reviews). The most common deviations from full systematic reviews were: (1) searching of a limited number of bibliographic databases; (2) lack of dual title and abstract screening; (3) abbreviated quality assessments at the study level; and (4) exclusively qualitative summaries.

Challenges included: (1) difficulty engaging busy key IMT staff to guide prioritization and formulation of review key questions; (2) co-ordination of ad hoc requests for reviews coming from multiple sources within WHO; and (3) managing expectations of IMT members regarding feasible scope and timelines for high-quality reviews.

For the efficient and optimal use of evidence to inform WHO guidance in this continuing outbreak and in for future emergencies there is a need: (1) to identify multiple members of the IMT with varied expertise who have the capacity to respond to urgent requests from the RRT; (2) for training and socialization of rapid review methods between emergencies; (3) to have standard operating procedures in place prior to the onset of an emergency; and (4) to have a cadre of rapid review teams 'on call' in the event of emergencies.

Conclusions: The RRT has co-ordinated a large number of high-quality reviews within a short period of time. It was challenging to orchestrate with busy IMT staff and to manage expectations. Detailed planning is needed between emergencies for the optimal implementation of an RRT.

Cochrane Task Exchange experience during the coronavirus pandemic

Borges do Nascimento IJ[1], Zakarija-Grkovic I[2], Poklepovic T[2], O'Mathúna D[3], Marcolino MS[1]

[1] Clinical Hospital, Federal University of Minas Gerais, Brazil; [2] Faculty of Medicine of Split, Cochrane Croatia, Croatia; [3] The Ohio State University, USA

Background: International research collaboration is desirable as it broadens perspectives, increases funding opportunities and improves visibility. However, in contrast to the view that geographic boundaries are falling due to globalization, international research collaboration can be difficult to establish. Research co-ordinators frequently require previous publications and experience from interested researchers or even rank applicants by affiliation. Cochrane Task Exchange plays a unique role in connecting researchers from around the world, regardless of academic titles, affiliations, publication records or previous experience. During pandemics, resulting networks can contribute to rapid-learning health systems.

Objectives: To describe and assess the experience of establishing a research network via Cochrane Task Exchange during the novel coronavirus pandemic.

Methods: Two posts were published on Cochrane Task Exchange, in January and March 2020, calling for researchers interested in collaborating with other Cochrane members on conducting a systematic review about COVID-19. All applications were appraised by the call issuer, based on candidates' interests, aspirations, and availability for short- and long-term contributions. Selected candidates were included in an online instant messaging group to discuss research purposes and ongoing projects. Assessment of performance will be based on research output. Participants' experience will be based on an online survey containing 14 open- and 2 closed-ended questions. We will upload the survey onto Google Forms to evaluate applicants' experience of using Cochrane Task Exchange and collaborating with the research group.

Results: Overall, 39 candidates applied, volunteering to join the research collaboration through Cochrane Task Exchange. Eleven candidates contacted the team supervisor directly and expressed willingness to help. Many candidates (63%) did not reply to further e-mail contact after their initial application. Finally, eleven candidates were selected to be part of the team based on their availability and personal interest in the topic. As of 1 April 2020, the group remains and has so far published two systematic review protocols, one scoping review with meta-analysis, submitted for publication one overview of systematic reviews, and is currently finishing an update of the first review. Experience, so far, has shown that clear descriptions of the team's expectations should be planned in advance, as well as descriptions of individual roles for each participant, in order to avoid a posteriori conflicts regarding involvement. Results from the survey will be presented at the Colloquium.

Conclusions: Cochrane Task Exchange is an excellent online platform for connecting people from all over the world who need help and are willing to collaborate in research projects, and by doing so contribute to efforts to make the world a better place.

Patient or healthcare consumer involvement: Not applicable.

Computers are your friend? Using artificial intelligence in rapid reviews

Shaw E[1], Robalino S[1], Lazur B[1], Harrod C[1]

[1] Center for Evidence-based Policy, Oregon Health & Science University, USA

Background: Rapid reviews often use methods not considered the 'gold standard' to expedite evidence reviews. Increasingly, systematic review software, such as DistillerSR, and online reviewing tools offer automated approaches to systematic review tasks that are traditionally conducted by human reviewers. The intent of automation is to speed up routine tasks (e.g. title and abstract screening); thus, freeing up the human capacity for tasks that require it. Automation also provides other approaches to improving the quality of reviews. In practice, reviewers can face practical and philosophical hurdles to using machines for systematic review tasks. Therefore, we ask, can machines help us as members of a rapid review team?

Objectives:

1) Explore how machine learning can help with rapid reviews.
2) Describe a case study of using machine learning in a rapid review environment.

Methods: Using Google Scholar, we searched for recent abstracts and published papers on the use of DistillerSR's artificial intelligence (AI) capacity in systematic reviews. We focused on papers published since 2019 to ensure the results were as applicable to the current version as possible. For the case study, we selected a rapid review conducted using our standard methods (i.e. a single reviewer with quality assurance from a senior researcher) and assessed how we could use DistillerSR's AI in the rapid review process. We also explored the potential benefits and limitations of our approach.

Results: DistillerSR has been evaluated as an automated tool in screening for a range of publication types, including for randomized controlled trials. However, one evaluation of the use of DistillerSR suggested that there are limitations to replacing human screening with automated screening alone. Based on the published evaluations, we decided that in our context, DistillerSR would be a useful second screener to quickly exclude references at the title and abstract screening stage. We also explored how DistillerSR could be used to 'check' the decisions of the human reviewer at the full-text stage. We will present our experience of using DistillerSR's AI tool, including the practical challenges we faced.

Conclusions: Automated processes for screening is increasingly promoted as an effective and efficient way to improve decision accuracy and reduce review time. While we found that machine screening can be useful in providing another level of certainty in reviewer decision, we also encountered practical challenges. Challenges included understanding strengths and limitations of machine selection and applying processes in real life. We also faced challenges of reviewer confidence in results. Are machines our friend? We think so, but we would like to get to know them better.

Patient or healthcare consumer involvement: Not relevant for this submission.

Defining Rapid Reviews: a systematic scoping review and thematic analysis of definitions and defining characteristics of rapid reviews

Hamel C[1], Michaud A[1], Thuku M[1], Skidmore B[1], Stevens A[1], Nussbaumer-Streit B[2], Garritty C[1]

[1] Ottawa Hospital Research Institute, Canada; [2] Cochrane Austria, Department for Evidence-Based Medicine and Clinical Epidemiology, Danube University Krems, Austria

Background: Rapid reviews (RRs) have become a pragmatic alternative to systematic reviews (SRs) because they can provide decision-makers timely answers to urgent health system questions. Implementing one or more shortcuts in the conduct of the review usually saves resources. However, there is currently no consensus around the definition of a RR.

Objectives: To perform a systematic scoping review to identify RR definitions used in published RRs (between 2017 and 2018) and to perform a thematic analysis of these definitions to allow for the creation of a preliminary definition to be further discussed with the SR and RR community.

Methods: An Information Specialist developed a search strategy in consultation with the review team, which was peer-reviewed. We performed the study selection in two stages: titles and abstracts were screened using the liberal accelerated method and full-text screening was performed independently, in duplicate. Data charting included copying the definition of a RR as stated in the included reviews, and the reference(s) (if applicable). We supplemented definitions from RRs with RR definitions from RR methods papers. We then thematically analyzed definitions and presented them graphically (e.g. radar chart).

Results: We found 2657 unique records, evaluated 422 at full-text, and included 216 RRs. A total of 158 RRs provided a definition, and the top four articles referenced were Khangura 2012 (n = 54), Ganann 2010 (n = 42), Tricco 2015 (n = 21) and Grant 2009 (n = 18). Including the definitions retrieved from the 90 RR methods papers identified in another scoping review, we thematically analyzed 204 definitions and mapped them to eight key themes: compare/contrast to a full traditional SR (n = 139), variation in methods shortcuts (n = 112), accelerated/rapid process or approach (n = 99), resource efficiency rationale (n = 99), stakeholder rationale (n = 72), systematic approach (n = 50), focus/depth/breadth of scope (n = 35), and bias/limitations (n = 19) (Figure 1).

Figure 1. Reporting of key themes
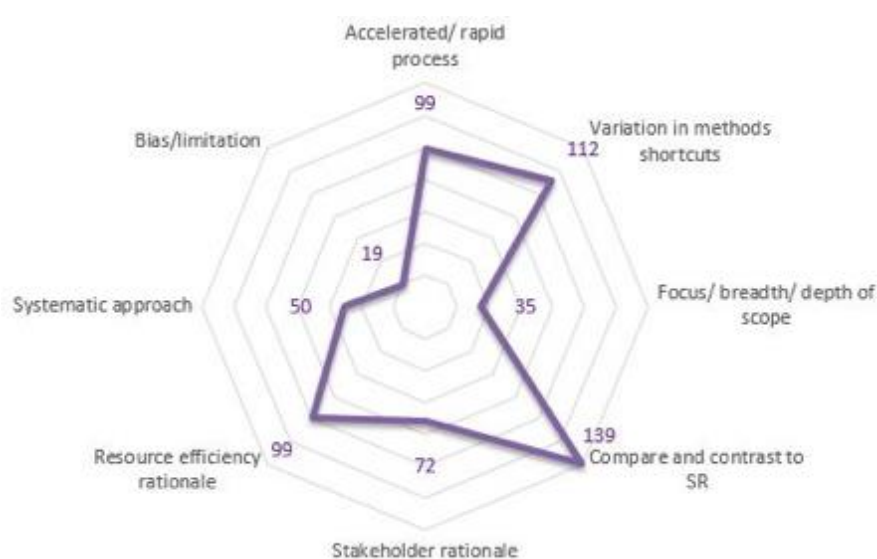
Conclusions: We identified eight key themes and propose a preliminary definition of a RR. This suggested definition, with additional caveats, can help guide discussions with the SR community to

develop a consensus definition. This definition will also serve to inform discussions within Cochrane regarding possible future implementation of RRs. Failure to achieve consensus on the definition or at least a minimum set of criteria of a RR will be a barrier to moving the science forward in this field.

Patient or healthcare consumer involvement: Although there was no direct patient or consumer involvement, the results from this scoping review aim to provide the first step in achieving consensus around a definition of a RR. This should provide developers of RRs and stakeholders to get high-quality information more quickly.

Effectiveness and safety of glucocorticoids to treat COVID-19: a rapid review and meta-analysis

Lu S[1], Zhou Q[2], Wang Z[2], Shi Q[2], Chen Y[3]

[1] School of Medicine, University of Electronic Science and Technology of China, Chengdu, China; [2] School of Public Health, Lanzhou University, Lanzhou, China; [3] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China

Background: Glucocorticoids are widely used in the treatment of various pulmonary inflammatory diseases, but they are often accompanied by significant adverse reactions. Published guidelines point out that low dose and short duration systemic glucocorticoids may be considered for patients with rapid disease progression in adults or children with COVID-19. However, there is still no effective evidence on the possible harms of this intervention.

Objectives: To systematically review the effectiveness and safety of glucocorticoids in COVID-19 patients.

Methods: We comprehensively searched electronic databases (from 2003), completed predefined handsearching, and contacted experts. Two review authors applied inclusion criteria. We included randomized controlled trials (RCTs) and cohort studies evaluating the effectiveness and safety of glucocorticoids in children and adults with COVID-19, SARS and MERS patients, and combined the same outcome indicators for meta-analysis and systematic review.

Results: Our search retrieved 23 studies, including one randomized controlled trial and 22 cohort studies, with a total of 13,815 patients. Our meta-analysis showed that in adults with COVID-19, the use of systemic glucocorticoids did not reduce mortality (risk ratio (RR) 2.00, 95% confidence interval (CI) 0.69 to 5.75, $I^2 = 90.9\%$) or the duration of lung inflammation (weighted mean difference (WMD) -1 day, 95% CI -2.91 to 0.91). A significant reduction was shown in the duration of fever (WMD -3.23 days, 95% CI -3.56 to -2.90). Using glucocorticoids did not reduce mortality (RR 1.52, 95% CI 0.89 to 2.60, $I^2 = 84.6\%$), duration of fever (WMD 0.82 days, 95% CI -2.88 to 4.52, $I^2 = 97.9\%$), or duration of lung inflammation absorption (WMD 0.95 days, 95% CI -7.57 to 9.48, $I^2 = 94.6\%$) in SARS patients. The use of systemic glucocorticoids significantly prolonged the length of hospital stay in COVID-19, SARS and MERS patients.

Conclusions: Glucocorticoid therapy significantly reduced the duration of fever but did not reduce the mortality, hospital stay, or lung inflammation absorption in patients with COVID-19. Long-term use of high-dose glucocorticoids increased the risk of adverse reactions such as infections. Routine use of systemic glucocorticoids for patients with COVID-19 population is therefore not recommended.

Patient or healthcare consumer involvement: None.

Efficacy and safety of antibiotic agents in children with COVID-19: a rapid review

Wang J[1], Tang Y[2], Estill J[3], Ma Y[4], Zhou Q[4], Luo Z[2], Liu E[2], Chen Y[1]

[1] School of Public Health, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, Lanzhou, China; [2] Department of Respiratory Medicine; Children's Hospital of Chongqing Medical University; Chongqing Key Laboratory of Pediatrics, Chongqing, China; [3] Institute of Global Health, University of Geneva; Institute of Mathematical Statistics and Actuarial Science, University of Bern, Switzerland; [4] WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, Lanzhou, China

Background: The outbreak of COVID-19 is the third introduction of a highly pathogenic coronavirus into the human population in the twenty-first century, after the Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS) epidemics. At present, there are no standardized or specific treatment schemes for COVID-19 patients, and the clinical treatment mainly focuses on symptomatic and supportive care.

Objectives: The aim of this review was to evaluate the efficacy and safety of antibiotic agents in children with COVID-19, as well as to introduce the present situation of antibiotics use and bacterial coinfections in COVID-19 patients.

Methods: We searched the Cochrane library, MEDLINE, Embase, Web of Science, CBM, Wanfang Data and CNKI from their inception to 29 February 2020. In addition, we searched related studies on COVID-19 published before 29 February 2020 through Google Scholar and preprint servers. We evaluated the risk of bias of included studies, and synthesized the results using a qualitative synthesis.

Results: Five case series and one cohort study met our inclusion criteria. Five studies on SARS showed an overall risk of death of 7.2% to 20.0%. One study of SARS patients who used macrolides, quinolones or beta lactamases showed that the mean duration of hospital stay was 14.2, 13.8 and 16.2 days respectively, and their average duration of fever was 14.3, 14.0 and 16.2 days. One cohort study on MERS indicated that macrolide therapy was not associated with a significant reduction in 90-day mortality (odds ratio (OR) = 0.84, 95% confidence interval (CI) 0.47 to 1.51, P = 0.56) and improvement in MERS-CoV RNA clearance (hazard ratio (HR) = 0.88, 95% CI 0.47 to 1.64, P = 0.68). Besides, 24 studies testified that the proportion of antibiotic use ranged from 16.0% to 19.4% in children and 12.4% to 100% in adults. The most commonly used antibiotic in adults was quinolones and in children were cephalosporins and macrolides, despite the lack of etiological evidence.

Conclusions: The benefits of antibiotic agents for adults infected with SARS or MERS were questionable in the absence of bacterial coinfections. There is no evidence to support the use of antibiotic agents for children with COVID-19 in the absence of bacterial coinfection.

Patient or healthcare consumer involvement: None

Evolution of an approach to rapidly synthesizing evidence about pressing health-system issues

Wilson M[1], Lavis J[1], Waddell K[2], Gauvin F[2], Moat K[1]

[1] Department of Health Evidence & Impact, McMaster University; McMaster Health Forum, Canada; [2] McMaster Health Forum, Canada

Background: Rapidly synthesizing evidence is critical for supporting evidence-informed policymaking about health systems. Since 2013 we have developed, implemented and iteratively refined a rapid-response program at the McMaster Health Forum that has (in days or weeks) identified and synthesized evidence for Canadian policymakers and stakeholders about more than 50 pressing health-system issues.

Objectives: Our objective was to document the evolution of this program to provide insights for others interested in rapidly synthesizing evidence.

Methods: We used a multi-method approach to document how and why our rapid-response program has evolved and to identify current and future challenges faced in efforts to provide robust yet rapidly synthesized evidence to inform pressing health-system issues. This included a detailed internal program review that was based on internal documentation and interviews with staff, a documentary analysis of products produced through the program and a focus group with those involved in the administration and scientific aspects of running the rapid-response program.

Results: Our experience with conducting rapid syntheses has evolved to: 1) incorporate longer timelines (e.g. 60- or 90-day requests); 2) address both health- and social-system issues; 3) better accommodate the types of complex questions often asked by policymakers (e.g. that synthesize evidence about policy problems, options, implementation considerations, and monitoring and evaluation plans); 4) expand the types of evidence and insights synthesized (e.g. by drawing on systematic reviews and primary studies, as well as from policy documents and key informant interviews); and 5) conduct and integrate multiple types of analyses such as policy, systems and political analysis.

Conclusions: While our approach to conducting rapid syntheses remains underpinned by a commitment to being systematic and transparent in identifying and synthesizing evidence and insights for health- and social-system leaders it has evolved in a way that allows us to go farther, faster in responding to urgent requests.

Patient or healthcare consumer involvement: Topics addressed in rapid syntheses are driven by requests from health-system policymakers and stakeholders, which can include consumer-driven groups. Moreover, when rapid syntheses are conducted for governments or other stakeholders, they are often generated through internal engagement processes within those organizations and generate findings that are relevant to patients and consumers (e.g. how to empower caregivers in home-based restorative care processes) and/or prioritize the need for patient/consumer engagement (e.g. through rapid-learning health systems that are anchored on patient needs, perspectives and aspirations, and focused on improving their care experiences and health at manageable per capita costs and with positive provider experiences).

Mental health and preterm birth: abstract submission for oral presentation

Tsoa E[1], Bourdages N[1], Mather J[1]

[1] Toronto Public Health, Canada

Background: According to the World Health Organization, preterm birth (PTB) is the leading cause of infant mortality and morbidity. Infants born preterm are at greater risk of health and developmental problems, with significant implications for public health. Over the past 10 years, Toronto's PTB birth rate of 9% has been relatively stable, and is higher than the provincial average of 8%. Recent evidence has shown that mental health concerns are a contributing factor to PTB. In Toronto, the rate of women experiencing any mental health concern during pregnancy increased from 8% in 2013 to 10% in 2016. Recent evidence has shown that mental health concerns are a contributing factor to PTB. In Toronto, the rate of women experiencing any mental health concern during pregnancy increased from 8% in 2013 to 10% in 2016. Given these findings, Toronto Public Health (TPH), Canada's largest local public health agency which provides public health programs and services to 2.8 million residents, prioritized a rapid review to determine whether mental health promotion strategies had an influence on PTB.

Objectives: To complete a rapid review of the recent evidence based on the following PICO (Population, Intervention, Comparison, Outcome) question: How do mental health promotion strategies relevant to public health practice influence preterm birth?

Methods: We conducted a rapid review. We searched various electronic databases for systematic reviews and meta-analyses. We screened 726 records based on the inclusion and exclusion criteria. We selected 11 articles for inclusion in the rapid review. Three review authors independently appraised the quality of each article.

Results: The overall results from this review found:

- a lack of evidence of public health interventions addressing mental health concerns during pregnancy that aim to reduce preterm labour and birth;
- that mental health concerns during pregnancy, such as depression, anxiety, and perceived stress can increase the risk for, or is significantly associated with PTB;
- a positive association between exposure to intimate partner violence and PTB.

Conclusions: This rapid review did not find evidence about effective mental health promotion strategies relevant to public health that address PTB. However, the evidence did show an association between depression, anxiety and perceived stress on PTB.

Patient or healthcare consumer involvement: One recommendation from this rapid review is to raise awareness with the general public about the association between PTB and mental health concerns. Toronto Public Health is exploring collaborations with other stakeholders to further pursue this work.

Methodological compromises and their implications in a rapid review of qualitative evidence: a worked example

Ames H[1], Zuske M[2], Bosch-Capblanch X[2]

[1] The Norwegian Institute of Public Health, Norway; [2] Swiss Tropical and Public Health Institute; University of Basel, Switzerland

Background: We were commissioned by the World Health Organization to review the qualitative evidence on community and drug distributor perceptions and experiences of mass drug administration for the elimination of lymphatic filariasis, in the context of a guideline development process.

Objective: To present and discuss the methodological compromises taken during a rapid review and their implications.

Method: We decided to stay as close to Cochrane methodology as possible but had to make compromises based on limited resources. These included single author screening, inclusion, data extraction, methodological assessments and analysis. We did not perform a GRADE-CERQual assessment. During the review process, the authors discussed the steps to be taken and compromises made.

Findings: The single author screening, inclusion and data extraction was fast. We discussed the included and full-text excluded studies. This was a good approach as we had a specific topic and inclusion was mostly related to methodology, not content. Authors with content expertise were consulted if the first author was unsure. We decided on a framework analysis as we felt this would be faster than a thematic analysis as the topics of interest for the guidelines were specific. Single author data extraction and analysis was quick but also a potential challenge as placing the data in the framework and deciding on findings or themes was their responsibility. We did not perform a GRADE-CERQual assessment of our findings. Therefore, we decided to present them as key-messages and create simplified tables with the key-message, contributing studies and study quality. Finally, by using key-message tables instead of a GRADE-CERQual assessment, we delegated the responsibility of understanding and weighing the possible impact of methodological weaknesses to the guideline committee.

Conclusions: We were able to complete a rapid review of evidence quickly. This allowed the inclusion of qualitative data in the guidelines process, something we feel was useful and important to decision-makers. However, the fact that the majority of the work was completed by a single author was seen as a possible limitation as potentially relevant issues could have been overlooked or underdeveloped. However, our work was based on the professional experience of our multidisciplinary team and we did not experience these limitations as compromising the final body of evidence. Nevertheless, guidelines development should be sensitive to methodological issues which could compromise the implementation of useful interventions.

Patient or healthcare consumer involvement: This abstract discusses a methodology that aims to include rich data in qualitative evidence synthesis that can provide a more detailed presentation of health consumers' thoughts, opinions and experiences. This allows synthesis authors to better interpret the meaning and context of findings presented in the primary studies.

Potential effectiveness and safety of antiviral agents in COVID-19: a rapid review

Shi Q[1], Zhou Q[1], Wang Z[2], Lu S[3], Ma Y[2], Xun Y[2], Chen Y[4]

[1] The First School of Clinical Medicine, Lanzhou University, China; [2] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, China; [3] School of Medicine, University of Electronic Science and Technology of China, China; [4] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, Lanzhou, China

Background: The COVID-19 outbreak presents a new, life-threatening disease. Data and evidence on the effectiveness and safety of antiviral agents for patients with COVID-19 are limited.

Objectives: To assess the potential effectiveness and safety of antiviral agents for COVID-19 patients.

Methods: We searched electronic databases from January 2002 to 29 March 2020 for randomized controlled trials, quasi-randomized controlled trials and cohort studies of interventions with antiviral agents for patients (adults and children) with COVID-19.

Results: We included 25 studies with 6124 patients after screening 2879 titles and abstracts and 80 full-text articles. The risks of bias in all studies were moderate to high in general. There is no evidence showing the effectiveness and safety of antiviral therapy for children with COVID-19. Evidence of very low to low-quality suggested that the effectiveness and safety of existing antiviral agents for adult patients with COVID-19 is uncertain: lopinavir/ritonavir had no effect on mortality (risk ratio (RR) 0.77, 95% confidence interval (CI) 0.45 to 1.30) and probability of negative PCR test (RR 0.98, 95 CI% 0.82 to 1.18). Arbidol had no benefit on probability of negative PCR test (RR 1.27, 95% CI 0.93 to 1.73). Hydroxychloroquine was significantly associated with increased probability of negative PCR result (RR 5.60, 95% CI 1.48 to 21.13). For other antiviral agents, we included indirect evidence of SARS and MERS, and evidence of very low to low-quality suggested interferon could reduce corticosteroid dose (weighted mean difference (WMD) -0.14 g, 95% CI -0.21 to -0.07) but with no effect on mortality (RR 0.72, 95% CI 0.28 to 1.88); ribavirin did not reduce mortality (RR 0.68, 95% CI % 0.43 to 1.06) and was associated with high risk of severe adverse reactions; oseltamivir had no effect on mortality (RR 0.87, 95% CI 0.55 to 1.38) in adults with SARS. Ribavirin combined with interferon was also not effective in adults with MERS and associated with adverse reactions. We will update the results as new evidence emerging and provide living evidence to the related stakeholders.

Conclusions: The effectiveness and safety of existing antiviral agents for patients with COVID-19 is still uncertain. We do not suggest clinical routine use of antivirals for COVID-19 (with the exception of clinical trials).

Patient or healthcare consumer involvement: None.

Rapid review methods: a systematic scoping review

Hamel C[1], Michaud A[1], Affengruber L[2], Skidmore B[1], Stevens A[1], Nussbaumer-Streit B[2], Garritty C[1]

[1] Ottawa Hospital Research Institute, Canada; [2] Cochrane Austria, Department for Evidence-Based Medicine and Clinical Epidemiology, Danube University Krems, Austria

Background: Rapid reviews (RRs) have become a pragmatic alternative to systematic reviews (SRs) because they can be completed in much less time and provide decision-makers quick answers to urgent health system questions. The savings in resources is usually achieved by implementing one or more "shortcuts" in the review stages. Various RR shortcut methods have been used to produce RRs, but there is lack of agreement how best to conduct and deliver more timely reviews.

Objectives: The aim of this work was to perform a systematic scoping review to identify studies that have assessed one or more shortcut methods applicable for undertaking RRs and mapping these to review conduct stages and MECIR guidance.

Methods: An Information Specialist developed a search strategy in consultation with the review team, which was peer-reviewed. We searched several electronic databases (e.g. MEDLINE, Embase) to identify the published literature, and searched for grey literature on websites of organizations that produce RRs. The initial database search produced over 30,000 records, so we implemented a more targeted strategy. We performed study selection in two stages: we screened titles and abstracts using the liberal accelerated method, with some screening using an artificial intelligence tool. We performed full-text screening independently, in duplicate. Data charting included the stage of conduct of the method evaluated, the area of research, details on the nature of the comparison/evaluation, and a synopsis of related results. We have presented information by stage of review conduct to identify existing and gaps in research. We also performed a mapping exercise to MECIR guidelines to show similarities and differences.

Results: The searches resulted in 1873 unique records, of which 156 were further evaluated, and 90 publications were included. Publications were divided into four categories:

1) Formal evaluation (n = 14);
2) Development, which included four subcategories (n = 65);
3) Comparison (n = 2); and
4) Applying reporting guidelines/critical appraisal tools (n = 3).

Six studies were classified as "SR surrogates". Four formal evaluation studies were composite evaluations, including more than one shortcut simultaneously. The remaining 10 studies evaluated searching, screening, data extraction, and "other" areas (e.g. involving stakeholders) (Figure 1). Due to complexities around shortcuts evaluated, in terms of methods and types of shortcuts, only a cursory mapping to MECIR criteria was possible.

Conclusions: Some methods shortcuts may be useful in the context of RRs, but there are limitations in the included studies that may limit their applicability. The results will serve to inform discussions within Cochrane regarding possible future implementation of RRs.

Patient or healthcare consumer involvement: Although there was no direct patient or consumer involvement, the results from this scoping review will provide needed information to RR producers to get high-quality information to those who need it in an expedited timeline.

Figure 1. Mapping to key dimensions
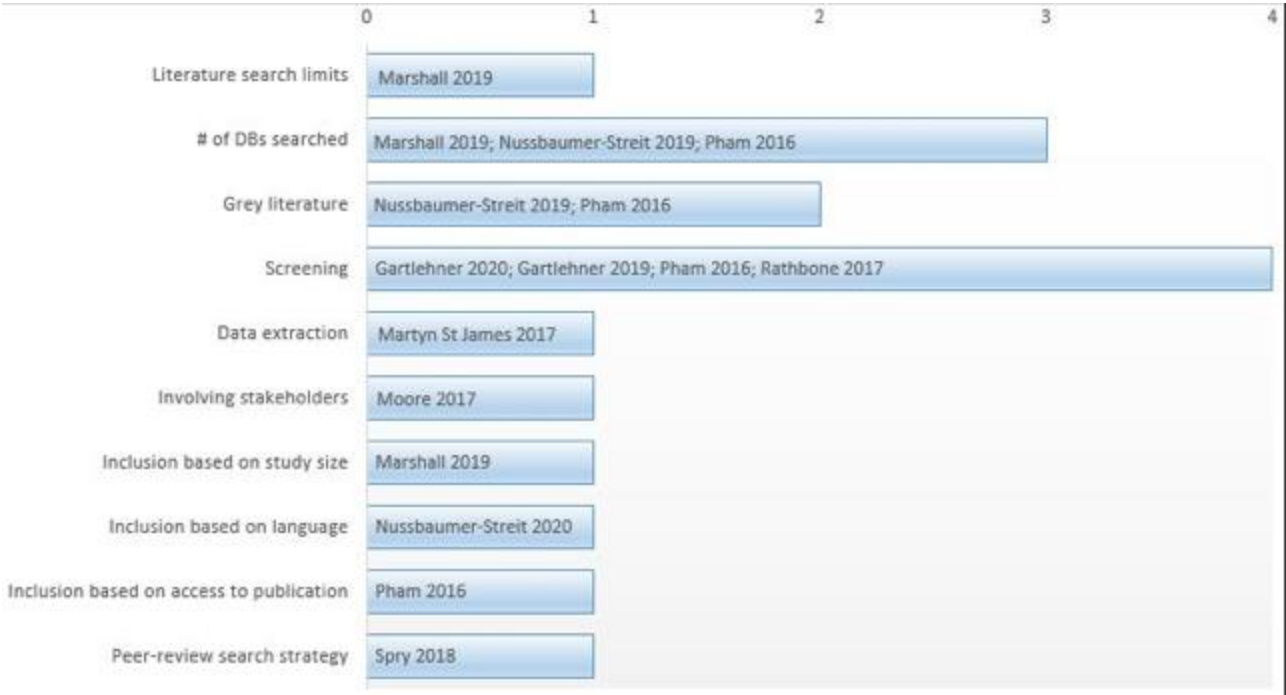
Rapid reviews in rapid learning health systems

Klugar M[1], Pokorná A[2], Slezáková S[1], Vrbová T[3], Líčeník R[1], Dušek L[4], Klugarová J[1]

[1] Czech National Centre for Evidence-Based Healthcare and Knowledge Translation (Cochrane Czech Republic, Czech CEBHC: JBI Centre of Excellence, Masaryk GRADE Centre), Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University, Czech Republic; [2] Czech National Centre for Evidence-Based Healthcare and Knowledge Translation and Department of Nursing, Faculty of Medicine, Masaryk University, Czech Republic; [3] Czech Health Research Council, Czech Republic; [4] Institute of Health Information and Statistics of the Czech Republic; Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University, Czech Republic

Background: In the information era, 'oversaturation' is an essential keyword, especially in the context of health systems and health policymakers. Some policymakers prefer to inform their statements and decisions based on some evidence. Our role as systematic reviewers and evidence-based healthcare advocates is to advocate for the best available evidence, which would ideally inform all decisions. However, the main question remains is whether the evidence used by health policymakers is unbiased, not confounded and rigorously and transparently synthesized. The usual answer is no. For most questions we do not have a high-quality rigorously and transparently developed evidence synthesis.

Objectives: The objective of this work is to explore if the rapid review is 'rapid' enough in rapid learning health systems.

Methods: Expert evidence of systematic reviewers from Czech National Centre for Evidence-Based Healthcare and Knowledge Translation (Cochrane Czech Republic, Czech Centre of Evidence-Based Healthcare: JBI centre of excellence, Masaryk University GRADE Centre) supporting health policymakers by timely evidence synthesis on the national level.

Results: There exist several equations to calculate how much time is needed to develop rigorous and unbiased synthesis of evidence. The greatest unknown in these equations is usually the number of retrieved studies after searching. However, searching itself, the search strategy and number of databases of published and unpublished studies would influence the total number of retrieved studies as well. The usual question of policymakers is: "When can your team synthesize the evidence for us?" The usual answer to this is "How much rigour would you like to have in your synthesis? It will take six months, and if we want to have a publication, probably sixteen months; or three months, if we do not search all possible databases, it will still be rigorous enough but no 'typical publication' will be possible." At this point the policymakers usually stop us and say, "we need it in a week's time". Our answer to this is, "Alright, if we stop doing anything else for the next week and our work is 'fast and dirty' the synthesis is possible and will still be rigorous enough."

Conclusions: While finishing this abstract, there are several Cochrane rapid reviews which are planned to be developed and published within two weeks because of the extraordinary COVID-19 pandemic. However, shouldn't we be able to accommodate such real rapid reviews in our normal lives? Although the situation in these days is unique, our health systems need 'real' rapid reviews. Hopefully, at least one positive lesson learned from the COVID-19 pandemic will be that 'real' rapid reviews are possible.

Patient or healthcare consumer involvement: It should be possible for all relevant stakeholders, including patients and healthcare consumers, to be involved in rapid review development, even if this is a 'real' rapid review developed and published in two weeks.

Rapid reviews of healthcare interventions (2018 to 2019): a scoping review

Parsons RN[1], Althuis M[1], Kimmel EK[1], Bowen JD[1], Lawrence W[1], Slutsky J[1]

[1] Patient-Centered Outcomes Research Institute (PCORI), USA

Background: Funders, clinicians, policy makers and healthcare organizations want high-quality evidence in a timely and cost-effective manner to support decisions about healthcare interventions. Rapid reviews (RRs) are an evidence synthesis product that streamline systematic review methods. Currently no consensus definition, official guidance for conducting RRs, or established criteria for when a RR should or could be used in place of a systematic review exists. Consequently, PCORI seeks to better understand the methods and reporting of RRs as well as to explore how this streamlined approach may be used to inform clinical effectiveness research.

Objectives: To perform a scoping review of published RRs of healthcare interventions in order to:

1) Describe review characteristics and frequency of streamlined methodologic approaches;
2) Assess reporting quality;
3) Understand stakeholder involvement and ascertain insights into product utility.

Methods: We searched Web of Science, PubMed, and Cochrane to identify reports described in the title or abstract as RRs. We report the number of RRs published in the last five years (2015 to 2019) and conduct a detailed data extraction and assessment of study characteristics and reporting quality for recent RRs (2018 to 2019).

Results: Published RRs of healthcare interventions doubled from 2015 to 2019. Recent publications (2018/2019; N = 57) show that RRs were most often produced by academic organizations (56%) and focus on nonpharmaceutical interventions (65%). Three-quarters (75%) cited disparate sources of informal RR guidance; 19% included meta-analysis; and 88% assessed risk of bias of individual studies. Broadly employed streamlined approaches to methods included narrowing the scope (100%), parallelization of tasks (26%), automating review tasks (4%), and using review shortcuts (100%). For the latter, approximately half of RRs employed one reviewer or one reviewer with a verifier to perform title/abstract screening; fewer rapid reviews applied limited reviewer methods for full text screening (30%), data extraction (42%), and critical appraisal (18%). Preliminary findings from 2019 RR publications suggest that reporting quality ranges widely with PRISMA compliance noted for 35% to 98% (median = 70%) of checklist items. Less than 20% specifically called out the streamlined approaches that distinguished the report as a RR and rationale for use. Additionally, stakeholder involvement and time to complete the review were infrequently reported (4% and 18%, respectively). RRs consistently identified research gaps or provided recommendations for future study (84%). This work is ongoing and additional results will be presented at the Colloquium.

Conclusions: Preliminary findings suggest that RRs focused on healthcare interventions use a wide range of streamlined approaches with variable reporting quality. Future RRs would benefit from structured reporting strategies and formal guidance.

Patient or healthcare consumer involvement: None.

Rapid synthesis for health policies: interventions to reduce unintended pregnancies among adolescents

Wachira V[1], Farinasso C[1], Lima A[1], Borges R[1], Andrade K[1], Regô D[1], Leão L[1], Sachetti C[1], Carvalho V[1]

[1] Ministry of Health Brazil, Brazil

Background: Pregnancy among adolescents is a contributing factor to the increase of infant and maternal mortality, school dropout rates, propagation of precarious health cycles and poverty. It's a common issue in all social classes with a predominance in the most disadvantaged social set ups.

Objectives: To identify and describe political options based on scientific literature to reduce unintended pregnancy among adolescents.

Methods: The problem, its magnitude and causes were interactively clarified through discussions between the authors, consultations with departments in the Brazilian Ministry of Health, key document reviews and research in a limited number of databases. To identify the options, we consulted PubMed, BSV-Adolec, BVS, Epistemonikos, Health Systems Evidence, the Cochrane Library, and e Embase. We included systematic reviews with successful interventions in coping with unintended pregnancy among adolescents, independent of the year of publication or language. We excluded studies that addressed specific adolescent populations or other types of reviews. The search strategy yielded 487 studies, of which 11 met the inclusion criteria. We assessed study quality using AMSTAR 2 (A MeaSurement Tool to Assess systematic Reviews).

Results: We identified four options: implementation of comprehensive educational programs for adolescents in schools and communities, promoting the supply of contraceptives in the school environment, implementation of psychosocial interventions for behavioral changes, and implementation of a group of interventions to increase knowledge about unintentional pregnancy. For each option, we listed the main elements, costs and cost effectiveness in relation to the context of the study location, uncertainties, the perceptions and experiences of the interested parties in the theme, and the considerations about equity and implementation.

Conclusions: The options are complementary since the nature of unintended pregnancy among adolescents is multifactorial. The options adopted should be adapted to the particularities, behaviors and local contexts. Special attention is necessary for adolescents with vulnerability to develop risky behaviors. Moreover, the context of implementation of each intervention should be carefully chosen to increase effectiveness. Any policy aimed at investing in programs related to the theme should complement the existing actions. It is of fundamental importance to encourage adolescents to plan achievements in all areas of life especially in the academic field as well as to facilitate their access to the interventions adopted to cope with unintended pregnancy.

Patient or healthcare consumer involvement: Adolescents should be made aware of the existing interventions as well as involve them in decision making in relation to the option that best fits them.

Safe management of bodies of deceased persons with suspected or confirmed COVID-19 infection: a rapid systematic review

Yaacoub S[1], Schünemann HJ[2], Khabsa J[1], El-Harakeh A[1], Khamis A[3], Chamseddine F[1], El-Khoury R[1], Saad Z[4], Hneiny L[5], Cuello-Garcia C[6], Muti-Schünemann GEU[7], Bognanni A[8], Chen C[9], Chen G[10], Zhang Y[11], Zhao H[12], Abi Hanna P[13], Loeb M[14], Piggott T[6], Reinap M[15], Rizk N[16], Stalteri R[6], Duda S[6], Solo K[6], Chu D[6], Akl EA[17]

[1] Clinical Research Institute, American University of Beirut, Lebanon; [2] WHO Collaborating Centre for Infectious Diseases; Research Methods and Recommendations and the Michael G. DeGroote Cochrane Canada Centre, McMaster University, Canada; [3] Hull York Medical School, University of Hull, UK; [4] Global Evidence Synthesis Initiative, American University of Beirut, Lebanon; [5] University Libraries, Saab Medical Library, American University of Beirut, Lebanon; [6] Department of Health Research Methods, Evidence, and Impact, McMaster University, Canada; [7] Medical School, Vita Salute San Raffaele University, Italy; [8] Department of Health Research Methods, Evidence and Impact, McMaster University, Canada; [9] The Fourth Clinical Medical College, Guangzhou University of Chinese Medicine, China; [10] Dongzhimen Hospital, Beijing University of Chinese Medicine, China; [11] Health Research Methods, Evidence & Impact, McMaster University, Ontario, Canada; [12] Institute of Acupuncture and Moxibustion, China Academy of Chinese Medical Sciences, China; [13] Infectious Disease Division, Beirut Rafic Hariri Hospital, Lebanon; [14] Department of Pathology and Molecular Medicine, McMaster University, Canada; [15] London School of Hygiene and Tropical Medicine, UK; [16] Division of Infectious Diseases, American University of Beirut, Lebanon; [17] Department of Internal Medicine, American University of Beirut, Lebanon;

Background: Proper strategies to minimize the risk of infection in individuals handling the bodies of deceased persons infected with 2019 novel coronavirus (COVID-19) are urgently needed.

Objectives: To systematically review the literature to first scope, and then assess the effects of, specific strategies for the management of the bodies of deceased persons with suspected or confirmed COVID-19 infection.

Methods: We searched five general electronic databases, four COVID-19 specific electronic databases, and three Chinese databases on 26 March 2020. We searched for guidance documents providing practical advice on the handling of bodies of deceased persons with suspected or confirmed COVID-19 infection. Then, we sought primary evidence of any study design on COVID-19 and other coronaviruses. We also searched for evidence relevant to contextual factors (i.e. acceptability, feasibility, resource use and impact on equity).

Results: We identified 23 guidance documents providing practical advice on the steps of handling the bodies: body preparation, packing, and others and advice on the handling of the dead bodies and the use of personal protective equipment (PPE) by individuals handling them. We did not identify COVID-19 evidence relevant to any of these steps. We identified one study proposing an uncommon strategy of handling autopsies for severe acute respiratory syndrome patients. The study provided very low certainty evidence that it reduced the risk of transmission.

Conclusions: While a substantive number of guidance documents propose specific strategies, we identified no study providing direct evidence for the effects of any of those strategies. While this review highlights major research gaps, it allows interested entities to build their own guidance based on the identified guidance documents.

Patient or healthcare consumer involvement: Patient involvement is not applicable. However, the findings of this review will inform the response of the World Health Organization on the safe management of the bodies of deceased persons with suspected or confirmed COVID-19 infection in this pandemic.

The protective measures for health-care workers during the outbreak of COVID-19: A rapid review

Ma Y[1], Wang J[2], Yang N[1], Zhang J[2], Zhang H[1], Chen Y[1]

[1] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, China; [2] School of Public Health, Lanzhou University; Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Cochrane China Network; Chinese GRADE Center, China

Background: On 11 March 2020, the World Health Organization (WHO) said that the COVID-19 outbreak can be characterized as a "pandemic" as the "SARS-COV-2" virus spreads increasingly worldwide. As the people who have the closest contact with patients, a number of healthcare workers have been infected. Healthcare workers are every country's most valuable resource during the outbreak of COVID-19; it is very important to protect their safety.

Objectives: To conduct a comprehensive literature search and summarize the existing evidence to show: 1) what are the current protective measures for allied health personnel during the outbreak of COVID-19, and 2) how effective are the protective measures for allied health personnel?

Methods: Considering the lack of search results on COVID-19, we will carry out the search independently in the following electronic databases from their inception to 24 March 2020: the Cochrane library, MEDLINE, Web of Science, China Biology Medicine disc (CBM), China National Knowledge Infrastructure (CNKI), and Wanfang Data. Moreover, we will also search Google Scholar, the preprint platforms, and the reference lists of the identified reviews for further potential studies. After selection, we will include randomized and non-randomized controlled trials that addressed the protection for health-care workers during the outbreaks of SARS, MERS and influenza. In addition, all types of studies directly targeting the protection of healthcare workers during the COVID-19 will also be included. The risks of bias of the included RCTs and non-RCTs will be assessed by using Cochrane's risk of bias tool and the ROBINS-I tool. We will perform synthesize-analysis using mixed methods. For RCTs and non-RCTs, if sufficient data are available, we will conduct the quantitative analysis for each outcome using forest plots; when effect sizes could not be pooled, we will report the study findings narratively. For other types of COVID-19 studies, we will conduct a qualitative synthesis in order to get comprehensive protection measures.

Results: This rapid review is ongoing, and results will be presented at the Cochrane Colloquium as available.

Conclusions: This rapid review is ongoing, and conclusions will be presented at the Cochrane Colloquium as available.

Patient or healthcare consumer involvement: None.

The Rapid Response Service: synthesized research evidence to inform HIV-related programs, services, and policies in Ontario

Gogolisvili D[1], Giliauskas D[1], Andruszkiewicz N[1], Wasdell M[1]

[1] The Ontario HIV Treatment Network (OHTN), Canada

Background: The Ontario HIV Treatment Network (OHTN) is a non-profit collaborative network that aims to create rapid learning systems to improve the health and lives of people living with and at risk of HIV. A core service of the OHTN is its Rapid Response Service (RRS), an internal program that delivers synthesized evidence to support decision making, intervention planning, and adoption of best practices. The RRS produces and disseminates evidence in an accessible format to partners including community-based AIDS service organizations (ASOs), clinics, internal stakeholders, as well as the Ontario Ministry of Health, its agencies, and other stakeholders in the sector.

Objectives: To describe the scope of the OHTN's RRS, including program methodology and its impact on HIV-related policy and practice.

Methods: Knowledge users can request a Rapid Response on the OHTN's website. The RRS team then engages with the requester to understand needs and expectations, fine-tune the research question, and determine the best approach to the review. The RRS then develops a search strategy and searches MEDLINE/PubMed, PsycInfo, and/or other relevant databases. The RRS team identifies relevant research evidence from the results and selects articles that fit within the scope of the research question. Evidence is drawn from systematic reviews, primary studies, and grey literature if necessary. The synthesized evidence is presented concisely, using accessible language.

Results: As of February 2020, the RRS Team has produced 145 Rapid Responses and published them on the OHTN website. Recent topics include HIV testing, substance use, sexually transmitted infections, and pre- and post-exposure prophylaxis. The number of rapid response downloads from the website increases each consecutive year, totalling 48,000 in 2019. Several of the Rapid Responses have contributed to policy changes at the provincial level in areas such as Narcan nasal spray use, HIV testing intervals, and supervised injection sites for preventing and responding to drug overdose. In our evaluation of the RRS, knowledge users found the service to be valuable, contributing to programmatic decision-making, informing stakeholder organizations' strategic directions, and policy development.

Conclusions: The RRS continues to be a valued resource for OHTN's partners by synthesizing and disseminating research evidence in an accessible format. This ensures that research evidence is usable for people living with HIV, network stakeholders, decision-makers, and other individuals working in the HIV sector in Ontario and beyond.

Patient or healthcare consumer involvement: The RRS aims to support HIV infrastructure in Ontario by providing synthesized evidence to policy makers and people living with HIV. Most Rapid Responses are requested by ASOs that represent consumers and people living with HIV who are actively involved in each stage of the Rapid Response development process. Policy and practice changes developed based on the Rapid Responses also directly benefit healthcare consumers.

The relationship between the type of rapid reviews and quality: in the background of COVID-19

Wang Z[1], Zhou Q[2], Shi Q[2], Xun Y[3], Zhao S[4], Chen Y[5]

[1] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China; [2] The First School of Clinical Medicine, Lanzhou University, Lanzhou, China; [3] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, China; [4] School of Public Health, Lanzhou University, Lanzhou, China; [5] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, Lanzhou, China

Background: In a public health emergency, like the outbreak of COVID-19, rapid reviews (RR) have emerged as a streamlined, instead of systematic review, approach to synthesizing evidence quickly, typically for the purpose of helping decision-makers in healthcare settings to make decisions expeditiously. According to WHO handbook, there are four types of RR: a traditional systematic review (conducted rapidly), rapid review of a systematic review, rapid review of systematic reviews plus primary studies, and rapid review of primary studies only. It is important to know the quality of the different types of rapid reviews for researchers so as to choose the best fit rapid review in a public health emergency.

Objectives: To evaluate the quality of different type of rapid reviews of COVID-19

Methods: We searched the rapid reviews of COVID-19 in electronic databases (including PubMed, Embase, Web of Science, the Cochrane Library) until 6 April 2020. We also searched Google Scholar and the preprint servers. Two researchers independently screened the records and extracted data, disagreements were resolved through discussion or by consulting a third researcher. We excluded non-COVID-19 rapid reviews. We used the Assessment of Multiple Systematic Reviews (AMSTAR) tool and Preferred Reporting Items for Systematic Reviews and Meta Analyses (PRISMA) to assess the methodology and reporting quality of the included rapid reviews.

Results: The result will be presented at the meeting.

Conclusions: The result will be presented at the meeting.

Patient or healthcare consumer involvement: None.

Ultra-rapid synthesis of secondary sources for hospital COVID-19 guidance

Mitchell M[1], Flores E[1], Connolly J[2], Goel R[2], Gottschalk A[2], Kavanagh N[2], Philipson B[2], Wang S[2], Wood C[2], Mull N[1]

[1] University of Pennsylvania Health System, USA; [2] University of Pennsylvania School of Medicine, USA

Background: The COVID-19 pandemic has brought about an urgent need for rapid evidence-based guidance to inform local decision making on infection control, procedures for safely discharging and following up on patients who have been hospitalized, and diverse other topics.

Objectives:

1) Rapidly identify existing guidance for topics such as use, extended use, and re-use of medical masks and respirators; and present that information in a way that facilitates rapid decision-making and action.
2) Provide updated review and synthesis to hospital stakeholders as new guidance is made available and new clinical questions emerge.

Methods: We adapted an existing rapid HTA synthesis product to summarize clinical guidelines, policies from other hospitals, and key society position statements where available. The new prototype product is called a Rapid Guidance Summary. Seven medical students volunteered to assist with gathering evidence and writing reports. Updates to guidance sources were monitored using a combination of technology support (VisualPing, Vancouver, BC, Canada) and daily manual review. We used simple evidence tables to describe the guidance, and concordance tables to report agreement and variations among different sources. Each report included a summary table on the first page reporting key recommendations and the quantity and source of support for those recommendations. We disseminated reports directly to chief medical officers and infection control staff at our affiliated hospitals and outpatient/home care entities and to all system providers via the hospital intranet site.

Results: In the first 10 days of the program, we completed six reports, one of which was an update and elaboration on the first report. Students served as lead analysts on two of those six reports. The mean time needed to complete a report was 3.5 days (range 1 to 5 days). Reports averaged 5 pages in length (range 2 to 9 pages), citing 4 or 5 major guidelines and 3 to 6 hospital policies. Very little guidance was found for some topics, such as criteria for deeming a COVID-19 patient ready for discharge. In other areas such as personal protective equipment (PPE) requirements, there was considerably more guidance, and our challenge was to organize and present that information as succinctly as possible. We identified several areas where guidelines and/or hospital policies were in disagreement. In other areas, hospital policies were frequently based on the same guidance from the Centers for Disease Control and Prevention and the World Health Organization, so their agreement represented less of an expert consensus than it would if the policies had been developed independently.

Conclusions: Evidence-based practice centers can provide timely guidance to clinicians and hospital administrators in a crisis, if centers are willing to adopt new methods as necessary and use non-traditional sources such as policy documents from peer hospitals.

Additional file: [Combined PPD update 404](#)

Using SWIFT-Active Screener to reduce the expense of evidence-based toxicology

Howard B[1], Tandon A[1], Phillips J[1], Maharana A[1], Shah R[1]

[1] Sciome LLC, USA

Background: Systematic review is a formal process used widely in evidence-based toxicology and environmental health research to identify, assess, and integrate the primary scientific literature with the goal of answering a specific, targeted question in pursuit of the current scientific consensus. We recently received Phase I Small Business Innovation Research (SBIR) funding to conduct research and development to enhance our web-based, collaborative systematic review software application, SWIFT-Active Screener.

Objectives: By employing a machine learning methodology called "Active Learning", and through a novel statistical method that can accurately estimate the percentage of relevant studies screened, Active Screener can significantly reduce the overall screening burden compared to traditional approaches.

Methods: We first investigated several improvements to our statistical algorithms used for article prioritization and recall estimation (Aim 1: improved statistical models). The resulting refinements further improve the performance of our algorithms and address critical technical issues that previously limited the applicability of our methods. Secondly, we explored ways in which our models and methods can be improved to handle the scenario in which an existing systematic review is updated with new data several years after its initial publication (Aim 2: new methods for systematic review updates). Finally, in order to ensure that our software is capable of supporting the full demand from our many users, we have reengineered the system to support hundreds to thousands of simultaneous screeners (Aim 3: software engineering for scalability, usability). During this research, our methods and software have been rigorously tested on 26 different systematic review datasets, demonstrating robust performance of Active Screener's prioritization and recall estimation methods in a variety of real-world scenarios.

Results: For reviews with 5000 or more documents, we report an average reduction in screening burden of 61% (to obtain 95% recall). Active Screener has been used successfully to reduce the effort required to screen articles for systematic reviews conducted at a variety of organizations including NIEHS, EPA, USDA, TEDX, and EBTC. These early adopters have provided us with an abundance of useful data and user feedback, and we have identified several areas where we can continue to improve our methods and software. Several new features have been planned for the software, and it will be developed, improved and maintained for the foreseeable future.

Conclusions: SWIFT-Active Screener is a valuable tool that reduces the human screening burden and increases the pace at which reviews are conducted. On average, users can save in excess of 50% of screening effort normally required, resulting in significant time and cost savings.

QUALITATIVE EVIDENCE SYNTHESIS

A new tool to assess the trustworthiness of evidence-based point-of-care information for healthcare professionals

Lenaerts G[1], Bekkering G[2], Goossens M[1], Deconinck L[1], Delvaux N[3], Cordyn S[4], Adriaenssens J[5], Vankrunkelsven P[2]

[1] Belgian Centre for Evidence Based Medicine (CEBAM), Belgium; [2] Belgian Centre for Evidence Based Medicine (CEBAM); Cochrane Belgium, Belgium; [3] Department of Public Health and Primary Care, KU Leuven, Belgium; [4] Federation of the White and Yellow Cross of Flanders, Belgium; [5] Belgian Health Care Knowledge Centre, Belgium

Background: User-friendly information at the point-of-care (POC) should be well-structured, rapidly accessible and comprehensive. The reliability of information and the associated methodological process must be clear. There is no standard tool to evaluate the trustworthiness of such POC information.

Objectives: To develop and validate a new tool to assess the trustworthiness of evidence-based POC information.

Methods: We performed a systematic review to analyze the available tools. None of the tools was assessed for both reliability and validity. Therefore, we decided to design a new tool based on the results of the systematic review. We listed the different criteria important for assessment of trustworthiness of POC information. A working group of healthcare professionals and methodologists defined the criteria in several re-iterations. Subsequently, two researchers piloted the usability of the criteria on existing information sources. Then, all criteria were subject to content validation with a Delphi study. We invited an international panel of 10 experts to rate their agreement with the relevance and wording of the different criteria and to give feedback in case of non-agreement. The process was in writing and anonymous. Consensus was reached when 70% of the experts agreed. When no consensus was reached, we reformulated the criteria based on the experts' comments for a next round of the Delphi study. We repeated this process until consensus was reached for each criterion. Our next step is to test the inter-rater reliability of the final tool.

Results: The new tool was designed with nine certification criteria. After the first round of the Delphi study, the nine certification criteria were assessed as relevant by the experts, but refinement of the wording was needed. Some criteria had to be split up, resulting in a final version with 11 certification criteria. The certification criteria relate to authorship, literature search, use of pre-appraised evidence, critical appraisal of evidence, expert opinions, peer review, timeliness and updating, conflict of interests, and commercial support.

Conclusions: We developed and validated a new tool to assess the trustworthiness of evidence-based POC information for healthcare professionals. The next steps are to test its reliability. The systematic use of this tool will enhance the quality of POC information and will support healthcare professionals to practice evidence-based medicine.

Reference:

1. Lenaerts G, Bekkering GE, Goossens M, De Coninck L, Delvaux N, Cordyn S, et al. Tools to Assess the Trustworthiness of Evidence-Based Point-of-Care Information for Health Care Professionals: Systematic Review. *Journal of Medical Internet Research* 2020;22(1):e15415. doi:10.2196/15415

Assessing conceptual richness in a meta-ethnography: an example from a qualitative evidence synthesis on factors influencing acceptance of childhood vaccination

Cooper S[1], Schmidt B[1], Swartz A[2], Colvin C[2], Leon N[3], Wiysonge CS[1]

[1] Cochrane South Africa, South African Medical Research Council, Cape Town, South Africa; [2] School of Public Health and Family Medicine, University of Cape Town, Cape Town, South Africa; [3] Health Systems Research Unit, South African Medical Research Council, Cape Town, South Africa

Background: There is growing recognition amongst qualitative evidence synthesis (QES) experts that meta-ethnography requires conceptually rich studies, as more descriptive studies usually have insufficient depth for an interpretive synthesis. However, understandings of 'conceptual richness' are diverse and often vague, and there is currently no established method for its assessment. We therefore developed and applied an approach for assessing conceptual richness for a Cochrane QES on acceptance of childhood vaccination.

Method: Our approach drew on Sandelowski and Barroso's (2007) typology of the type and nature of qualitative findings. This typology conceives qualitative findings as being located along a spectrum based on the degree of data abstraction or 'transformation'. At one end of the spectrum are more descriptive findings that describe patterns in the data. At the more transformative end of the spectrum are interpretive or explanatory findings. These findings have a high-level of abstraction and provide theoretical interpretations or explanations, often across multiple patterns within the data. We created a 5-point scale to categorise studies on this spectrum and developed clear definitions for each score. We agreed that studies with a score of ≥ 3 would be included in our QES.

Results: A total of 136 studies (165 articles) met the inclusion criteria for our QES and were assessed for conceptual richness using our 5-point scale. Initially, 25 random studies were scored independently by two investigators who were both experienced qualitative researchers. Thereafter, one investigator performed the assessment on the remaining eligible studies, a sample of which was checked by a second investigator. Both investigators experienced challenges with the process, and comparison of their assessments revealed some variation. On reflection, several strategies may have helped reduce these challenges and variations. These include identifying a 'prototype' study for each score before beginning the assessment process to serve as a reference point; reducing the scale to 3-points; refining our definition of conceptual richness to incorporate relevance to the synthesis objectives and to more clearly distinguish it from methodological quality; and developing a more systematic approach for gauging overall ranking for studies with multiple articles of varying richness. The value of having two (or more) investigators may be less about achieving inter-rater reliability and more about facilitating the making, challenging and articulating of what are often far from straightforward judgements.

Conclusion: While assessments of conceptual richness are partly intuitive and subjective, there is a need for greater debate and transparency regarding how we define and judge richness. Our method provides one potentially fruitful approach which could serve as a foundation for further work and practical application.

Patient or healthcare consumer involvement: None

Does qualitative data on a research topic become saturated? Empirical observations from adherence to antiretroviral drugs in Africa

Rohwer A[1], Hendricks L[1], Oliver S[2], Garner P[3]

[1] Centre for Evidence-based Health Care, Faculty of Medicine and Health Sciences, Stellenbosch University, South Africa; [2] UCL Institute of Education, University College London, UK; [3] Centre for Evidence Synthesis in Global Health, Liverpool School of Tropical Medicine, UK

Background: We have published a qualitative evidence synthesis (QES) on adherence to anti-retroviral treatment (ART) in people living with HIV in Africa with a search date of 2016. Just prior to its publication, we carried out an additional search and conducted a rapid appraisal of the new qualitative literature, and decided no new themes had emerged. This made us wonder whether there were circumstances in qualitative research synthesis where there is sufficient knowledge to render additional research on the question redundant.

Objectives: To explore the influence of new research on the findings of an existing systematic review. To do this we evaluate data saturation by means of assessing whether the literature identifies new themes, enriches existing themes, or leads to their modification in the published review.

Methods: We used the same search strategy and eligibility criteria as for the existing review. Two authors independently selected studies for inclusion; assessed quality and coded included studies. We coded studies deductively, using the list of codes generated in the existing review and adding to this list in case new codes emerged. New codes were flagged. We had regular meetings to discuss emerging codes and themes. During this iterative process, we considered whether new codes 1) fit into the existing themes and subthemes of the framework; 2) enriched the current themes with new subthemes; or 3) added new themes to the existing framework.

Results: Our search identified 3947 new citations between December 2016 and November 2019. After removal of duplicates, we screened 3830 titles and abstracts, and 307 full texts. We piloted our approach to extract and analyse data on 8 included studies. We are at an early stage of the analysis, but in the few studies examined, we have identified some codes that added to our existing themes, but did not identify any new themes. We will present comprehensive results at the Colloquium.

Conclusions: In our case study, we examine more closely whether, in the rapidly moving policy area of HIV treatment and adherence in Africa, additional qualitative primary research on barriers and facilitators of adherence to ART is justified. This study will help reveal whether data saturation in QES is a true phenomenon and thus further primary studies not justified, or whether we can recognise attributes of a study that could justify updating a QES.

Patient or healthcare consumer involvement: Patients and healthcare consumers were not involved in the conduct of this methodological work. However, we have a multi-disciplinary team with backgrounds in epidemiology, infectious disease, nursing, and social science that have provided insights and expertise into this process.

Factors influencing the implementation of mental health recovery into services: A mixed studies systematic review

Piat M[1], Wainwright M[2], Sofouli E[1], Vachon B[3], Deslauriers T[3], Prefontaine C[3]

[1] McGill University, Canada; [2] Douglas Mental Health University Institute, Canada; [3] University of Montreal, Canada

Background: Recovery is a worldwide paradigm in mental health. Emerging from the consumer survivor movement personal recovery is described as "a way of living a satisfying, hopeful, contributing life, despite the psychiatry disability or symptoms". While traditional services focus on professional control, patient dependency, self-stigma and hopelessness, the focus of recovery-oriented services is on client empowerment, collaborative professional/client relationships, and community integration.

Objectives: We conducted a mixed studies review on the operationalization of recovery into services for adults. The review questions were: How has mental health recovery been implemented into services for adults, and what factors influence the implementation of recovery-oriented services?

Methods: We searched Ovid- MEDLINE, Ovid-Embase, EBSCO-CINAHL Plus with Full Text, ProQuest Dissertations and Theses, the Cochrane Library, and Scopus from 1998 to July 2018. We included peer-reviewed studies on implementation process, factors and experience when implementing new efforts to transform services for adults with mental illness towards a recovery-orientation. All studies were independently screened over two stages for inclusion by two review authors using Distiller SR software. We applied a Best-Fit Framework Synthesis approach to synthesis and used the consolidated framework for implementation research (CFIR). We used the mixed-methods appraisal tool (MMAT) to appraise all included studies. We used both the PRISMA and ENTREQ reporting guidelines. We used NVivo12 for data synthesis.

Results: Of the 70 included studies, 54 were qualitative, 7 mixed-methods, 6 quantitative, 2 (quantitative and qualitative) and 1 RCT. Studies were from: Hong Kong (1), Denmark (1), Japan (1), Republic of Ireland (1), Norway (2), Germany (2), Canada (4), UK (15), USA (19), and Australia (24). Sixty-eight were in English, and two in German. Two sets of findings will be presented: (1) distribution of data extracted across CFIR domains and constructs and (2) conceptualization of studies into similar types of innovations and common issues effecting implementation. Seven recovery innovations: 1) E-Innovations; 2) Family-Focused Innovations; 3) Peer Workers; 4) Personal Recovery Planning; 5) Recovery Colleges; 6) Service Navigation and Co-ordination and 7) Staff Training. Common implementation issues are: flexibility, relationship building, inclusion of lived experience, challenges with medical model, risk management, embedding innovations in wider organization, and early engagement with stakeholders.

Conclusions: To date reviews in mental health recovery have been on conceptualizing personal recovery, measure instruments and intervention effectiveness. This is the first review on the implementation of recovery-oriented services and the factors known to effect implementation, and common factors that influence implementation.

RCT rehabilitation checklist RCTRACK: project for a reporting guideline for RCTs in rehabilitation

Negrini S[1], Arienti C[2], Patrini M[2]

[1] Department of Biomedical, Surgical and Dental Sciences, University of Milan "La Statale", Milan; IRCCS Istituto Ortopedico Galeazzi, Milan, Italy;
[2] IRCCS Fondazione Don Carlo Gnocchi, Milan, Italy

Background: During the first Cochrane Rehabilitation Methodological Meeting, held in Paris in July 2018, a series of methodological problems in rehabilitation research were discussed and the results were published in a special issue of the European Journal of Physical and Rehabilitation Medicine. Afterwards, a scoping review has listed these methodological issues, and the REREP study has shown the very low clinical replicability of RCTs in rehabilitation. This preliminary work highlighted that randomized controlled trials (RCTs) in rehabilitation face specific methodological challenges related to the practice itself that are not faced in the classical reporting checklists. Precise reporting in these domains is essential to allow interpretation and quality evaluation of published work.

Objectives: To develop a checklist of items to be followed in the reporting of RCTs in rehabilitation.

Methods: The RCT Rehabilitation Checklist (RCTRACK) project has been developed and deposited in the EQUATOR Network. Working areas (WAs) have been identified at the launch meeting and studied with systematic or scoping reviews to identify methodological needs. WAs included: PICOs elements, blinding, statistical analysis and appropriate randomization, attrition, follow up and protocol deviation, research question and study design. During a Consensus Conference held in Orlando in March 2020 we discussed the results of the WAs work and proposed the first draft version of RCTRACK. This will be submitted to a series of Delphi Rounds involving all rehabilitation journals' Editorial Boards, authors of RCTs and methodologists/epidemiologists of the area to achieve the final checklist version by the end of 2020.

Results: The preliminary results of systematic reviews and scoping reviews performed by each WAs highlighted a series of items to be added to current CONSORT checklists. These rehabilitation needs relate to the items objective, participants, interventions, and outcomes, but also statistical analysis.

Conclusions: The RCTRACK checklist will include a set of items that directly address the methodological issues of rehabilitation research. Further, it will be a useful educational tool for authors, reviewers and clinicians to improve the quality of evidence in this specific field.

Patient or healthcare consumer involvement: Not applicable.

Using a matrix analysis to integrate findings from a qualitative evidence synthesis and a review of effectiveness: a worked example

Ames H[1], Glenton C[1], Leon N[2]

[1] The Norwegian Institute of Public Health, Norway; [2] South African Medical Research Council, South Africa

Background: We conducted a qualitative evidence synthesis exploring clients' perceptions and experiences of targeted digital communication via mobile devices for reproductive, maternal, newborn, child, and adolescent health. As part of the process we had to integrate the results from our synthesis with the related intervention reviews. We used a matrix analysis to see whether implementation barriers identified in our synthesis had been addressed in the trials included in the related Cochrane Intervention (effectiveness) reviews.

Objective: To present how we conducted our matrix analysis.

Method: To create the matrix we did the following.

1) We selected the synthesis findings that we had assessed as having high or moderate confidence and that presented potential barriers to the implementation of targeted client communication programmes.
2) We created 10 questions reflecting these barriers, and created a table.
3) We assessed whether the trials included in the two related intervention reviews reported any attempt to address these implementation barriers.

To perform this assessment, we examined the publications included in the intervention reviews. We also performed a search for additional publications that could be related to the trials. We did this by 1) examining the reference lists of the main trial publication; and 2) searching for each trial in PubMed, and doing an advanced search for 'Similar articles'. This search used the trial's first author to identify possible related studies that had this author as a co-author, and selected any that appeared to be related to the trial.

Findings: Our matrix (table 1) suggests that important implementation barriers identified through the qualitative research are ignored in many trials. This may be due to poor reporting by trial authors. We were initially going to base our matrix analysis only on the publications included in the intervention reviews. However, we realized that more information could be contained in protocols, process evaluations and/or related qualitative studies. This process was time-consuming and ended in a large matrix with 125 publications in 68 trials. However, by conducting this additional search we felt more confident in the results.

Conclusion: The matrix allowed us to identify potentially important problems in the designs of the interventions assessed in trials. The use of this approach can help explore reasons for intervention effectiveness. By searching beyond the primary studies included in the related intervention reviews we feel that the analysis was more complete. Integration of qualitative review findings with effectiveness reviews is time intensive and requires good planning and co-ordination of both review processes.

Patient or healthcare consumer involvement: This abstract discusses a methodology that aims to include rich data in qualitative reviews which can provide a more detailed presentation of health consumers' thoughts, opinions and experiences allowing authors to better interpret the meaning and context of findings presented in the primary studies.

Additional files: Table 1 Matrix analysis

Why systematic review production and update processes are resource intensive: a phenomenological qualitative study protocol

Ellen M[1], Sfetcu R[2], Baladia E[3], Spijker R[4], Nußbaumer-Streit B[5]

[1] Department of Health Systems Management, Guilford Glazer Faculty of Business and Management and Faculty of Health Sciences, Ben-Gurion University of the Negev, Israel; [2] National School of Public Health, Management and Professional Development, Romania; [3] Red de Nutrición Basada en la Evidencia (RED-NuBE), Academia Española de Nutrición y Dietética, Spain; [4] Cochrane Netherlands, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, The Netherlands; [5] Cochrane Austria, Department for Evidence-Based Medicine and Evaluation, Danube University Krems, Austria

Background: Systematic reviews are labour intensive and time consuming. Previous research describes the large variation of resources needed to conduct a good systematic review and the influence of various factors on the quality and time frame to complete the review. The diversity of available methods and tools and the exponentially increasing number of systematic reviewers will likely result in various research practices, with some being more efficient than others. To the best of our knowledge, a qualitative investigation of systematic review practices, as well as the perceived areas where a gain in effectiveness can be achieved, has not been conducted to date.

Objectives: To understand why some steps in the systematic review production and update processes are perceived as resource intensive.

Methods: We will conduct in-depth, semi-structured interviews with experts who have actively contributed to the production or update of systematic reviews on health-related topics including clinical, health services, public health and health policy research. The focus of the interviews will be to explore which steps in the systematic review production and update process are resource intensive and participants' perceptions of potential methods and technologies to prioritize and expedite elements of the process. We will aim to conduct approximately 20 interviews and they will be audio-recorded, transcribed, coded and thematically analysed using a deductive approach, guided by the 16 steps of a systematic review process. Ethics approval will be obtained.

Discussion: The results of our study will provide an overview of factors influencing resource intensity of different steps in the systematic review production and update process. Results of this project will feed into a Delphi study that aims to prioritize areas in the systematic review process and methods that are most relevant and promising for expediting the review process. This should guide future methods improvement and validity studies in this area and ultimately help accelerate systematic review production without compromising quality. We anticipate being able to identify functional insights about critical bottlenecks in conducting time-efficient and academically relevant systematic reviews. This qualitative evaluation of systematic review research efforts and challenges can increase the dissemination of high-quality health-related research evidence.

Patient or healthcare consumer involvement: As this is a methodological study, patients and healthcare consumers will not be involved directly. Participants will represent a wide range of stakeholders in the field of systematic review production. With the interview group we aim to create a diverse sample with respect to geographic diversity, experience, content area and types of reviews conducted that will allow us to thoroughly investigate all angles of the phenomenon.

NETWORK META-ANALYSIS

A Bayesian approach to detect outliers in network meta-analysis

Metelli S[1], Mavridis D[2], Chaimani A[1]

[1] Inserm Research Center in Epidemiology and Statistics (CRESS-U1153), Université de Paris, France; [2] Department of Primary Education, University of Ioannina, Greece

Background: Network meta-analysis (NMA) is rapidly reaching the forefront of healthcare research. To avoid misleading conclusions and provide valuable information for clinical decision making, NMAs need to be assessed thoroughly against potential sources of bias. A potentially major threat of the validity of results from NMA are studies with markedly different or extreme effect sizes, namely outlying studies. Such studies could substantially influence and bias the conclusions of the NMAs and thus need proper investigation. Whilst several outlier detection methods have been developed for standard meta-analysis, little work has been done in the field of NMA.

Objectives: To propose an intuitive Bayesian model that captures deviating studies within a network of interventions and to explore the influence of such studies in the NMA results under different modelling scenarios.

Methods: We define outliers as studies with 'shifted' effect sizes and based on this, we introduce a Bayesian NMA mean-shifted model, which assumes shifted effects sizes for each study. Then, we use Bayes factors to test whether each study is more compatible to the conventional or to the mean-shifted NMA model. In the latter situation the study is considered as an outlier. Furthermore, detection of outliers is not straightforward when there is a cluster of outliers. To mitigate this issue, we embed the whole procedure in a leave-one-out cross validation scheme where we restrict our search to groups of studies comparing either the same treatments or the same class of treatments.

Results: We explored the performance of our method using simulated networks contaminated by artificial outliers and a real network of 112 randomized controlled trials comparing second-line treatments for advanced non-small cell lung cancer. The method successfully identified existing outliers in the simulated data. In the lung cancer network, we identified one clear and two potential outliers corresponding to a very large and two moderate Bayes factors respectively. The impact and influence of each of these studies has been assessed via contribution matrices and sensitivity analysis. In both cases, results suggest that two out of the three potential outliers are affecting the results, with one study being particularly influential.

Conclusions: Our method offers an effective diagnostic tool for the identification of outlying and influential studies in a network of interventions. Sensitivity analysis is used to exclude outliers and assess result robustness. This has the clear potential to avoid inappropriate NMA conclusions while aiding robust clinical judgments and correct interpretation of results.

Patient or healthcare consumer involvement: None.

Advantages and disadvantages of the GRADE system and CINeMA approach to rating the certainty of network meta-analysis

Pan B[1], Wang Y[2], Wang Q[2], Ge L[2]

[1] Gansu Provincial Hospital, China; [2] Lanzhou University, China

Background: The popularity and influence of network meta-analysis (NMA) has been rapidly growing. Currently, two approaches – GRADE system and CINeMA framework – are available to rate the certainty of NMA. A study applying GRADE to a network meta-analysis of antidepressants showed that application of GRADE highlighted varying evidence certainty, led to more conservative conclusions, and potentially avoided unwarranted strong inferences based on low certainty of evidence. However, the advantages and disadvantages of these two approaches remain unclear.

Objectives: To perform an empirical study to compare the differences between the results of these two approaches in rating the certainty of network meta-analysis.

Methods: We performed a systematic survey of the literature and included a sample of NMA of randomized controlled trials that used GRADE system or/and CINeMA approach to rate the certainty of network meta-analysis. Two review authors independently screened the title and abstract, and further screened the full text to identify eligible studies. Eligible studies have to meet the following criteria: (1) use GRADE system or/and CINeMA approach to rate the certainty of network meta-analysis; (2) provide enough information to re-assess the certainty of network meta-analysis evidence; and (3) provide enough data to re-run the meta-analysis.

We will rate the certainty of evidence using the GRADE system or/and CINeMA platform to compare the differences of the results and the advantages and disadvantages of the two approaches.

Results: This study is ongoing, and results will be presented at Colloquium as available.

Conclusions: This study is ongoing, and results will be presented at Colloquium as available.

Patient or healthcare consumer involvement: Not applicable

# Antiepileptic drug monotherapy for epilepsy: an updated network meta-analysis of individual participant data

Nevitt S[1], Tudur Smith C[1], Marson A[1]

[1] University of Liverpool, UK

Context: Epilepsy is a common neurological condition, in which people experience recurrent, unprovoked seizures and accounts for 1% of the total global burden of disease. Two types of epileptic seizures are studied within this review, focal seizures that start in one area of the brain, and generalized onset tonic-clonic seizures that start in both cerebral hemispheres simultaneously. With effective drug treatment, up to 70% of individuals with active epilepsy have the potential to become seizure free and go into long-term remission of seizures shortly after starting therapy with a single antiepileptic drug (AED) in monotherapy. Currently in the UK, the National Institute for Health and Care Excellence (NICE) guidelines for adults and children recommend carbamazepine or lamotrigine as the first treatment options for individuals with newly diagnosed focal seizures and sodium valproate for individuals with newly diagnosed generalized tonic-clonic seizures. However, a range of other AEDs are available.

Methods: The choice of the first antiepileptic drug for an individual with newly diagnosed seizures is of great importance and should be made after considering high-quality evidence of how effective the drugs are at controlling seizures and whether they are associated with side effects. It is also important that drugs appropriate for different seizure types are compared to each other. An individual participant data network meta-analysis (IPD-NMA) can provide relative estimates of how all relevant AEDs compare to each other, incorporating direct and indirect evidence, while taking account of important time-to-event outcomes and different epileptic seizure types. Our previous Cochrane IPD-NMA published in 2017 considered the time to treatment failure, remission and first seizure of 10 AEDs used as monotherapy in children and adults with focal or generalized onset seizures. Results of this previous review supported current NICE guidelines, and also demonstrated that newer AED levetiracetam may be a good first treatment for focal epilepsy. We will present the results of an update to our previous review, including two new drugs within the IPD-NMA and new studies published since 2017. A network plot of the evidence included in the updated review is provided in Figure 1. Results of the IPD-NMA will inform 66 pairwise comparisons of the 12 AEDs of interest, including 34 pairwise comparisons which have never been made in head-to-head trials. Challenges related to retrieval of IPD from a range of sources will also be presented.

Figure 1. Network plot of 12 anti-epileptic drugs (monotherapy)
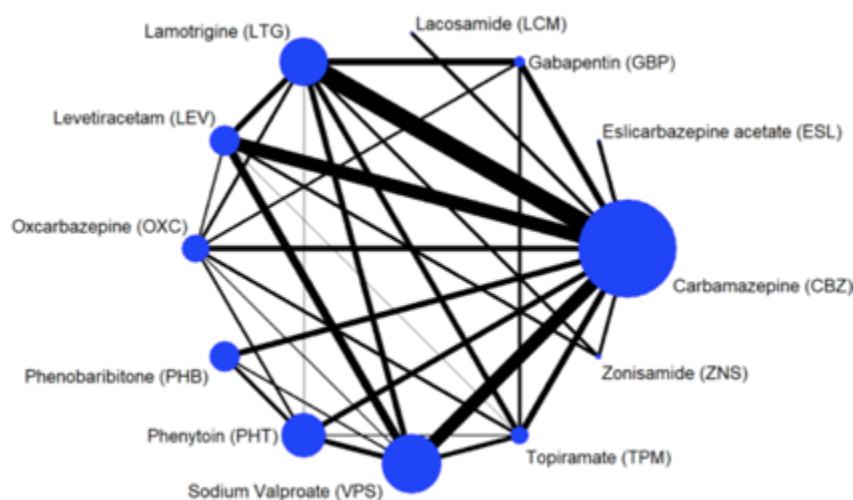
Patient or healthcare consumer involvement: NICE guidelines within the UK are in the process of being updated. The results of our updated review will provide up-to-date and high-quality evidence to directly inform these guidelines and therefore the treatment of individuals with newly diagnosed seizures within the UK. Results of this updated review will also provide wider, up-to-date and high-quality evidence to inform a choice for decision-makers, clinicians or individuals with epilepsy globally.

Artemisinin-based combination therapies for malaria: a systematic review and network meta-analysis

Chu X[1], Zhang N[1], Feng L[1], Yan P[2], Yang K[1]

[1] Lanzhou University, China; [2] Institute of Clinical Research and Evidence-Based Medicine, Gansu Provincial Hospital, China

Background: Malaria is one of the serious global problems of our time. Artemisinin-based combination therapies (ACTs) have contributed substantially to the reduction in the global burden of malaria. However, Artemisinin and partner-drug resistance in *Plasmodium falciparum* are major threats to malaria control and elimination. Recently, the Triple artemisinin-based combination therapies were found to provide effective treatment and delay emergence of antimalarial drug resistance.

Objectives: To assess the current latest evidence on Artemisinin-based combination therapies for malaria as comprehensively as possible.

Methods: We searched the Cochrane Library, Embase, PubMed and Web of Science from the earliest publication date available. We included randomized controlled trials comparing ACTs for malaria. We also screened the reference lists of relevant reviews. No language restrictions were applied. Two authors independently included studies, extracted data, and assessed risk of bias. We used the weighted mean difference (WMD) as the effect size for the continuous variables. We presented summary risk ratios (RR) with 95% confidence intervals (CI) if the results were binary variables.

Results: Final results will be available by the time of the Cochrane Colloquium.

Conclusions: This systematic review and network meta-analysis will provide efficacious, well tolerated, and safe choice of ACTs for malaria. We will present our findings simply and concisely for informed decision-making.

Patient or healthcare consumer involvement: None.

Certainty of evidence in 'Summary of findings' tables from Cochrane network meta-analyses

Saiz LC[1], Leache L[1], Erviti J[1], Gutiérrez-Valencia M[1]

[1] Innovation and Organization Unit, Navarre Health Service, Spain

Background: As per the traditional GRADE recommendations, indirect comparisons should normally imply lowering the quality of the evidence due to indirectness by at least 1 point. Hence, network meta-analyses (NMA) that mix direct and indirect comparisons would be expected to downgrade their scores for that matter but new approaches to better match GRADE to NMA do not consider downgrading indirect evidence by default.

Objectives: To outline the quality of the evidence presented in the 'Summary of findings' tables (SFT) from the published Cochrane NMA. Furthermore, to examine the rationale for lowering the certainty by GRADE domains. To contrast the certainty from the combined (direct and indirect) comparisons with that obtained from the direct comparisons. To assess the degree of agreement among the different NMA when reporting the SFT and how transitivity and incoherence/inconsistency are evaluated.

Methods: We searched the Cochrane Library website using "Network Meta-Analysis" as a MeSH term and within titles, abstracts or keywords on 13 March 2020 to locate all published Cochrane NMA. We collected all the comparisons listed in the SFT of the identified reviews in order to evaluate the implementation of the GRADE criteria for assessing the quality of the evidence.

Results: We retrieved 41 NMA published between January 2016 and March 2020. Fourteen reviews were excluded due to lack of reporting of at least one indirect or combined comparison in SFT format. We evaluated 859 comparisons. The certainty of the NMA evidence was "high" in 11.1% of the combined comparisons, "moderate" in 18.3%, "low" in 21.8% and "very low" in 22.4%. There were no data for 26.4% of all the analyzed comparisons. The reasons for downgrading the certainty of the evidence were: imprecision (49.6%), risk of bias (30.1%), indirectness (10.7%), inconsistency/incoherence (7.5%), others (2.1%). One review did not report any reason for downgrading the certainty regarding combined comparisons. Only in four NMA (14.8%) was quality systematically lessened due to indirectness and none was reduced two levels due to this reason. In seven of the 27 NMA (26%) certainty can be explored separately from direct and combined evidence in the SFT. In sixteen comparisons (11%), certainty of the evidence was graded higher for the NMA with respect to direct evidence. Only in 22 comparisons (15%) was direct evidence graded higher. No explicit assessment of transitivity and incoherence/inconsistency was found in seven (26%) and one (4%) reviews respectively.

Conclusions: In general, quality of the NMA evidence was not routinely downgraded due to indirect comparisons. Indirectness was rarely considered as a reason for lowering the quality of the evidence. There is great heterogeneity among NMA in assessing the certainty of the evidence and also in showing the results. Only a few NMA published results from direct, indirect and combined evidence separately. Transitivity was not thoroughly explored in some reviews.

Patient or healthcare consumer involvement: None.

Clustering physical interventions in network meta-analyses

Ernst M[1], Folkerts A[2], Roheger M[2], Liebermann-Jordanidis H[2], Krohm F[2], Droz M[2], Chakraverty D[2], Adams A[3], Eggers C[4], Monsef I[1], Dresen A[5], Skoetz N[6], Kalbe E[2]

[1] Cochrane Haematology, Department I of Internal Medicine, Center for Integrated Oncology Aachen, Bonn, Cologne, Dusseldorf, Faculty of Medicine and University Hospital, University of Cologne, Cologne, Germany; [2] Medical Psychology | Neuropsychology and Gender Studies and Centre for Neuropsychological Diagnostics and Intervention (CeNDI), Faculty of Medicine and University Hospital, University of Cologne, Cologne, Germany; [3] Institute of Medical Statistics and Computational Biology, Faculty of Medicine and University Hospital, University of Cologne, Cologne, Germany; [4] Clinic for Neurology, University Hospital Giessen and Marburg, Marburg, Germany; [5] Institute of Medical Sociology, Health Services Research and Rehabilitation Science, Faculty of Human Sciences and Faculty of Medicine and University Hospital, University of Cologne, Cologne, Germany; [6] Cochrane Cancer, Department I of Internal Medicine, Center for Integrated Oncology Aachen, Bonn, Cologne, Dusseldorf, Faculty of Medicine and University Hospital, University of Cologne, Cologne, Germany

Background: Non-pharmacological treatment options, especially physical interventions, are known to be effective in the management of movement disorders such as Parkinson's disease, improving motor functioning and quality of life. In this broad field, it is still not clear which types of physical intervention (e.g. gait training, Tai Chi, cycling) are most effective in specific patient groups. Network meta-analyses (NMA) have become a popular method to address this question. However, clustering these interventions when conducting NMA can be challenging due to their high complexity, overlapping intervention components and missing information on the specific intervention content.

Objectives: To present an approach of clustering physical interventions for patients with Parkinson's disease within a NMA that allows the integration of interventions that are highly diverse with respect to the training modality, environment, use of devices, and further features.

Methods: We conducted a systematic search for randomized controlled trials (RCTs) of physical interventions for patients with Parkinson's disease and clustered eligible trials using pre-existing approaches to categorize physical exercise for the elderly. We took an approach that had been developed for falls prevention trials in the elderly (ProFaNE taxonomy, Lamb 2011) as a basis, adapted the original categories and added new categories to integrate eligible interventions that could not be matched clearly to any of the original categories.

Results: The original ProFaNE taxonomy specified five categories of structured exercise: Gait, balance, functional training; strength/resistance; flexibility; three-dimensional (3D) exercise (e.g. Tai Chi, dance); and endurance. We separated the original category 3D into mind-body and dance which we considered distinct interventions, and added water-based training as a third 3D category in order to integrate interventions delivered in an aquatic setting. For the integration of structured physical interventions delivered using a virtual reality (VR) device which was not covered by any of the existing categories, we added the category VR.

Conclusions: Our adaptation of the pre-existing taxonomy allows clustering of a wide range of physical interventions in NMA. Therefore, a more realistic picture of current non-pharmacological physical interventions can be represented in analyses comparing several treatment approaches. Our operationalization of each cluster may help trial investigators describe their interventions more precisely. The adapted system may be used when synthesizing evidence on physical interventions in other diseases.

Patient or healthcare consumer involvement: Within the scope of our overall project investigating physical interventions in patients with Parkinson's disease, we separately conduct focus group discussions with patients and providers of physical interventions to get further insight on the potential and the subjective meaning of physical interventions as a treatment option for Parkinson's disease.

Comparative safety and efficacy of cognitive enhancers for Alzheimer's dementia: an individual patient data network meta-analysis

Veroniki AA[1], Ashoor H[2], Rios P[2], Seitidis G[1], Mavridis D[1], Holroyd-Leduc J[3], Straus S[2], Tricco A[2]

[1] Department of Primary Education, School of Education, University of Ioannina, Ioannina, Greece; [2] Knowledge Translation Program, Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, Canada; [3] Departments of Medicine and Community Health Sciences, University of Calgary, Calgary, Canada

Background: Alzheimer's dementia (AD) is the most common type of dementia. However, it is unclear which cognitive enhancer is optimal for severe AD. Patient-level data from people with AD can be helpful to explore patient-level variation per treatment response. Pooling individual patient data (IPD) from multiple randomized controlled trials (RCTs) of clinical interventions is considered the 'gold standard' analysis.

Objectives: To examine the comparative efficacy and safety of cognitive enhancers by patient characteristics, such as AD severity and sex, and to assess treatment-by-covariate interactions through IPD network meta-analysis (NMA).

Methods: We searched for RCTs with adults comparing cognitive enhancers. The primary outcome was cognition using the Mini-Mental State Examination (MMSE), and the secondary outcome was serious adverse events (SAEs). For eligible RCTs, we requested IPD from authors, sponsors and data sharing platforms. We assessed for consistency between results from published RCTs and provided IPD. We applied an available case analysis for each study, but we plan to explore the impact of missing data through the use of informative missingness parameters in NMA. We captured reasons for missing participants and time to SAE if this was available. We conducted a two-stage analysis: at the first stage we aggregated IPD from included studies to study-level summary; at the second stage the trial parameter estimates were synthesized in a random-effects NMA. We summarized evidence using the odds ratio (OR) and mean difference (MD), respectively. In the main analysis, we used crude ORs and MDs and did not adjust for any patient characteristics. In a further analysis we included patient-level covariates with interaction terms in the model including the patient characteristics that were provided. We combined aggregated data from RCTs for which we were unable to obtain IPD. We performed subgroup and meta-regression NMA for all potential effect modifiers requested from data providers, whenever data were provided.

Results: We included 108 RCTs and received IPD for 17 (16%) RCTs. Of the 17 RCTs, we were able to include 12 RCTs in our NMA with complete data. Access to IPD via proprietary sponsor-specific platforms restricted us from combining IPD in a one-stage NMA model. In most IPD, we encountered a high dropout rate (up to 72%), for which most publications used the last observation carried forward imputation method. We will present NMA results including IPD and/or aggregate data at the Cochrane Colloquium.

Conclusions: An advantage of our IPD-NMA is that we were able to include outcome data, which were not reported in the original publications. Our study will provide insight on personalized medicine for patients with AD.

Patient or healthcare consumer involvement: People with Alzheimer's Dementia require personalized medicine to optimise their healthcare. Evidence from high quality systematic reviews and patient-level network meta-analyses influence patient care since they are used to tailor decision making.

Considering the weight of loops for assessing the certainty of indirect evidence of network meta-analysis

Yao L[1], Guyatt G[1], Brignardello-Petersen R[1]

[1] Department of Health Research Methods, Evidence and Impact, McMaster University, Canada

Background: Indirect evidence needs to be assessed when rating the certainty of evidence of a network meta-analysis (NMA). To increase efficiency, the current GRADE approach focuses the assessment of the indirect evidence on the dominant first order loop. This may lead, however, to ignoring an important proportion of the evidence and to inappropriate ratings of the certainty of evidence.

Objectives: To determine the extent to which current GRADE guidance results in important errors in the assessment of certainty of the evidence.

Methods: We will use a sample of 30 NMAs, using data from previous work conducted in the Department of Health Research Methods, Evidence and Impact at McMaster University and research teams that collaborate with us. Using a contribution matrix, we will estimate the weight of the first order loop and other loops in the body of indirect evidence. For each comparison, we will assess the certainty of indirect evidence 1) using current GRADE guidance, and 2) considering all the loops that contribute at least 10% of the indirect evidence, until we have accounted for at least 90% of all the indirect evidence.

Results: This is work in progress. We will present the main characteristics and the weight of the dominant first order loop of included NMAs. We will summarize the proportion of comparisons in which the certainty of the evidence differs when using current GRADE guidance versus when considering other loops in addition to the first order loop, and characterize the nature of the discrepancies. We will explore if discrepancies are related to the contribution of the first order loop to the indirect evidence.

Conclusions: The results of this study will inform if current GRADE guidance should be revised.

Cumulative network meta-analyses, practice guidelines, and actual prescriptions for post-menopausal osteoporosis: a meta-epidemiological study

Kataoka Y[1], Luo Y[2], Chaimani A[3], Onishi A[4], Kimachi M[5], Tsujimoto Y[6], Murad MH[7], Li T[8], Cipriani A[9], Furukawa TA[2]

[1] Hospital Care Research Unit, Hyogo Prefectural Amagasaki General Medical Center, Japan; [2] Department of Health Promotion and Human Behavior, School of Public Health in the Graduate School of Medicine, Kyoto University, Japan; [3] Paris Descartes University; Inserm, UMR1153 Epidemiology and Statistics, Sorbonne Paris Cité Research Center (CRESS), METHODS Team; Cochrane France, France; [4] Department of Rheumatology and Clinical Immunology, Kobe University Graduate School of Medicine, Japan; [5] Department of Healthcare Epidemiology, School of Public Health in the Graduate School of Medicine, Kyoto University, Japan; [6] Department of Nephrology and Dialysis, Kyoritsu Hospital, Japan; [7] Mayo Clinic–Preventive Medicine, USA; [8] Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, USA; [9] Department of Psychiatry, University of Oxford, UK

Background: The evidence practice gap, or the delay in incorporating research results into practices, is receiving the increasing attention of clinicians and consumers alike. One of its many possible causes is a time delay between the publication of individual study results and the recommendations provided in clinical practice guidelines. Another cause may be a time delay between evidence, either as individual studies or as guideline recommendations, and the real-work prescriptions by the clinicians.

Objectives: To compare the results of cumulative network meta-analyses (NMAs) with the recommendations in post-menopausal osteoporosis practice guidelines and actual prescribing practices in the US.

Methods: We searched MEDLINE, Embase, the Cochrane Central Register of Controlled Trials (CENTRAL), Web of Science, and Scopus to retrieve randomized controlled trials (RCTs) in July 2017. We searched the Agency for Healthcare Research and Quality's National Guideline Clearinghouse and associated society websites to retrieve guidelines in June 2018. We used the Medical Expenditure Panel Survey (MEPS) to analyze prescription data from 1996 to 2015. Two independent investigators selected eligible RCTs. One investigator selected potential eligible guidelines, which were confirmed by another investigator. Two independent investigators extracted data from included RCTs. One investigator extracted recommendations from guidelines, which were confirmed by another investigator. (Registration: UMIN000031894)

Results: We analyzed data from 1995, 2000, 2005, 2010, and 2015. We chose hip fracture as the primary outcome of cumulative NMAs. We included 51 trials, 17 guidelines, and 5099 post-menopausal osteoporosis patients from the MEPS. Bisphosphonate, including alendronate, and combination of vitamin D and calcium (vDCa) were consistently recommendable from an efficacy viewpoint in NMAs and recommended in guidelines. Alendronate was the most prescribed drug (more than 30% over the observation period); however, vDCa was seldom prescribed. The maximum proportion was 5.3% from 2011 to 2015.

Conclusions: In postmenopausal osteoporosis, there was no apparent discrepancy between guideline recommendations and drug prescribing rankings, with the exception of vDCa, when we used cumulative NMAs as references.

Patient or healthcare consumer involvement: No.

Additional file: tables

Dealing with retrieval bias for an evidence-informed individual patient data network meta-analysis

Veroniki AA[1], Ashoor H[2], Rios P[2], Seitidis G[1], Mavridis D[1], Straus S[2], Tricco A[2]

[1] Department of Primary Education, School of Education, University of Ioannina, Ioannina, Greece; [2] Knowledge Translation Program, Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, Canada

Background: The synthesis of individual patient data (IPD) from randomized controlled trials (RCTs) can strengthen evidence used for decision-making. Network meta-analyses (NMA) modelling IPD usually includes non-sponsored or publicly sponsored RCTs. Evidence suggests that IPD sharing may depend on study characteristics, such as funding type, RCT size, RCT risk of bias, and treatment effect. However, retrieval bias in IPD-NMA of sponsored RCTs has not been assessed before.

Objectives: To explore retrieval bias in IPD-NMAs of sponsored RCTs and address challenges and barriers.

Methods: We contacted authors and sponsors of RCTs eligible for 2 IPD-NMAs to obtain IPD. If a study had multiple sponsors, we contacted all of them. To facilitate IPD retrieval, we contacted data sharing platforms. All IPD were checked for consistency with results from published RCTs. We explored whether IPD studies suggested different findings with those of studies not sharing IPD and outlined the IPD availability from sponsors. We noted all barriers and resource requirements associated with the IPD acquisition during the author and sponsor contact processes.

Results: We included 137 RCTs and received IPD for 29 (21%) RCTs (1058 total waiting days). None of the 137 authors shared their IPD. Instead, 17 sponsors for 107 studies were contacted and 7 (41%) sponsors shared their data through proprietary sponsor-specific platforms. The 7 sponsors held data for 94 RCTs and we obtained data from 31% (29/94) of these RCTs. Of the 29 RCTs, we were able to include 23 RCTs in our NMA due to incompleteness of provided data. For example, a study included only IPD for the placebo arm and thus was excluded from the NMA. A big challenge in the IPD was the high dropout rate (up to 72%) from the RCTs, for which many original authors applied inappropriate imputation methods. Hence, our findings differed from published RCT results. We also encountered outcome reporting bias; specifically, some outcomes were missing from the publications but were available as IPD. The use of different platforms restricted us from combining IPD in a single NMA model and a one-stage analysis was impossible. Time restriction to remote-access platforms and frequent changes of these platforms added another challenge to the analysis, given that IPD from different RCTs were available at different time points.

Conclusions: Retrieval bias can severely impact NMA findings and decision-making. Our study highlighted challenges encountered during an IPD-NMA of sponsored RCTs.

Patient or healthcare consumer involvement: Personalized medicine is required to optimize health care. Well-conducted meta-analyses of IPD are considered the 'gold-standard' and influence patient care since patient-level data can be provided to facilitate tailored decision making. However, results from meta-analyses of IPD are likely subject to retrieval bias and awareness of these limitations and their potential impact on findings is required.

Effectiveness of professional interventions to improve antibiotic prescription: a network meta-analysis

Ahmadi A[1], Mallah N[2], Takkouche EB[3]

[1] Knowledge Utilization Research Center; Tehran University of Medical Sciences, Iran; [2] Department of Psychiatry, Radiology and Public Health; University of Santiago de Compostela, Spain; [3] Department of Preventive Medicine, University of Santiago de Compostela, Spain

Background: There are several interventions to promote the 'rational prescription of antibiotics' in physicians. For most of these interventions, there is no enough evidence for the head to head comparison.

Objectives: To compare effectiveness of professional interventions to promote antibiotic prescription in physician by using network-meta analysis

Methods: Search strategy: ('Anti-Infective Agents'(related terms/OR)) AND ('physicians'(related terms/OR)) AND ('prescription'(related terms/OR)).Types of study: Randomized controlled trials. Population: Physicians and Dentists. Intervention: Groups of interventions based on EPOC categorization of professional interventions. Primary outcome: indicators of change in physician' prescriptions.

Study selection and data extraction: We screened titles and abstracts against the inclusion criteria. We appraised selected studies using the Cochrane risk-of-bias tool. Two review authors independently extracted data. Disagreements that arose were resolved through discussion and, if deemed necessary, by referral to a third review author. Data analysis: we performed a random-effects NMA within a Bayesian framework. The routine care/no intervention arm was considered as common comparator. We calculated the posterior Relative Risk (RR) and Confidence Intervals (CI) and the posterior mean ranks for each arm. We used WinBUGS v1.4.3 for the analyses.

Results: We identified 994 unique records, of which 30 studies were eligible. Seven studies contributed in quantitative analysis. There were four distinct types of intervention components: 1) distribution of educational material, 2) audit and feedback, 3) reminders, 4) provider incentives. Various combinations of these components were applied in five arms in the seven included studies: A: no intervention or routine care; B: distribution of educational material and audit and feedback; C: distribution of educational material and audit and feedback and reminders and provider incentives; D: distribution of educational material and audit and feedback and reminders and provider incentives and educational meetings; and E: Distribution of educational material.

The relative risk to increase the change of the behavior in physicians for arms B, C, D and E in comparison with the group A were (RR 3.29; CI:2.09 to 7.56), (RR 1.54; CI -0.15 to 4.17), (RR 1.64; CI:-0.11 to 4.20) and (RR 2.05; CI:-0.03 to 5.37) respectively. The highest mean score of relative rank for effectiveness was for arm B:4.88 and then E:3.51, D:2.56, C:2.52, A:1.46. The surface under the cumulative ranking curve (SUCRA) is depicted for interventions.

Conclusions: The intervention combination of educational material and audit and feedback shows the highest probability of being the most effective intervention to improve antibiotic prescription in physicians.

Patient or healthcare consumer involvement: To make our evidence more applicable, the research idea was discussed with infectious specialist physicians. 'Consumer-informed' material has been produced to improve physicians and policy makers knowledge about effective interventions.

Additional files: figure 1; figure 2; figure 3

Empirical evaluation of ranking metrics in network meta-analysis

Chiocchia V[1], Nikolakopoulou A[1], Papakonstantinou T[1], Egger M[1], Salanti G[1]

[1] Institute of Social and Preventive Medicine, University of Bern, Switzerland

Background: Network meta-analysis (NMA) can produce ranking metrics that lead to a hierarchy of medical interventions from the most to the least preferable. Existing ranking metrics can be non-probabilistic, such as the estimated relative treatment effect, or probabilistic, where probabilities are derived using the distribution of the relative treatment effects. Probabilistic ranking metrics include the probability of each treatment ranking first, second, third, etc., the mean rank, the median rank, and the surface under cumulative ranking curve (SUCRA) or its frequentist equivalent, the P-score. A specific definition of the best treatment leads to a distinctive treatment hierarchy problem and can be addressed with a different ranking metric.

Objectives: To empirically evaluate the level of agreement between treatment hierarches produced by different ranking metrics.

Methods: We re-analysed 232 networks of four or more interventions from randomized controlled trials, published between 1999 and 2015. We produced treatment hierarchies using the following ranking metrics: the probability of producing the best value (pBV), the surface under the cumulative ranking curve (SUCRA) from both a frequentist and a Bayesian framework, and the relative treatment effects using an alternative parametrization of the network meta-analysis model that estimates relative treatment effects against a fictional treatment of average performance. To estimate the level of agreement between treatment hierarchies we used Spearman's ρ, Kendall's τ correlation, and the Yilmaz τ and Average Overlap to give more weight to agreement on higher ranks. Finally, we assessed how the amount of the information present in a network affects the agreement between treatment hierarchies, using the average variance, the relative range of variance, and the total sample size over the number of interventions of a network.

Results: Overall, the pairwise agreement was high for all treatment hierarchies obtained by the different ranking metrics (Table 1). The highest agreement was observed between SUCRA and the relative treatment effect for both correlation and top-weighted measures whose medians were all equal to one. The agreement between rankings decreased for networks with less precise estimates and the hierarchies obtained from pBV appeared to be the most sensitive to large differences in the variance estimates. However, such large differences were rare in practice.

Table 1. Pairwise agreement between treatment hierarchies obtained from the different ranking metrics measured by Spearman ρ, Kendall τ, Yilmaz τ_{AP} and Average Overlap.

	p_{BV} vs $SUCRA_F$	$SUCRA_F$ vs relative treatment effect	p_{BV} vs relative treatment effect	$SUCRA_F$ vs $SUCRA_B$
Spearman ρ	0.9 (0.8, 0.96)	1 (0.99, 1)	0.9 (0.8, 0.97)	1 (0.98, 1)
Kendall τ	0.8 (0.67, 0.91)	1 (0.95, 1)	0.8 (0.69, 0.91)	1 (0.93, 1)
Yilmaz τ_{AP}	0.78 (0.6, 0.9)	1 (0.93, 1)	0.79 (0.65, 0.9)	1 (0.93, 1)
Average Overlap	0.85 (0.72, 0.96)	1 (0.91, 1)	0.88 (0.79, 1)	1 (0.94, 1)

Medians, 1ˢᵗ and 3ʳᵈ quartiles are reported. p_{BV}: probability of producing the best value; $SUCRA_F$: surface under the cumulative ranking curve (calculated in frequentist setting); $SUCRA_B$: surface under the cumulative ranking curve (calculated in Bayesian setting); relative treatment effect stands for the relative treatment effect against fictional treatment of average performance.

Conclusions: Different ranking metrics address different treatment hierarchy problems, but they produced similar rankings in the published networks. Therefore, researchers reporting NMA results can use the ranking metric they prefer, unless there are imprecise estimates or large imbalances in the variance estimates. In this case treatment hierarchies based on both probabilistic and non-probabilistic ranking metrics should be presented. This project is funded by the Swiss National Science Foundation under grant agreement No.179158.

Patient or healthcare consumer involvement: Not relevant.

Evaluation of network meta-analysis in clinical practice guidelines for percutaneous coronary intervention: a systematic review

Zhang L[1], Gong X[1], Chen Q[2], Zhang M[1]

[1] West China Hospital of Sichuan University, China; [2] Chengdu Center for Disease Control & Prevention, China

Background: Clinical practice guidelines (CPGs) have an important role in guiding choices among the numerous options in percutaneous coronary interventions (PCIs). Meta-analysis provides much valuable clinical evidence for CPGs but little is known about the influence of network meta-analysis (NMA).

Objectives: To assess whether NMA in clinical practice guidelines for PCI are consistent with current evidence and whether the consistency of the guidelines depends on the quality of guideline development.

Methods: We searched Web of Science, MEDLINE, Embase, and the Cochrane Library from inception to August 2016. Two review authors independently screened citations to identify English-language guidelines on PCI. Review authors assessed whether the guidelines addressed and agreed with conclusions from these NMAs. Two review authors independently rated NMA quality by using AMSTAR (A MeaSurement Tool to Assess systematic Reviews).

Results: Of the 803 screened citations, 29 NMA met the inclusion criteria. Most of the NMAs (86.2%) published from America and Europe. Twelve NMAs (41.4%) were cited by CPGs. Two NMAs included observational studies and the rest included randomized controlled trials. All-cause mortality, myocardial infarction, target vessel revascularization and stent thrombosis were the main outcomes. NMAs with higher quality are much more likely to be cited by CPGs, which were highly consistent with the evidence-based conclusions.

Conclusions: Not all clinical practice guidelines on PCI were consistent with available evidence from NMA. Guidelines judged to be of higher quality contained more recommendations consistent with evidence-based conclusions. The quality of guideline development processes varied substantially.

Patient or healthcare consumer involvement: Clinical practice guidelines are the product of combining different dimensions such as patients, doctors and medical policies.

Framework for evaluating reporting bias in network meta-analysis

Chiocchia V[1], Nikolakopoulou A[1], Mavridis D[2], Egger M[1], Salanti G[1]

[1] Institute of Social and Preventive Medicine, University of Bern, Switzerland; [2] Department of Primary Education, University of Ioannina, Greece

Background: Reporting bias, or "non-reporting bias" as defined in the latest Cochrane Handbook for Systematic Reviews of Interventions (Version 6, 2019), can seriously compromise the results of systematic reviews and meta-analysis and, as a consequence, potentially affect clinical decision-making. Various graphical and statistical methods are available to assess the risk of reporting bias. However, these approaches have mostly been developed for pairwise meta-analysis, making it difficult to assess the impact of reporting bias on the results from network meta-analysis (NMA).

Objectives: To develop a conceptual and methodological framework for evaluating the impact of reporting bias on NMA results.

Methods: The framework combines comparison-adjusted funnel plots, regression techniques, selection models and threshold analysis. We produce comparison-adjusted funnel plots where the direction of potential bias in each comparison is informed by pairwise contour-enhanced funnel plots and regression slopes for small-study effects. The limit meta-analysis model to adjust for small-study effects (Rücker et al, 2011) is extended to multiple treatment comparisons. To explore the impact of publication bias, we use the extension for NMA of the Copas selection model (Mavridis et al, 2014). For comparisons with less than 10 studies a qualitative assessment of the risk of bias is performed following the framework described in Chapter 13 of the Cochrane Handbook. The threshold analysis to assess the sensitivity of treatment recommendations to bias (Philippo et al, 2019) is also applied where, for each relative effect, a threshold is calculated indicating how much the pairwise evidence could change due to bias before a different treatment is favoured. Then, the plausibility of this change is judged qualitatively.

Results and Conclusions: We present the feasibility and applicability of the methods using illustrative examples of previously published NMAs accessed through the nmadb R package (Papakonstantinou, 2019). We plan to implement these strategies in the Confidence in Network Meta-Analysis (CINeMA) framework (Nikolakopoulou et al., 2019) and web-application (https://cinema.ispm.unibe.ch/). This will allow a more systematic evaluation of the reporting bias domain and produce better informed confidence ratings of the NMA findings. This project is funded by the Swiss National Science Foundation under grant agreement No. 179158.

Patient or healthcare consumer involvement: Not relevant

GRADE for network meta-analysis: a new framework for reporting and interpreting results

Martins C[1], Brignardello-Petersen R[2], Firmino R[3], Carrasco-Labra A[4], Ge L[5], Riva J[2], Granville-Garcia AF[6], Colunga-Lozano LE[2], Schünemann H[2]

[1] Federal University of Minas Gerais; McMaster University, Brazil; Canada; [2] McMaster University, Canada; [3] Federal University of Minas Gerais; McMaster University, Brazil; Canada; [4] McMaster University; University of North Carolina at Chapel Hill, Canada; USA; [5] McMaster University; Lanzhou University, Canada; China; [6] State University of Paraíba, Brazil

Background: Ranking probabilities are one statistical step of the network meta-analysis (NMA) that rank treatments from the highest to the lowest probability of being the best treatment. However, if the certainty of the evidence if not considered in the interpretation of the results of the NMA, the inferences can mislead the clinician and the patient in choosing a treatment with high ranking probability but very low or low certainty of evidence.

Objectives: To describe the methodological aspects of assessing the certainty of evidence using the GRADE approach for network meta-analysis (NMA) and to introduce a new framework for presenting and interpreting NMA results.

Methods: The methodology described here was used in a random Bayesian NMA of randomized controlled trials (RCTs) conducted to determine the effect of desensitizing toothpastes on dentin hypersensitivity (DH) (PROSPERO #CRD42018086815). We assessed the certainty of evidence and used a new framework proposed by the GRADE working group to present and interpret NMA. We chose a comparator as a reference category, and we separated other treatments into the following categories: 1) those that were more or less effective against the comparator and 2) those with similar effect against the comparator. We next determined the magnitude of the effects following Cohen's classification. For NMA interpretation, we graded treatments according to certainty of evidence, and we checked consistency with effect estimates and ranking according to the surface under the cumulative ranking curve (SUCRA).

Results: We included 90 trials in our NMA evaluating 16 treatment arms. The SUCRA ranking value suggested that arginine (87.2%) was one of the best treatments for pain relief due to tactile stimulus. However, with low certainty of evidence due to problems in risk of bias and incoherence, we cannot be confident in the final estimate. With the new approach, high-to-moderate treatments were considered effective against the comparator.

Conclusions: This NMA reported a new GRADE framework for presenting and interpreting results. The judgement that places interventions in categories relies primarily on the magnitude of the effect estimates, the certainty of evidence for those estimates, and secondarily their order in the ranking. This approach can avoid misleading inferences based solely on SUCRA ranking.

Patient or healthcare consumer involvement: Our approach demonstrates that the interpretation of data following SUCRA rankings can lead to misleading inferences. Instead, the interpretation based on certainty of evidence and the magnitude of the effect estimates can help the clinician and the patient to a shared decision-making related to the best treatment option for the patient.

Funding: CAPES; CNPq; FAPEMIG.

Methodological review of items for assessing the risk of bias in network meta-analyses provides groundwork for the creation of an extension to the ROBIS tool

Lunny C[1], Andrea T[2], Wright J[1], Whiting P[3]

[1] Cochrane Hypertension Group, University of British Columbia, Canada; [2] University of Toronto, Canada; [3] University of Bristol, UK

Background: To decide the best treatment for a patient, healthcare providers and patients need a synthesis of evidence for all possible treatments for a given condition. Network meta-analysis (NMA) emerged due to the limitations of pairwise meta-analyses to provide comparative effectiveness of multiple treatments for the same condition. Conventional meta-analyses only average the randomized trials comparing two treatments. NMA can help patients and their care providers choose the treatment that is most important to them based on side effects and efficacy of all treatments. Tools are available for most study designs to make quality assessment easier for a knowledge user. For example, ROBIS can be used to assess the risk of bias of systematic reviews (SRs). However, there is currently no risk-of-bias tool for network meta-analyses (NMA). As the main ROBIS domains are applicable to systematic reviews, ROBIS' synthesis domain can be adapted as an extension for assessing NMAs.

Objectives: To conduct a methodological scoping review with the aim to develop a list of items relating to risk of bias in network meta-analyses.

Methods: We searched MEDLINE, Embase, the Cochrane library as well as grey literature databases including the EQUATOR Network, websites of evidence synthesis organizations (Cochrane, the US Institute of Medicine, the Campbell Collaboration, and the Joanna Briggs Institute), as well as methods collections (e.g. Cochrane Methodology Register, Meth4ReSyn library, AHRQ Effective Health Care Program). We included any article describing or reporting items related to risk of bias in NMAs. We also included studies that assessed the methodological quality of NMAs. To identify other potentially relevant studies, we examined the reference lists of included studies and undertook forward citation searches of seminal articles using Google Scholar. Two review authors independently reviewed titles, abstracts and full-text articles. We extracted data on items, criteria or guidance that was potentially relevant to the risk of bias or quality of NMAs of interventions. Sources were ordered and extracted by year of publication, and when a new source was reviewed, items already extracted were revised iteratively or added if they were unique. The final list of items deemed unique (i.e. same conceptual or methodological issue) were retained. The items were categorized based on the synthesis domain in the ROBIS structure.

Results: A list of items was developed and categorized into broad themes based on the ROBIS tool. When items related to the same conceptual or methodological issue they were combined, and other unique concepts were split into separate items. Many items were reworded as signalling questions so that each item is phrased so "yes" is good, and each item only covers a single concept.

Conclusions: This review provides groundwork for the creation of an extension to the ROBIS tool for assessing risk of bias in NMAs. Knowledge users need the highest quality evidence to make decisions about which treatments should be used in healthcare.

Penalized regression in network meta-analysis: a new approach for analyzing networks of interventions with rare events

Evrenoglou T[1], Mavridis D[2], Chaimani A[1]

[1] Inserm Research Centre in Epidemiology and Statistics (CRESS-U1153), Université de Paris, France; [2] Department of Primary Education, University of Ioannina, Greece

Background: Network meta-analysis (NMA) allows comparing simultaneously multiple interventions and, under certain conditions, provides the highest possible level of evidence for the development of clinical guidelines. However, conventional NMA models for dichotomous outcomes can provide biased and imprecise results when the available studies are small or there are few events. Cochrane suggests analyzing at least one efficacy and one safety outcome and the latter typically includes few events. Methods that allow proper analysis of individual studies with low event rates do exist but have never been considered in the context of NMA.

Objectives: To allow more accurate and less biased conclusions from NMAs evaluating rare safety outcomes by extending appropriate statistical methods used in the analysis of individual studies into full networks of interventions.

Methods: We developed a new statistical approach for low event rate binary data forming a full network of interventions. Following ideas previously suggested in the literature for the analysis of a single study, we reduce the bias of NMAs with rare events by modifying the likelihood function of the model. We evaluate the performance of our approach using various simulation scenarios and two real datasets: a network comparing the safety of different drugs for chronic plaque psoriasis and a network comparing interventions for decreasing blood loss and blood transfusion requirements during liver transplantation. To facilitate the use of our method, we have implemented it in R.

Results: In comparison to three alternative NMA models that have been suggested for handling rare events, our method gave on average more precise and less biased relative effect estimates in most simulated scenarios. In the real examples, our model led to much smaller confidence intervals than the other methods particularly for comparisons involving only one or two studies. This is explained partly because the method gives more precise results for trials with small event rates and partly because our approach allows the inclusion of all available studies no matter the number of events per arm. It this way, we also avoid the risk of ending up with disconnected networks due to the exclusion of studies with zero events without making arbitrary continuity corrections.

Conclusions: When we have studies with small event rates we should employ appropriately tailored methods to synthesize them. The suggested methodology offers a reliable and more informative alternative to existing approaches for the analysis of networks of interventions with rare events. Nevertheless, in the presence of small event rates we should always place less confidence in our effect estimates.

Patient or healthcare consumer involvement: None.

Pharmacological interventions for primary insomnia in adults: a systematic review and network meta-analysis

Pan B[1], Yang K[2], Hou L[2], Lai H[2], Wang Q[2], Ge L[2]

[1] Gansu Provincial Hospital, China; [2] Lanzhou University, China

Background: Insomnia has become a major public health problem with significant social burden, which is associated with an increased risk of medical or mental disorders, including injury, non-alcoholic fatty liver disease, hypertension, cardiovascular diseases, dementia, depression, mortality from all-cause and cardiovascular diseases. Cognitive behavioural treatment and pharmacological treatment are available. However, pharmacotherapy is generally prescribed for insomnia in primary care. Different drug classes and individual drugs are recommended by current national and international guidelines, but the recommendations are thus inconclusive. Previous pairwise meta-analyses only compared the effect for one or two of them. And thus it is difficult to generate the clear hierarchies and superiority for the efficacy and safety of available agents.

Objectives: This network meta-analysis aims to compare different pharmacologic treatments for primary insomnia on both broad classes of drugs and individual drugs, and to enable patients and clinicians to make more precise decisions on the best pharmacological treatments for insomnia.

Methods: We searched PubMed, Embase, and Cochrane Central Register of Controlled Trials (CENTRAL) to identify randomized control trials focused on efficacy and safety of the agents for the treatment of primary insomnia. Using standardized study eligibility forms, teams of three review authors independently screened all the retrieved bibliographic records according to eligibility criteria. Data extracted from included studies including general information, drug information, and outcomes of interests (adverse event, sleep onset latency, total sleep time, and wake time after sleep onset,). The methodological quality of included study was assessed using the Cochrane risk-of-bias tool. We use R software to conduct a Bayesian framework random-effects network meta-analysis. We also rated the certainty (quality) of evidence using the GRADE approach.

Results: This study is ongoing and results will be presented at Colloquium as available.

Conclusions: This study is ongoing and results will be presented at Colloquium as available.

Patient or healthcare consumer involvement: Not Applicable.

Use of the GRADE approach to rate the certainty of evidence from network meta-analyses: a cross-sectional survey

Pan B[1], Wang Q[2], Lai H[2], Hou L[2], Wang Y[2], Wang J[1], Ge L[2]

[1] Gansu Provincial Hospital, China; [2] Lanzhou University, China

Background: Network meta-analysis (NMA) could address the comparative effectiveness of multiple interventions by way of combining direct and indirect estimates of effect, and thus is rapidly growing popularity and influence. But application of NMA's results requires understanding the quality of the evidence. In 2014, the Grading of Recommendations Assessments, Development, and Evaluation (GRADE) working group developed a framework to assess the certainty (quality) of the evidence from NMA. The GRADE approach to NMA has been widely used in recent years but the details of its application remain unknown.

Objectives: This study aims to investigate the general characteristics and the usage of GRADE approach of NMA that use GRADE approach to rate the certainty of evidence.

Methods: We conducted a comprehensive search of PubMed and Cochrane library to identified NMA that use GRADE approach to rate the certainty of evidence published from 2014 to 2020 in English. Two review authors independently screened the title and abstract, and further screened the full-text to identify eligible studies. Data extracted from NMA included the first author, publication year, journal of publication, country, institution, article type, research topic. We summarized the above characteristics using descriptive statistics.

Results: This study is ongoing, and results will be presented at Colloquium as available.

Conclusions: This study is ongoing, and results will be presented at Colloquium as available.

Patient or healthcare consumer involvement: Not applicable.

DIAGNOSTIC TEST ACCURACY REVIEW METHODS

Chest computed tomography for the diagnosis of patients with coronavirus disease 2019 (COVID-19): evidence and effectiveness

Lv M[1], Wang M[2], Luo X[1], Estill J[3], Liu Y[1], Ren M[1], Wang X[1], Chen Y[4]

[1] School of Public Health, Lanzhou University, Lanzhou, China; [2] Department of Radiology, The First Affiliated Hospital, Lanzhou University, China; [3] Institute of Global Health, University of Geneva, Switzerland; [4] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China

Background: The outbreak of the coronavirus disease 2019 (COVID-19) has had a massive global impact. As computed tomography (CT) has been widely used in the diagnosis of this novel pneumonia, it is essential to understand the role of CT for the diagnosis and the main imaging manifestations of patients with COVID-19.

Objectives: To estimate the probability of positive findings in the initial CT examination and to investigate the main imaging manifestation in patients with COVID-19.

Methods: We conducted a systematic review and meta-analysis on studies about the use of chest CT for the diagnosis of COVID-19. We comprehensively searched databases and preprint servers on chest CT for patients with COVID-19 between 1 January 2020 and 29 February 2020. The primary outcome is the initial sensitivity CT examination. We also conducted subgroup analyses and evaluated the quality of evidence using the GRADE approach.

Results: We included 56 studies with 3380 patients. A meta-analysis of 37 studies estimated the probability of positive results in the initial CT examination to be 98% (95% confidence interval (CI) 0.94 to 1.00). This probability was slightly higher in Hubei (99%, 95% CI 0.97 to 1.00) than outside of Hubei (95%, 95% CI 0.89 to 0.99). If case reports were excluded, the probability was 94% (95% CI 0.90 to 0.98). The positive rate of initial CT examination in children was only 50% (5% CI 0.07 to 0.94). The most common imaging manifestation was GGO which was found in 63% (95% CI 0.53 to 0.73) of the patients. The pooled probability of bilateral involvement was 84% (95% CI 0.78 to 0.88). The quality of evidence was low across all outcomes.

Conclusions: In conclusion, this meta-analysis indicated that almost all patients with COVID-19 presented abnormal findings in their initial CT scan. Therefore, CT can potentially be used to assist in the diagnosis of COVID-19. However, considering low sensitivity of CT scan in children with COVID-19 and the risk it may pose, we do not recommend it as a routine diagnostic tool for pediatric patients.

Patient or healthcare consumer involvement: None.

Comparative reviews of diagnostic test accuracy in imaging research: evaluation of current practices

Dehmoobad Sharifabadi A[1], Leeflang M[2], Treanor L[1], Kraaijpoel N[3], Salameh JP[4], Alabousi M[5], Asraoui N[1], Choo-Foo J[1], Takwoingi Y[6], Deeks J[6], McInnes M[4]

[1] University of Ottawa, Canada; [2] University Medical Centre (UMC), Amsterdam, The Netherlands; [3] Academic Medical Center, The Netherlands; [4] Ottawa Hospital Research Institute, Canada; [5] McMaster University, Canada; [6] University Hospitals Birmingham NHS Foundation, UK

Background: Research comparing the diagnostic accuracy of two or more imaging tests is essential to determine optimal diagnostic pathways and downstream treatment decisions. High-quality comparative methods would include primary studies that apply the index tests to every study participant or randomly allocate participants to receive one of the index tests. If comparative imaging reviews largely rely on non-comparative primary studies, the conclusions regarding the comparative accuracy of index tests may be biased.

Objectives: To determine the extent to which comparative imaging systematic reviews of diagnostic test accuracy (DTA) use primary studies with comparative or non-comparative designs.

Methods: We used MEDLINE to identify DTA systematic reviews published in imaging journals between January 2000 and May 2018. The systematic reviews compared at least two index tests (one of which was imaging-based). We extracted review characteristics and evaluated study design and other characteristics of primary studies included in the systematic reviews.

Results: We included 103 comparative imaging reviews. Eleven (11%) included only comparative studies, 12 (11%) included only non-comparative primary studies, and 80 (78%) included both comparative and non-comparative primary studies. For reviews containing both comparative and non-comparative primary studies, the median proportion of non-comparative primary studies was 81% (interquartile range (IQR) 57% to 90%). Of 92 reviews that included non-comparative primary studies, 86% did not recognize this as a limitation. Furthermore, among 4182 primary studies, 3438 (82%) were non-comparative and 744 (18%) were comparative in design.

Conclusions: Most primary studies included in comparative imaging reviews are non-comparative in design and awareness of the risk of bias associated with this is low. This may lead to incorrect conclusions about the relative accuracy of diagnostic tests and be counter-productive for informing guidelines and funding decisions about imaging tests.

Patient or healthcare consumer involvement: Medicine today relies on cost effectiveness analyses, by comparing DTA imaging measures as the cornerstone of medical diagnoses. The demand for accurate comparative data combined with minimal awareness of valid comparative study designs may lead to counter-productive research and inadequately supported clinical decisions. Using comparative accuracy imaging reviews with a high risk of bias to inform guidelines and funding decisions may have detrimental impacts on patients.

Additional file: figure 1

Computer-aided diagnosis for prostate cancer based on magnetic resonance imaging: a systematic review with diagnostic meta-analysis

Li M[1], Yao L[2], Yan P[3], Cao L[1], Lu Z[4], Hui X[1], Yang K[4]

[1] Evidence-Based Social Science Research Center, School of Public Health, Lanzhou University, Lanzhou, China; [2] Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, China; [3] Institute of Clinical Research and Evidence-Based Medicine, The Gansu Provincial Hospital, Lanzhou, China; [4] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China

Background: A computer-aided detection (CAD) system for accurate and automated prostate cancer diagnosis has been developed, however, the diagnostic test accuracy of different CAD systems is still controversial.

Objectives: To assess the diagnostic accuracy of CAD systems based on magnetic resonance imaging (MRI) for prostate cancer.

Methods: We searched the Cochrane Library, PubMed, Embase and China Biology Medicine disc until March 2019 for original diagnostic studies. Two independent review authors selected studies on CAD based on MRI diagnosis of prostate cancer and extracted the requisite data. We calculated pooled sensitivity, specificity, and the area under the summary receiver operating characteristic (SROC) curve to estimate the diagnostic accuracy of the CAD system.

Results: We included fifteen studies involving 1945 patients in our analysis. The diagnostic meta-analysis showed that overall sensitivity of CAD system ranged from 0.47 to 1.00 and, specificity from 0.47 to 0.89. The pooled sensitivity of CAD system was 0.87 (95% confidence interval (CI) 0.76 to 0.94), pooled specificity 0.76 (95% CI 0.62 to 0.85), and the area under curve (AUC) 0.89 (95% CI 0.86 to 0.91). Subgroup analysis showed that the support vector machines (SVM) produced the best AUC among the CAD classifiers, with sensitivity ranging from 0.87 to 0.92, and specificity from 0.47 to 0.95. Among different zones of prostate, the CAD system produced a better AUC in the transitional zone than in the peripheral zone and central gland; sensitivity ranged from 0.89 to 1.00, and specificity from 0.38 to 0.85.

Conclusions: CAD system can help improve the diagnostic accuracy of prostate cancer especially using the SVM classifier. Whether the performance of the CAD system depends on the specific locations of the prostate needs further investigation.

Determining the primary endpoints for diagnostic test accuracy reviews of substance use disorders

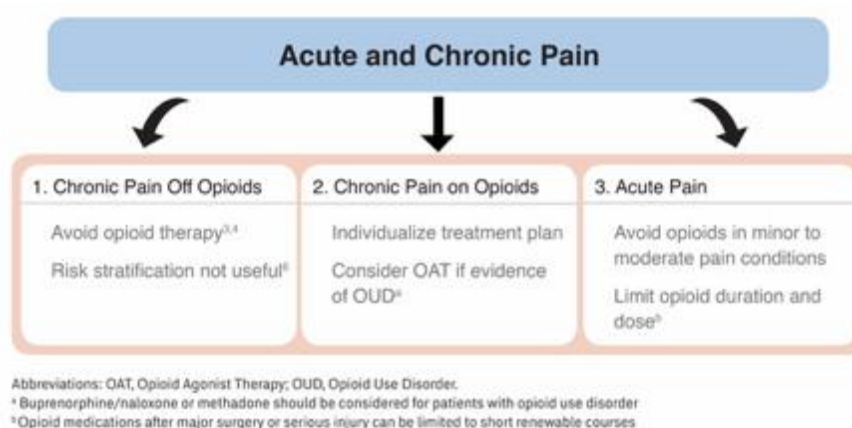Klimas J[1], Dormuth C[2], Wood E[3]

[1] University of British Columbia, British Columbia Centre on Substance Use, Canada; [2] Therapeutics Initiative, University of British Columbia, Canada; [3] University of British Columbia, BC Centre on Substance Use, Canada

Background: Diagnostic test accuracy reviews increasingly are being used in substance use research, yet the majority of the studies considered in these reviews do not account appropriately for the primary endpoints of interest. Recently, the DSM-5 criteria have been largely criticized as poorly able to properly distinguish those with problematic use from those with substance use disorders.

Objective: To review the main implications of adopting alternate primary endpoints in diagnostic test accuracy reviews of substance use disorders and highlight a practical application of appropriate analytical techniques.

Methods: The application of literature-derived primary endpoints is demonstrated through the use of empirical data from two substance use diagnostic test accuracy reviews. We derived primary endpoints from literature in a review of alcohol-withdrawal patient assessment (*JAMA* 320(8):825-33), and patient risks of developing prescription opioid use disorders when initiating opioid analgesics for pain among opioid naïve patients (Figure 1, *JAMA Network Open* 2(5):e193365). We evaluated primary endpoint options in existing substance use studies. This evaluation included surveys of the literature for endpoints and measurement approaches, followed by assessment of endpoint choices against diagnostic test accuracy issues, population characteristics, tests of sensitivity and specificity. We discuss the advantages of this practical approach.

Figure 1. Evidence-based opioid sparing pain management strategy

Abbreviations: OAT, Opioid Agonist Therapy; OUD, Opioid Use Disorder.
a Buprenorphine/naloxone or methadone should be considered for patients with opioid use disorder
b Opioid medications after major surgery or serious injury can be limited to short renewable courses

Conclusions: Inappropriate definitions of primary endpoints in diagnostic test accuracy studies can lead to the presentation of inaccurate results and hence potentially misleading conclusions. We have demonstrated that adjustment for literature-derived endpoints, as opposed to using the criticized DSM-5 criteria, can be useful in these studies and we encourage more judicious use of the established diagnostic categories to enhance accuracy of reviews and meta-analyses.

Patient/healthcare consumer involvement: British Columbia Centre on Substance Use Network of Family Members and Caregivers reviewed the meta-analyses and facilitated findings' dissemination. It is comprised of families who are affected by and want to change the existing substance use system, including representatives from various support and advocacy groups.

Diagnosis and prediction of different machine learning methods for knee osteoarthritis: a systematic review and meta-analysis

Cao L[1], Yao L[2], Li M[3], Yang K[3]

[1] School of Public Health, Lanzhou University, China; [2] Health Research Methodology I Department of Health Research Methods, Evidence and impact, McMaster University, Canada; [3] Evidence Based Social Science Research Center, School of Public Health, Lanzhou University, China

Background: Knee osteoarthritis (OA) is a chronic and progressive joint disease with a high contribution to global disability, mainly in the elderly and particularly in women. The available diagnostic approaches such as x-ray, computed tomography and magnetic resonance imaging have large precision errors and low sensitivity. Machine learning (ML) is the application of probabilistic algorithms to train a computational model to make predictions; it has great potential to become a valuable clinical diagnostic tool.

Objectives: To determine the diagnostic and prediction accuracy of different machine learning methods for knee OA

Methods: We searched four electronic databases from inception to July 2019. We used the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool for risk of bias and applicability assessment. The outcomes we assessed were test characteristics such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC). We used RevMan 5.3 software to pool data and to carry out the meta-analysis where possible.

Results: Based on objective selection, we included six studies that validated the performance of their machine learning method using either a new group of patients or retrospective datasets. We identified six methods for machine learning able to diagnose and predict knee OA. Five studies reported sensitivity and specificity of 73.2% to 94.4% and 73.9% to 100.0%, respectively. Two studies report accuracies of 66.71% and 75%. Three study provides an area under the receiver operating curve (AUC) of 0.81, 0.93 and 0.972. In addition to diagnostic performance, two study also reported sensitivity of 77.97% and 88.9%, specificity of 78% and 82% for prediction of knee OA.

Conclusions: Of the currently included studies, machine-learning algorithms have demonstrated promising results and certainly have the potential to aid radiologists with the detection and screening of knee OA. We should interpret the findings of these reviews with caution, considering the problem of over-fitting in the machine learning method, and large datasets need to be built to verify these findings in the future.

Patient or healthcare consumer involvement: Not applicable

Diagnosis and prediction of different machine learning methods for knee osteoarthritis: a protocol for a systematic review

Cao L[1], Liang Y[2], Yang K[3]

[1] School of Public Health, Evidence-Based Social Science Research Center, Lanzhou University, China; [2] Health Research Methodology I Department of Health Research Methods, Evidence and impact, McMaster University, Canada; [3] Evidence-Based Social Science Research Center, School of Public Health, Lanzhou University, China

Background:

Objectives: To determine the diagnosis and prediction accuracy of different machine learning methods for knee osteoarthritis

Methods: Two review authors systematically searched Cochrane, PubMed, Embase, and Web of Science (last updated in March 2019) for eligible articles. To identify potentially missed publications, we screened manually the reference lists of the final included studies. The outcomes we assessed were test characteristics such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (ROC). We will use the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool to assess the risk of bias and applicability. Two independent review authors will conduct all procedures of study selection, data extraction, and methodological assessment. We will resolve any disagreements through consultation with a third review author. We will use RevMan 5.3 software and Stata V15.0 to pool data and to carry out the meta-analysis if it is possible.

Results: This systematic review will provide a high-quality synthesis of machine learning for diagnosing osteoarthritis of the knee from various evaluation aspects including accuracy, sensitivity, specificity and area under the ROC curve.

Conclusions: The findings of this systematic review will provide the latest evidence of diagnosis and prediction of different machine learning for patients with knee Osteoarthritis.

Patient or healthcare consumer involvement: Not applicable

Diagnostic accuracy of Canadian C-spine rule and NEXUS for detecting clinically important cervical spine injury following blunt trauma

Nitzsche R[1], Saragiotto BT[1], Maher CG[2], Lin CC[3], Verhagen AP[4], Goergen S[5], Michaleff ZA[6]

[1] Masters and Doctoral Programs in Physical Therapy, Universidade Cidade de São Paulo, Brazil; [2] Sydney School of Public Health, The University of Sydney, Australia; [3] Musculoskeletal Health Sydney, Sydney School of Public Health, The University of Sydney, Australia; [4] Department of Physiotherapy, University of Technology Sydney, Australia; [5] Diagnostic Imaging, Monash Medical Centre, Southern Health, Melbourne, Australia; [6] Arthritis Research UK Primary Care Centre, Research Institute for Primary Care & Health Sciences, Keele University, Keele, UK

Background: Cervical spine trauma can lead to clinically important injuries such as fractures, dislocations and ligament instability. It represents approximately 3.5% of trauma occurring in emergency departments around the world. Cervical injuries can have severe consequences such as spinal cord injury and even death. For this reason, the diagnosis must be made early. The diagnosis is usually done by imaging exams such as radiography, nuclear magnetic resonance and computed tomography. The latter is the most used on initial examination in polytrauma patients or those with a high suspicion of injury. Even with a low prevalence of injuries, these tests are performed on between 60% to 90% of stable patients in alert state (Glasgow = 15). In addition, they generate high costs and sometimes unnecessary exposure to radiation. Clinical decision rules can help clinicians to rule out serious injuries without imaging. The "Canadian cervical spine rule" (CCR) and "The National Emergence X-Radiography utilization study" (NEXUS) are the most commonly used clinical decision rules after blunt cervical trauma.

Objectives: To describe and compare the diagnostic accuracy between CCR and NEXUS in patients with suspected cervical injury followed by sudden trauma.

Methods: We searched on MEDLINE, Embase, CINAHL and LILACS. The criteria for considering studies in this review were: being a prospective cohort or cross-sectional study; have a population of adults with suspected cervical injury after sudden trauma; have an analysis of diagnostic accuracy with CCR, NEXUS or both rules; the rules be compared with a standard exam of reference. We used the modified version of the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool to analyse the methodological quality of the studies. The data from each study was used to generate the contingency table and the true positives, true negatives, false positives and false negatives, and to calculate the sensitivity, specificity and 95% confidence interval of each test. To estimate the summary of sensitivity and specificity, we performed a meta-analysis using the bivariate logistics model.

Results: We found 1090 articles, of which 24 were included in this review (Figure 1). Most included studies presented moderate methodological quality. This review is still ongoing and at the moment we have analysed data from 15 studies. The sensitivity of the CCR ranged from 0.90 to 1.00 compared to 0.83 to 1.00 by NEXUS. The specificity of the CCR was between 0.01 to 0.77 compared to 0.02 to 0.46 by NEXUS.

Conclusions: The preliminary results of this review showed that the diagnostic accuracy of CCR appears to be greater compared to NEXUS to assess possible severe cervical injuries in patients following blunt trauma.

Additional files: [Figure 1: flow diagram](#)

Diagnostic value of seven biomarkers for breast cancer: an overview with evidence mapping and indirect comparisons of diagnostic test accuracy

Gao Y[1], Liu M[1], Shi S[1], Zhang J[2], Tian J[1]

[1] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China; [2] Evidence-Based Medicine Center, Tianjin University of Traditional Chinese Medicine, Tianjin, China

Background: Breast cancer is the second most common cause of cancer-related mortality, causing 15.00% of cancer deaths in 2018. In general, early breast cancer has a good prognosis with a five-year survival rate of more than 80.00%. However, in a resource-poor environment, the five-year survival rate for breast cancer is very low, ranging from 10.00% to 40.00%, as most breast cancer patients are diagnosed at an advanced stage. Therefore, there is an urgent need to find an effective diagnostic method for detecting breast cancer at an early stage to achieve a better prognosis. Several meta-analyses have evaluated the value of biomarkers in diagnosing breast cancer, but which biomarker has the optimal diagnostic value remains unclear.

Objectives: This overview aimed to compare the accuracy of different biomarkers in diagnosing breast cancer.

Methods: We searched PubMed, Embase, the Cochrane Library of Systematic Reviews, and Web of Science. We used AMSTAR-2 (A MeaSurement Tool to Assess systematic Reviews) to assess the methodological quality and Preferred Reporting Items for Systematic Reviews and Meta-analysis diagnostic test accuracy (PRISMA-DTA) for reporting quality. We performed pairwise meta-analyses to estimate the pooled results for each biomarker and conducted indirect comparisons of diagnostic accuracy between biomarkers.

Results: We included 11 systematic reviews (SRs) involving 218 original studies. All SRs were of critically low methodological quality; 3 SRs had minimal reporting flaws and 8 SRs had minor flaws. The pooled sensitivity and specificity were 0.77 and 0.87 for miRNA, 0.70 and 0.87 for circulating cell-free DNA, 0.29 and 0.96 for APC gene promoter methylation, 0.69 and 0.99 for 14-3-3 σ promoter methylation, 0.63 and 0.82 for CA153, 0.58 and 0.87 for CEA, and 0.73 and 0.56 for PSA. Compared with CA153 and PSA, miRNA had a higher sensitivity and specificity. The sensitivity of miRNA was higher than circulating cell-free DNA and CEA, although they had the same specificities. APC gene promoter methylation and 14-3-3σ promoter methylation were more specific than miRNAs, but they had unacceptably low sensitivity.

Conclusions: MiRNA had better diagnostic accuracy than the other six biomarkers. But due to the low quality of included SRs, the results need to be interpreted with caution. Further study should investigate the diagnostic accuracy of different biomarkers in direct comparisons and focus on the value of combined biomarkers.

Patient or healthcare consumer involvement: No.

Evaluation of 'spin' in systematic reviews of diagnostic accuracy studies in high impact factor journals

McGrath TA[1], Bowdridge JC[2], Prager R[2], Frank RA[1], Treanor L[3], Dehmoobad Sharifabadi A[3], Salameh J[4], Leeflang MMG[5], Korevaar DA[6], Bossuyt PMM[5], McInnes MDF[1]

[1] University of Ottawa Department of Radiology, Canada; [2] University of Ottawa Department of Medicine, Canada; [3] University of Ottawa Faculty of Medicine, Canada; [4] Ottawa Hospital Research Institute, Canada; [5] Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Amsterdam Public Health, Amsterdam UMC, University of Amsterdam, The Netherlands; [6] Department of Respiratory Medicine, Amsterdam UMC, University of Amsterdam, The Netherlands

Background: 'Spin' of study findings is common in reports of diagnostic accuracy studies. Multiple studies have shown that clinicians may view journals more favourably based on a higher impact factor. There is a growing body of evidence signalling there may be better methodological quality and higher reporting standards in higher impact factor journals in the diagnostic accuracy literature.

Objectives: To evaluate the frequency of 'spin' in reports of systematic reviews of diagnostic accuracy studies in high impact factor journals, with the hypothesis that the incidence of 'spin' may be lower compared to a series of reviews from 'all-comer' journals previously analyzed.

Methods: We searched MEDLINE from January 2010 to January 2019. We included systematic reviews of diagnostic accuracy studies if they reported a meta-analysis and were published in a journal with an impact factor above 5. Two investigators independently scored each included systematic review for positivity of conclusions as well as actual and potential overinterpretation practices.

Results: Of 137 included systematic reviews, 63(46%) contained one or more forms of actual overinterpretation in the abstract; 52 (38%) in the full-text report; 108 (79%) contained a form of potential overinterpretation. Comparing to the previously assessed series, reviews in this series were less likely to contain one or more forms of actual overinterpretation in the abstract and full-text report or one or more forms of potential overinterpretation ($P < 0.001$ for all comparisons). Significance of these comparisons did not persist for actual overinterpretation in sensitivity analysis in which Cochrane systematic reviews were removed. Reviews published in the Cochrane Database of Systematic Reviews were less likely to contain actual overinterpretation in the abstract or the full-text report than reviews in other high-impact journals ($P < 0.001$ for both comparisons).

Conclusions: Reviews of diagnostic accuracy studies in high-impact journals are less likely to contain overinterpretation or spin. This difference is largely due to the reviews published in the Cochrane Database of Systematic Reviews, which contain spin less often than reviews published in other high-impact journals.

Diagnostic testing is ubiquitous in clinical medicine, from the physical examination, basic blood work and radiographs, to advanced diagnostic testing such as MRI. Systematic reviews are considered a high level of evidence, which may influence both individual clinician practice patterns and population health decisions. Therefore, systematic reviews of diagnostic accuracy studies should be reported as their results justify, free of 'spin', so clinicians and larger health authorities can make proper, unbiased, evidence-based clinical decisions.

Geriatric depression scale-4 and geriatric depression scale-5 for screening of depression among older adults: a systematic review

Brañez-Condorena A[1], Soriano-Moreno DR[2], Solis-Chimoy B[1], Taype-Rondan A[3]

[1] Faculty of Medicine, National University of San Marcos, Lima, Peru; [2] School of Human Medicine, Peruvian Union University, Lima, Peru; [3] St. Ignatius of Loyola University, Research Unit for the Generation and Synthesis of Evidence in Health, Lima, Peru

Background: Older adults are vulnerable to mental health problems such as depression. Screening tools are usually used for the identification of persons with depressive symptoms and the geriatric depression scale (GDS) is one of the most widely used instruments in this population. There are short versions of the GDS such as the GDS-4 and GDS-5 that can help for faster detection of this condition.

Objectives: To assess the accuracy of the GDS-4 and GDS-5 for screening depression in older adults.

Methods: During February 2020, we performed a systematic search in PubMed, PsycINFO, Scopus, and Google Scholar; to identify studies that reported diagnostic performance of the GDS-4 or GDS-5, compared to any reference assessment, in older adults from any setting. Two review authors in parallel performed study selection, data extraction according to the inclusion and exclusion criteria, and risk of bias evaluation of the included studies using the QUADAS-2 tool. We performed meta-analyses for sensitivity and specificity and evaluated the certainty of evidence using the GRADE methodology. The protocol is available at https://bit.ly/2yD1ScZ.

Results: We identified 202 records, of which 21 studies were included. Seventeen studies evaluated the GDS-4 and eight the GDS-5. We identified several versions of the GDS-4 and GDS-5.

Regarding GDS-4: when comparing different cut-off points, the threshold ≥ 1 had the highest pooled sensitivity (88%, 95% confidence interval (CI) 84% to 92%) but low specificity (67%, 95% CI 58% to 75%). In the subgroup analysis, Van Marwijk's GDS-4 had the best balance between sensitivity (86%, 95% CI 76% to 96%) and specificity (77%, 95% CI 67% to 87%).

Regarding GDS-5: when comparing different cut-off points, the threshold ≥ 2 had the best relation between sensitivity (87%, CI 95% 81 to 93) and specificity (80%, CI 95% 74 to 87). In the subgroup analysis, Hoyl's GDS-5 has the best sensitivity (87%, CI 95% 80 to 95) and specificity (81%, CI 95% 73 to 90). In general, the certainty of the evidence was very low for sensitivity and specificity of both versions (Table 1).

Conclusions: Among the assessed GDS versions, Hoyl's GDS-5 with a cut-off point ≥ 2 had the best performance in sensitivity and specificity. It is important to make an early diagnosis of depression to start correspondent treatment early. Screening tools for depression with good performance and reduced application time are required.

Table 1. Summary of findings of sensitivity and specificity of GDS-4 and GDS-5 compared to any reference.

Diagnostic tests	Reference standards	Number of studies (participants)	Summary of sensitivity % (95% CI)	Summary of specificity% (95% CI)	Certainty of the evidence	Consequences in a population of 1000 and a prevalence of depression of 18.1%*	
						Underdiagnosed (false negatives)	Overdiagnosed (false positives)
GDS-4 (Van Marwijck) cut-off point ≥1	DSM-V, DSM-IV, DIS	4 (1317)	86 (76-96)	77 (67-87)	S: Very low [b,c,d,e] E: Very low [b,c,d]	25	188
GDS-4 (D'Ath) cut-off point ≥1	DSM-IV, BAS, ICD-10, GMS	6 (532)	89 (83-95)	66 (54-77)	S: Moderate [b] E: Low [b,c]	20	287
GDS-4 (Cheng) cut-off point ≥1	ICD-10, DSM-III	2 (310)	89 (85-93)	45 (39-50)	S: Moderate [b] E: Low [b,g]	20	278
GDS-4 (Van Marwijck) cut-off point ≥2	DSM-IV, DIS, GDS-15	4 (1696)	76 (62-90)	78 (64-91)	S: Very low [a,c] E: Very low [a,c]	42	180
GDS-4 (D'Ath) cut-off point ≥2	DSM-IV, BAS, GDS-15, GMS	6 (1316)	65 (52-79)	85 (76-95)	S: Very low [a,c] E: Very low [a,c]	65	90
GDS-4 (Cheng) cut-off point ≥2	ICD-10, DSM-III	2 (310)	74 (69-79)	63 (58-69)	S: Low [a,b,c,d] E: Very low [a,b,c]	47	303
GDS-4 (Cheng) cut-off point ≥3	ICD-10, DSM-III	2 (310)	58 (52-63)	80 (76-84)	S: Very low [a,c] E: Moderate [a,c]	76	164
GDS-5 (Hoyl) cut-off point ≥2	DSM-IV, GDS-15	6 (1864)	87 (80-95)	81 (73-90)	S: Very low [a,c] E: Very low [a,c]	24	156

GDS: Geriatric Depression Scale, DSM: Diagnostic and Statistical Manual of Mental Disorders, ICD: International Classification of Diseases, DIS: Diagnostic Interview Schedule, BAS: Brief Assessment Scale, GMS: Geriatric Mental State, CI: Confidence interval, S: sensitivity, E: specificity.

*We considered the prevalence reported by a previous systematic review (Bao, 2017, doi: 10.1016/j.neubiorev.2017.01.032) which evaluated depression in community-dwelling older adults.

Explanations of the decrease in the certainty of the evidence:

a. Risk of bias: Observational studies show less than 70% of QUADAS-2 items as low risk.

b. Indirect evidence: Not directly related to clinical outcomes.

c. Inconsistency: Heterogeneity is greater than 40%

d. Inconsistency: Heterogeneity is greater than 80%

e. Imprecision: The confidence interval of the sensitivity crosses the value of 80.

f. Imprecision: The confidence interval of the sensitivity crosses the value of 65.

g. Imprecision: The confidence interval of the specificity crosses the value of 50.

Latent class models for individual participant data meta-analyses of diagnostic test accuracy studies with imperfect reference standards

Negeri ZF[1], Levis B[2], Wu Y[3], Thombs BD[2], Benedetti A[1]

[1] Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Canada; [2] Lady Davis Institute for Medical Research, Jewish General Hospital, Canada; [3] Department of Psychiatry, McGill University, Canada

Background: Depression accounts for more years of "healthy" life lost than any other medical condition. Typically, questionnaire-based screening tools and clinically administered diagnostic interviews are used to screen for and diagnose major depressive disorders. However, neither the screening tools nor the diagnostic interviews accurately screen or diagnose depressive symptoms because of the imperfect nature of the diagnostic interviews used as reference standards. Systematic reviews and meta-analyses results based on such imperfect reference standards may lead to misleading conclusions that misinform both clinicians and other decision-makers. Latent class models have been commonly applied to correct for imperfect reference or gold standards in conventional or aggregate data diagnostic test accuracy studies. Most of these latent class models used a Bayesian analysis approach to estimate unknown model parameters. To the best of our knowledge, there have been no methodological studies that attempted to account for imperfect reference standards in the context of individual participant data meta-analyses (IPDMA) of diagnostic test accuracy studies.

Objectives: To propose and validate latent class models for IPDMA to estimate the actual diagnostic test accuracy of both screening tools and imperfect reference standards for depression screening and diagnosis.

Methods: We will develop and evaluate latent class analysis-based models by exploring both Frequentist and Bayesian approaches to the problem of imperfect reference standards in IPDMA of diagnostic test accuracy data. We will illustrate the models using our database that consists of more than 100 studies and 46,000 participants on the most commonly used tool for detecting major depression in primary care – the Patient Health Questionnaire-9 (PHQ-9). In this database, the PHQ-9 is compared to diagnostic interviews such as the Structured Clinical Interview for DSM (SCID), Composite International Diagnostic Interview (CIDI), and the Mini International Neuropsychiatric Interview (MINI).

Anticipated Results: We expect that our models will generate more realistic test characteristics of depression screening tools and depression diagnosing clinical interviews by correcting for biases in results due to the imperfect nature of reference standards being used. This would better inform stakeholders about the correct diagnostic accuracy of the depression screening tools and diagnosing interviews.

Conclusions: Our proposed methods will have implications beyond IPDMA of depression screening tools and diagnostic interviews.

Patient or healthcare consumer involvement: There was no direct patient or healthcare consumer involvement in this study. Nevertheless, the outcome of this study will be a welcome addition to the body of knowledge and clinicians and policy-makers concerned with the accuracy of depression screening tools and interviews.

Meta-analyses of diagnostic test accuracy could not be reproduced

Stegeman I[1], Leeflang M[1]

[1] Department of Clinical Epidemiology, Biostatistics and Bioinformatics, The Netherlands

Background: In 2005, John Ioannides shook the scientific world with his landmark publication "Why most research findings are false". This finding seems to be true for all types of research, including clinical research. Mistakes in methodology or statistical approach can lead to false conclusions. When using these false claims as evidence for medical decisions, these mistakes may have major impact on clinical care.

Objectives: To investigate the reproducibility of diagnostic accuracy meta-analyses, as reported in published systematic reviews.

Methods: We selected all systematic reviews of diagnostic test accuracy containing a meta-analysis, published in January 2018 and retrieved in MEDLINE through OVID. All reviews reported a summary estimate of sensitivity and specificity. We requested the protocol from the review authors and used the protocol and the information in the published review to reproduce the reported meta-analysis. We evaluated the following items of included studies:

1) the availability of a 2x2 table;
2) availability of full text papers;
3) correctness of information included in the primary papers; and
4) the correctness of the pooled estimate.

Full replication was met if all four items were reported and the point estimate differed < 1% point from the reported point estimates. We studied reproduction of MAs in which the 2x2 tables were available. If the information for the 2x2 tables was not available we returned to the primary papers to search for information to comprise the 2x2 tables.

Results: Of the 51 included reviews, 16 had a protocol registered in PROSPERO and five of those responded to our request for a protocol. Nineteen reviews (37%) provided the 2x2 tables that were included in the meta-analysis. In 14 of those, the outcome of the meta-analysis could be reproduced. In 32 (63%) MAs the information to comprise 2x2 tables was not available. In 17 (33%) of those the primary papers were available. Of those 0 analyses could be replicated. Considering the correctness of the numbers from the primary papers and the complete reporting of the search strategy, only one meta-analysis was fully replicable.

Conclusions: Published meta-analyses of diagnostic test accuracy were poorly replicable. This was partly because of lack of information about the methods and data used; and partly because of mistakes in the data extraction or data reporting.

Patient or healthcare consumer involvement: Because of the nature of this study no patients were involved.

Methods and characteristics of network meta-analyses using diagnostic test accuracy studies

Veroniki AA[1], Tsokani S[1], Rücker G[2], Takwoingi Y[3], Mavridis D[1]

[1] Department of Primary Education, School of Education, University of Ioannina, Ioannina, Greece; [2] Institute for Medical Biometry and Statistics, Faculty of Medicine and Medical Center, University of Freiburg, Freiburg, Germany; [3] Test Evaluation Research Group, Institute of Applied Health Research, University of Birmingham, Birmingham, UK

Background: The diagnosis of a clinical condition is usually the first and more crucial step before initiating treatment. Diagnostic tests are routinely used for confirming or excluding a target condition. Although most diagnostic test accuracy (DTA) studies have focused on assessing a single index test, studies and systematic reviews are increasingly comparing the accuracy of multiple index tests to facilitate the selection of the best performing test(s) for patient care. For example, HPV DNA, HPV mRNA, and co-testing can be used for cervical cancer diagnosis. But which test is the most sensitive (or specific)? Since studies that directly compare test accuracy are not always available and comparisons between multiple tests constitute a network, DTA network meta-analysis (DTA-NMA) has been proposed.

Objectives: To identify and summarize the properties of DTA-NMA methods for comparing the accuracy of multiple diagnostic tests. To describe the network characteristics of empirical DTA-NMA studies.

Methods: We conducted a methodological review of statistical and empirical studies that performed, described, or evaluated a DTA-NMA of at least three diagnostic tests. We searched PubMed, JSTOR, and Web of Science. Studies of any design published in English were eligible. We also included relevant unpublished material. Several methods have been suggested for modelling DTA studies in a NMA, which vary in complexity. We will present the approaches together with a critique of their strengths and limitations and will identify gaps where methodology is lacking. We will present the characteristics of previously conducted DTA-NMAs, and the methods that have been applied. Quantitative data will be summarized using medians and interquartile ranges, while categorical data will be summarized using frequencies and percentages. We will use cervical cancer as a case study, to present an application of the DTA-NMA methods and to determine the most promising test (in terms of sensitivity and specificity) for use as the primary screening test for cervical cancer.

Results: We included 38 studies. Evaluations are ongoing and results will be ready by June 2020 for presentation at the Cochrane Colloquium.

Conclusions: Our study will provide a comprehensive overview of the methods for conducting a DTA-NMA, insight into the characteristics of DTA-NMA applications, and propose recommendations for appropriate use and reporting.

Patient or healthcare consumer involvement: While there has been no direct patient involvement in this study, well conducted and reported DTA meta-analyses influence patient care directly since they are used in decision-making and for developing guidelines. When alternative tests that can be used at the same point in the diagnostic pathway exist, studies that compare several tests and estimate differences in sensitivity and specificity are more informative than those that evaluate the accuracy of a single index test. Therefore, the comparison of multiple diagnostic tests using NMA-DTA can impact clinical decisions and patient health.

QUADAS-C: a tool for assessing risk of bias in comparative diagnostic accuracy studies

Yang B[1], Whiting P[2], Davenport C[3], Deeks J[3], Hyde C[4], Mallett S[3], Takwoingi Y[3], Leeflang M[1]

[1] Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Amsterdam UMC, University of Amsterdam, The Netherlands; [2] Population Health Sciences, Bristol Medical School, UK; [3] Test Evaluation Research Group, Institute of Applied Health Research, University of Birmingham, UK; [4] Exeter Test Group, Institute of Health Research, College of Medicine and Health, University of Exeter, UK

Background: Comparative diagnostic test accuracy studies assess the accuracy of multiple tests in the same study and compare their accuracy. While these studies have the potential to yield reliable evidence regarding comparative accuracy, shortcomings in the design, conduct and analysis may bias their results. The currently recommended quality assessment tool for diagnostic accuracy studies, QUADAS-2, is not designed for the assessment of test comparisons.

Objectives: We developed QUADAS-C as an extension to QUADAS-2 to assess the risk of bias in comparative diagnostic test accuracy studies.

Methods: Through a four-round Delphi study involving 24 international experts in test evaluation and a face-to-face consensus meeting, we developed a draft version of QUADAS-C which will undergo piloting in ongoing systematic reviews of comparative diagnostic test accuracy.

Results: QUADAS-C retains the same four-domain structure of QUADAS-2 (patient selection, index test, reference standard, flow and timing) and is comprised of additional questions to each QUADAS-2 domain. A risk of bias judgement for comparative accuracy requires a risk of bias judgement for each test (QUADAS-2), and additional criteria specific for test comparisons. Examples of such additional criteria include whether patients either received all index tests or were randomized to index tests, and whether index tests were interpreted blinded to other index tests.

Conclusions: QUADAS-C will be useful for systematic reviews of diagnostic test accuracy addressing comparative accuracy questions. Furthermore, researchers may use this tool to identify and avoid risk of bias when designing a comparative diagnostic test accuracy study. Currently a draft version of QUADAS-C is being piloted and the tool will be finalized by the time of the conference. This tool was developed together with the QUADAS-C Advisory Group.

Patient or healthcare consumer involvement: Patients or healthcare consumers were not involved in the design and execution of this study.

Searching practices and inclusion of unpublished studies in systematic reviews of diagnostic accuracy

Korevaar D[1], Salameh J[2], Vali Y[3], Cohen J[4], McInnes M[5], Spijker R[6], Bossuyt P[7]

[1] Department of Respiratory Medicine, University Medical Center (UMC), University of Amsterdam, The Netherlands; [2] Clinical Epidemiology Program, Ottawa Hospital Research Institute, Canada; [3] Department of Clinical Epidemiology and Biostatistics and Bioinformatics, University Medical Center (UMC), University of Amsterdam, The Netherlands; [4] Department of General Pediatrics and Pediatric Infectious Diseases, Necker – Enfants Malades Hospital, Assistance Publique – Hôpitaux de Paris, France; [5] Department of Radiology, Clinical Epidemiology Program, Ottawa Hospital Research Institute, University of Ottawa, Canada; [6] Cochrane Netherlands, Julius Center for Health Sciences and Primary Care, Utrecht University, The Netherlands; [7] Department of Clinical Epidemiology and Biostatistics and Bioinformatics, Amsterdam UMC, University of Amsterdam, The Netherlands

Background: Many diagnostic accuracy studies are never reported in full in a peer-reviewed journal. Searching for unpublished studies may avoid bias due to selective publication, enrich the power of systematic reviews, and thereby help to reduce research waste.

Objectives: To assess searching practices among recent systematic reviews of diagnostic accuracy, with a special focus on the identification and inclusion of unpublished studies.

Methods: We included systematic reviews if they had evaluated the diagnostic accuracy of one or more index tests against a reference standard in humans. We extracted data from 100 non-Cochrane systematic reviews of diagnostic accuracy indexed in MEDLINE and published between October 2017 and January 2018, and from all 100 Cochrane Reviews of diagnostic accuracy published by December 2018, irrespective of whether meta-analysis had been performed.

Results: Non-Cochrane and Cochrane Reviews searched a median of 4 (interquartile range (IQR) 3 to 5) and 6 (IQR 5 to 9) databases, respectively; most often MEDLINE/PubMed (n = 100 and n = 100) and Embase (n = 81 and n = 100). No language restrictions were applied in 37 and 90 reviews, and efforts to contact authors in case of incomplete or unclear data were announced or reported by 31 and 78 reviews. Additional efforts to identify studies beyond searching bibliographic databases were performed in 76 and 98 reviews, most often through screening reference lists (n = 71 and n = 96), review/guideline articles (n = 18 and n = 52), citing articles (n = 3 and n = 42), or contacting authors/experts (n = 6 and n = 37). Specific sources of unpublished studies were searched in 22 and 68 reviews, for example conference proceedings (n = 4 and n = 18), databases only containing conference abstracts (n = 2 and n = 33), or trial registries (n = 12 and n = 39). At least one unpublished study was included in 17 and 23 reviews. Overall, 39 of 2082 studies (1.9%) included in non-Cochrane reviews were unpublished, and 64 of 2780 studies (2.3%) in Cochrane Reviews, most often conference abstracts (97/103).

Conclusions: Searching practices vary considerably across systematic reviews of diagnostic accuracy. Cochrane Reviews seem to make more efforts to identify studies, both published and unpublished. Unpublished studies are a minimal fraction of the evidence included in recent reviews. This represents avoidable research waste, may introduce bias, and could have negative effects on patient care.

Selective cut-off reporting in depression screening accuracy studies: a comparison of meta-analysis of published cut-offs only versus all cut-offs

Neupane D[1], Levis B[1], Bhandari PM[1], Thombs BD[1], Benedetti A[1], DEPRESSD Collaboration

[1] McGill University, Canada

Background: Selectively reporting accuracy results from only well-performing cut-offs would be expected to result in biased accuracy estimates in meta-analyses. It is not known whether the extent of bias differs depending on the availability of a well-defined standard cut-off.

Objectives: We compared (1) bias in accuracy estimates and (2) cut-off reporting patterns in studies on the diagnostic accuracy of the Patient Health Questionnaire-9 (PHQ-9; well-defined standard cut-off of ≥ 10) and the Edinburgh Postnatal Depression Scale (EPDS; no standard cut-off, common cut-offs = ≥ 10 to ≥ 13).

Methods: We analyzed a subset of datasets from two separate individual participant data meta-analyses (IPDMAs) on PHQ-9 and EPDS accuracy for screening to detect major depression. Separately, for the PHQ-9 and EPDS, we used bivariate random effects meta-analysis to compare accuracy estimates based on published cut-offs only versus all cut-offs from all studies. To assess cut-off reporting patterns, we compared the number of published cut-offs below and above the standard cut-off (or common range) when the study-specific optimal cut-off was lower or higher than the standard cut-off (or common range).

Results: Compared to IPDMA of all cut-offs, PHQ-9 sensitivity estimates based on published cut-offs only were underestimated for cut-offs below ≥ 10 and overestimated for cut-offs above ≥ 10 (median differences: -0.06 and 0.07). EPDS sensitivity estimates were similar for cut-offs below ≥ 10 but higher for cut-offs above ≥ 13 (median differences: 0.01 and 0.14). PHQ-9 studies with optimal cut-offs below ≥ 10 reported more cut-offs below ≥ 10 and those with optimal cut-offs above ≥ 10 reported more cut-offs above ≥ 10 (mean cut-offs: 8.8 and 11.8). EPDS studies with optimal cut-offs below ≥ 10 did not report more cut-offs below 10 but those with optimal cut-offs above ≥ 10 reported more cut-offs above 10 (mean cut-offs: 9.9 and 11.8).

Conclusions: Selective cut-off reporting and resulting bias in accuracy estimates were more pronounced for the PHQ-9 than for the EPDS. Researchers evaluating diagnostic accuracy of screening tools should report results for all relevant cut-offs.

Patient or healthcare consumer involvement: There was no patient or healthcare consumer involvement in the present study.

Study designs for comparative diagnostic test accuracy: a methodological review and classification

Yang B[1], Olsen M[1], Vali Y[1], Langendam M[1], Takwoingi Y[2], Hyde C[3], Bossuyt P[1], Leeflang M[1]

[1] Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Amsterdam UMC, University of Amsterdam, The Netherlands; [2] Test Evaluation Research Group, Institute of Applied Health Research, University of Birmingham, UK; [3] Exeter Test Group and South West CLAHRC, University of Exeter Medical School, UK

Background: Systematic reviews of diagnostic test accuracy (DTA) addressing comparative questions include studies comparing the accuracy of two or more index tests (i.e. comparative DTA studies). Well-conducted comparative DTA studies represent the most reliable evidence for determining the relative accuracy of tests. However, the range of available study designs indicates varying risk of bias, and inconsistent labeling of designs complicates study identification and classification.

Objectives:

1) To examine the variability of comparative DTA study designs and to propose a study design classification scheme; and
2) To describe study design labels used by comparative DTA study authors.

Methods: A methodological review of 100 comparative DTA studies published in 2015, 2016 and 2017. These were randomly sampled from comparative DTA studies included in 238 comparative DTA systematic reviews indexed in MEDLINE in 2017. From each study, we extracted six design features (direction of data collection, number of groups sampled, sampling method, allocation of participants, reference standard and verification of disease status) and labels used by authors.

Results: Most studies (n = 57) enrolled a single group of participants, with each participant receiving all index tests. We classified the studies into six study design categories based on how participants were allocated to each index test: 'paired' (n = 78), 'partially paired, random subset' (n = 0), 'partially paired, nonrandom subset' (n = 2), 'unpaired randomized' (n = 1), 'unpaired nonrandomized' (n = 3) and 'externally controlled' (n = 1). The allocation method of 15 studies were unclear. Sixty-one studies reported 33 unique study design labels, but only nine labels conveyed information that there was a test comparison in the study.

Conclusions: Our classification scheme for comparative DTA study designs may help systematic review authors when assessing risk of bias and interpreting results. In addition, researchers can use the scheme to select optimal designs for future studies. Further work is needed to develop an agreed set of informative labels for comparative DTA studies.

Patient or healthcare consumer involvement: Patients or healthcare consumers were not involved in this study.

Use of GRADE in the non-Cochrane systematic reviews of diagnostic test accuracy

Shi Q[1], Wang Z[2], Zhou Q[1], Xun Y[2], Zhao S[3], Chen Y[4]

[1] The First School of Clinical Medicine, Lanzhou University, China; [2] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, China; [3] School of Public Health, Lanzhou University, China; [4] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, Lanzhou, China

Background: Systematic reviews of diagnostic test accuracy summarize the accuracy of sensitivity and specificity and are important to inform evidence-based use of diagnostic tests in clinical practice. When there is a meta-analysis, the GRADE approach is being used to assess the quality/certainty of evidence, to interpret findings, and to draw conclusions from randomized or non-randomized studies of interventions in 'Summary of findings' tables.

Objectives: To analyse how many non-Cochrane systematic reviews of diagnostic test accuracy used the GRADE approach, and how GRADE was used.

Methods: This was a methodological (research-on-research) study. We systematically retrieved non-Cochrane systematic reviews of diagnostic test accuracy from inception to March 2020 in the following databases: MEDLINE (via PubMed), Embase, Web of Science, CBM (China Biology Medicine), CNKI (China National Knowledge Infrastructure), and Wanfang Data. Two review authors independently undertook study selection and data extraction. We extracted information about methods used for quality of evidence assessment, and if they used GRADE, we analyzed their methods and compared them with the GRADE guidelines. We used descriptive statistics to present the data of the included studies.

Results: The results will be presented at the meeting.

Conclusions: The results will be presented at the meeting.

Patient or healthcare consumer involvement: None.

SYSTEMATIC REVIEWS OF INTERVENTIONS

A bibliometric analysis and visualization of studies based on SEER database research

Li R[1], Li Y[1], Zhang W[1], Li M[1], Bing Z[2], Yang K[2]

[1] School of Public Health, Lanzhou University, China; [2] School of Basic Medical Sciences, Lanzhou University, China

Background: In order to reduce the cancer burden on the population, the National Cancer Institute (NCI) established the Surveillance, Epidemiology, and End Results Program Database (SEER) for tumor patients in the country in 1973. The huge amount of information in the SEER database provides a powerful data support for in-depth study of tumors. Therefore, there are a large number of medical studies that use the data provided by SEER for analysis and research.

Objectives: To comprehensively analyze the current status of studies based on SEER databases, and to understand the research hotspots and development trends of these studies.

Methods: We retrieved publications and their literature information from the Web of Science Core Collection database and the time span was defined as "all years". We used Microsoft Excel 2013 to detect the trend of annual numbers of publications, and used VOSviewer 1.6.9 software as the bibliometric method to analyze the research areas, countries/regions, institutions, authors, journals, research hotspots and frontiers, and development trends.

Results: We included in the bibliometric analysis 7249 related studies based on the SEER database. In 1980, the first related research based on the SEER database was published, and the number of publications with an increasing trend, even reaching 1048 studies in 2019. More than half of the studies was produced after 2015. The studies were published in 1084 journals, and a total of 19,740 authors from 89 countries or regions participated in the relevant research. The most prolific country and institution were the USA and NCI, respectively. Karakiewicz PI was the most productive author, while Cancer was the most prolific journal. Relevant literature mainly focused on the field of oncology.

Conclusions: Interest in the SEER database is increasing year by year, and big data-oriented cohort research has become a research hotspot. Researchers should attach importance to the role of controlled study in data analysis. The integration and transformation of biomedical big data can help generate evidence-based scientific information.

Patient or healthcare consumer involvement: Not applicable.

A systematic literature review on implementation of cervical cancer screening and associated factors in Nepal: study protocol

Shrestha A[1], Greibe Andersen J[1], Neupane D[2], Ghimire S[3], Campbell C[4], Kallestrup P[1]

[1] Center for Global Health, Department of Public Health, Aarhus University, Denmark; [2] Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA; [3] Nepal Cancer Care Foundation, Nepal; [4] Usher Institute, University of Edinburgh, UK

Background: Cervical cancer is a major cause of cancer death among women in Nepal. The high burden of this disease necessitates identifying relevant evidence to inform policy development and guidelines. This protocol describes a planned systematic review that will collate and report the experiences of implementation of cervical cancer screening and associated factors in Nepal over the last two decades to identify knowledge gaps to inform future implementation strategies.

Objectives:

- identify relevant peer-reviewed and grey literature describing cervical cancer screening delivery and coverage among women in Nepal;
- describe and synthesize the experiences of implementing cervical cancer screening in Nepal;
- identify any barriers to and facilitators of cervical cancer screening coverage in Nepal described in the literature; and
- collate major findings and recommendations on cervical cancer screening from studies conducted in Nepal from 1 January 2000 to 31 December 2018.

Methods: This protocol was developed according to the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA_P) statement and will apply eligibility criteria to screening and select peer-reviewed research articles and grey literature. The systematic review has been registered with PROSPERO (CRD42019144645). The study protocol has been published in the Journal of Global Health reports and cited as, "Shrestha AD, Andersen JG, Neupane D, Ghimire S, Campbell C, Kallestrup P. Protocol for systematic literature review on implementation of cervical cancer screening and associated factors in Nepal from 2000 to 2018. Journal of Global Health Reports. 2020;4:e20"0023". A computer-based search will be conducted for each type of publication in the PubMed/MEDLINE, CINAHL, Scopus and Embase databases using various search terms. We will modify search terms according to each database and screen the reference lists of the included studies to identify additional relevant materials. Data synthesis will use narrative synthesis and meta-analysis where appropriate.

Ethics and dissemination: This study does not require ethical approval as only secondary data from published and grey literature will be assessed. The review will be published in a peer-reviewed scientific journal.

Patient or healthcare consumer involvement: Patients and the public are not involved in this study.

Caregiver-provided and home-based individual cognitive stimulation: the cultural adaptation process guided by the systematic review method and community involvement approach

Silva R[1], Bobrowicz-Campos E[2], Costa P[1], Cardoso D[2], Gil I[1], Almeida M[2], Apóstolo J[2]

[1] Catholic University of Portugal, Porto; Health Sciences Research Unit, Nursing, Nursing School of Coimbra, Portugal; [2] Health Sciences Research Unit, Nursing, Nursing School of Coimbra, Portugal

Background: Promoting successful aging in the community is a societal priority. Cognitive stimulation is one of the interventions with a positive impact on older adults' health and well-being, as it preserves cognitive functioning, contributing to the maintenance of autonomy and quality of life.

Objective: To culturally adapt the Making a Difference 3 (MD3), an individual Cognitive Stimulation program (iCSP) delivered by caregivers to people with mild to major Neurocognitive Disorder (PwNCD) in a home-based setting.

Methods: The cultural adaptation of the MD3 was organized in five phases. Initially, we conducted the systematic review (SR) on the effectiveness of the iCSP on older adults' cognitive performance, according to the Joanna Briggs Institute methodology. In phase II, the iCSP-MD3 was translated into the European Portuguese language and culturally adapted based on the information generated by the SR and by the academics, researchers, health professionals and caregivers involved (n = 12). Phase III consisted of reviewing the culturally adapted MD3 with eight experts in nursing, psychology, occupational therapy and social working using the Delphi method and with caregivers (n = 10) through focus groups. In phase IV, to test the feasibility and effects of the MD3 Portuguese version, we conducted a randomized controlled trial (RCT) involving 52 dyads (PwNCD and a caregiver, 28 allocated in the experimental group and 24 in the control group. Phase V consisted of exploring through interviews (conducted with 2 PwNCD, and 2 caregivers) the meanings attributed by the participants to the iCSP-MD3.

Results: The SR on the effectiveness of the caregiver-delivered iCSP has identified beneficial effects in several cognitive domains, and thus generated evidence that underpinned the achievement of the following phases. Phase II resulted in the iCSP preliminary version. Phase III allowed for the adaptation of the verbal and structural contents of the iCSP program, creating an accessible and fitting version with meaning to the target population. As for the RCT conducted (phase IV), results from the intention-to-treat analysis revealed significant improvements in cognition (orientation and order comprehension) of PwNCDs. Beneficial effects of the intervention were also observed in the quality of life of the PwNCDs, but only from the caregiver's perspective. Qualitative data collected in phase V showed that participation in the iCSP-MD3 was evaluated as very positive and significant.

Conclusions: Through a thorough SR process, the research team was able to delineate essential methodological choices based on previous studies conducted in this area, such as the active inclusion of relevant stakeholders during the cultural adaptation of the iCSP-MD3. Moreover, the research team as able to identify the main domains that can be potentially improved through iCSP, as well as identify potential facilitators and barriers during its implementation in home-based settings. Overall, the iCSP-MD3 proved to be a feasible and meaningful intervention for the Portuguese population.

D-mannose for preventing and treating urinary tract infections in adults and children

Cooper T[1], Teng C[1], Howell M[1], Teixeira-Pinto A[1], Tong A[1], Wong G[1]

[1] Cochrane Kidney and Transplant Review Group, Australia

Background: Urinary tract infections (UTI) are common in global populations. Approximately 50% of females experience an episode in their life and 20% of adults and children suffer chronic symptomatic UTIs (≥ 2 episodes in 6 months or 3 in 12 months). Long-term antibiotics may lead to antibiotic resistance, adverse effects, significant patient burden and health costs. D-mannose is a sugar part of normal human diets that plays a role in the glycosylation of most secretory proteins. It attaches to bacteria, prevents adherence to the urothelial cells, and may have a role in the prevention and treatment of UTI in at-risk individuals.

Objectives: To assess the benefits and harms of D-mannose for preventing and treating urinary tract infections in adults and children, in any setting.

Methods: We searched the Cochrane Kidney and Transplant Register (which includes CENTRAL, MEDLINE, Embase, and ICTRP) to 1 March 2020 for RCTs of D-mannose in any formula, route. There were no restrictions on language, dates, or blinding. Screening, data extraction and analyses were performed by three independent review authors.

Outcomes: presence, or recurrence, of symptomatic bacteriuria UTI, symptomatic UTI, asymptomatic bacteriuria; any changes to previous treatment; and pain.

Results: We included six RCTs (704 adults) (Figure 1). Interventions varied in dose and frequency: D-mannose (2 RCTs), D-mannose plus vitamins (4 RCTs); antibiotics; vitamins; placebo; no treatment. Risk of bias was judged for the six included studies overall to be at high risk (Figure 2). Most concerns were around the lack of allocation concealment and blinding (open-label studies), or limited details in abstracts. No two studies investigated comparable intervention arms, so no meta-analysis was undertaken (Table 1). Separately, studies reported some improvement from D-mannose (Table 2): D-mannose plus vitamins found a slightly lower incidence of recurrent cystitis compared to other combinations of D-mannose plus vitamins at 3 months (N = 92). D-mannose plus vitamins found no difference in UTI incidence compared to vitamins at 6 months (N = 95). D-mannose plus vitamins found a difference in reduction of UTI compared to placebo at 3 months (N = 31). Both D-mannose and antibiotics compared to no treatment found a lower incidence of recurrent UTI, but no difference in mean time to UTI at 6 months (N = 308). D-mannose found a difference in time to UTI compared to antibiotics at 6 months (N = 60). Thirty-seven participants reported diarrhoea, with some nausea, headache, skin rash, and vaginal burning. For all comparisons and outcomes, GRADE was rated as very low certainty evidence, downgraded twice for very serious limitations to study design (high risk of bias), and once for sparse data (Table 3).

Conclusions: There is no evidence to support or refute the use of D-mannose to prevent or treat UTIs in adults and children. Further high-quality research through RCTs, is required to evaluate the benefits and harms.

Additional files: References; Figures; Tables

Effectiveness of psychotherapy for geriatric negative emotions: a systematic review and meta-analysis of 12 randomized controlled trials

Zhang W[1], Li X[1], Yan P[2], Hou L[1], Ma B[3], Shi X[4], Liu X[1], Yang K[3]

[1] Evidence-Based Social Science Research Center, School of Public Health, Lanzhou University, China; [2] Institute of Clinical Research and Evidence-Based Medicine, The Gansu Provincial Hospital, China; [3] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, China; [4] Institute for Evidence-Based Rehabilitation Medicine of Gansu Province, China

Background: Psychotherapy is a common treatment for geriatric loneliness, anxiety and depression, but the effectiveness and the long-term efficacy of it remains controversial.

Objectives: To analyze and summarize the effectiveness of psychotherapy, including reminiscence therapy, cognitive behavioral therapy (CBT), group therapy, mindfulness-based cognitive therapy, Internet cognitive behavioral therapy (ICBT), for geriatric loneliness, anxiety and depression.

Methods: We searched the Cochrane Library, PubMed, Embase, Web of Science and SinoMed from inception to July 2018. We included randomized controlled trials (RCTs) that evaluated the effectiveness of psychotherapy in the treatment of geriatric loneliness, anxiety and depression. The primary outcomes included the relief of depression, anxiety or loneliness symptoms. We used Cochrane's risk-of-bias tool to assess the methodological quality of the included studies and performed meta-analysis using Review Manager 5.3.

Results: We included a total of 12 RCTs comprising 962 individuals. Meta-analysis results showed significant effects of psychotherapy in relieving depression (n = 576; standardized mean difference (SMD) = -1.39; 95% confidence interval (CI) = -2.04 to -0.73; P < 0.05), loneliness (n = 506; SMD = -1.21; 95% CI = -2.04 to -0.38; P < 0.05) and anxiety (n = 217; SMD = -1.45; 95% CI = -2.59 to -0.31; P < 0.05). Subgroup analysis revealed that CBT had little effect on relieving anxiety (n = 124; SMD = -1.53; 95% CI = -4.05 to 0.99; P > 0.05). None of the studies met all the items for a high methodological quality assessment. And 58% of them were of high risk of bias in blinding of participants and personnel.

Conclusions: Psychotherapy might improve geriatric depression, loneliness and anxiety of the elderly without severe mental or physical disease. But the results still need more well-designed RCTs to confirm.

Patient or healthcare consumer involvement: No

Effects of exercise on adults with systemic lupus erythematosus: a systematic review and meta-analysis

Xun Y[1], Liu X[2], Yang M[2], Li Y[2], Si W[2], Chen Y[1]

[1] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, Lanzhou, China; [2] School of Public Health, Lanzhou University, Lanzhou, China

Background: Systemic Lupus Erythematosus (SLE) is an autoimmune disease associated with widespread inflammation and tissue damage. Although great progress has been made in the treatment of SLE, the quality of life of SLE patients has not been effectively improved. Evidence suggests that in some cases, exercise therapy is as effective as medication, and even more effective in specific cases.

Objectives: To evaluate the effects of different exercise patterns on SLE patients.

Methods: We systematically searched PubMed, Embase, the Cochrane library, Web of Science and China Biology Medicine disc, China National Knowledge Infrastructure, Wanfang Data from their inception to 31 October 2019. We also searched Google and Baidu academics. Two researchers then independently screened literature, extracted available data, and evaluated quality studies from the included studies. We only selected randomized controlled trials. We used the risk of bias (ROB) tool to appraise the quality of each included study. The mean differences (post-pre) with 95% confidence intervals (CIs) were used to analyze the effect size of the studies. All data were calculated by a random-effects model and the mean difference (MD) was preferred, the standardized mean difference (SMD) was used to calculate the same results measured at different scales. Statistical heterogeneity among the studies was examined with inconsistency (I^2). We used Review Manager 5.3 to perform the statistical analyses.

Results: We included 11 randomized controlled trials and four quasi-randomized controlled trials. The results of this systematic review and meta-analysis showed that exercise is safe and feasible for SLE patients, it could effectively reduce fatigue and depression, improve quality of life, and effectively improve cardiovascular capacity to varying degrees. Compared to sedentary, the effect of aerobic exercise were statistically significant on fatigue (MD -0.51, 95% CI -0.66 to -0.37), depression (SMD -0.56, 95% CI -0.88 to -0.24), gas exchange ratio of cardiovascular function (MD 0.05, 95% CI 0.01 to 0.09). Resistance exercise could improve patients' quality of life in all areas (except on Vitality). Combined exercise could relieve fatigue (MD -1.21, 95% CI -2.15 to -0.26), improve mental health. There was no statistical significance between aerobic exercise and resistance training in disease activity, fatigue, depression and other quality of life, cardiovascular function.

Conclusions: Although the best exercise plan has yet to be proven, according to current research, aerobic exercise whose exercise mode mainly by walking is generally recommended. Clinicians and relevant healthcare professionals should encourage SLE patients to change their sedentary lifestyle and start exercising.

Patient or healthcare consumer involvement: None.

Effects of low-carbohydrate diets versus low-fat diets on body weight loss and metabolic risk factors: a meta-analysis of randomized controlled clinical trials

Jin Y[1], Li J[1], Li X[1], Yang K[1]

[1] Lanzhou University, China

Background: Obesity is a 21st-century major public health challenge, not only in Western countries but also in Asian countries. In recent years, there has been growing interest as to whether low-carbohydrate diets are as effective as, or perhaps better than, traditional low-fat/low-energy diets for weight management. However, findings from various studies of low-carbohydrate and low-fat diets for weight loss remain controversial.

Objectives: To compare the effects of a low-carbohydrate or ketogenic diet with a conventional low-fat diet on weight loss in overweight adults.

Methods: We searched sources including PubMed, the Cochrane Library, Embase, and Web of Science. We included randomized controlled trials (RCTs) of low-carbohydrate diet and low-fat diet on weight loss. Two review authors independently performed study selection, performed data extraction and assessed the risk of bias in the included trials. The primary outcome was body weight change. We used a random-effects model to calculate mean differences or standardized mean differences for continuous data, with 95% confidence intervals. We assessed the quality of the evidence using Cochrane's risk-of-bias tool.

Results: We included nine studies with 895 participants. This systematic review and meta-analysis demonstrates that low-carbohydrate diets are more efficient than low-fat diets for the treatment of obesity. Low-carbohydrate diets led to greater body weight loss. Furthermore, triglycerides decreased more, and high-density lipoprotein cholesterol levels increased more after low-carbohydrate diets. There was no significant difference between low-carbohydrate diets and low-fat diets for changes in levels of total or low-density lipoprotein cholesterol. The most common adverse events after low-carbohydrate diets were constipation, bad breath, and dry mouth. One participant on the low-carbohydrate diet died from complications of hyperosmolar coma. This was thought to be due to poor compliance with drug therapy for diabetes.

Conclusions: A low-carbohydrate diet is associated with better weight loss than a low-fat diet. But there are more side effects, for example constipation, bad breath, dry mouth.

Effects of physical exercise programs and their FITT Elements on cardiac rehabilitation: a review of systematic review and meta-analysis

Bizzozero-Peroni B[1], Castro-Piñero J[2]

[1] Higher Institute in Physical Education, University of the Republic, Uruguay; [2] Higher Institute in Physical Education, Faculty of Education Sciences, University of Cadiz, Puerto Real, Cádiz; Biomedical Research and Innovation Institute of Cádiz (INiBICA), Cádiz, Spain

Background: Coronary artery disease (CAD) affects 17.5 million people each year, being the leading cause of death worldwide with 7.4 million deaths in 2012 and a forecast of 9.2 million deaths by 2030. More than 7 million people worldwide experience a myocardial infarction (MI) every year, accounting for 10% mortality and 20% re-infarction. Cardiac rehabilitation (CR) is the secondary treatment and prevention of heart disease (Montalescot et al., 2013). The inclusion of a physical exercise program (PEF) within the CR is a main factor for its success (Anderson et al., 2016).

Objectives: To know the influence of PEF and its FITT principle (frequency, intensity, type and time) on the secondary prevention of CR in patients with CAD or MI, or both.

Methods: We carried out a review in PubMed and Web of Science, with articles published to 1 March 2019. We used the 'Assessment of Multiple Systematic Reviews 2' tool (AMSTAR-2) to assess the methodological quality of the included systematic reviews and/or meta-analyses (B. J. Shea et al., 2017). To be selected in this review, the articles had to meet the following inclusion criteria:

1) studies of PEF in CR of people with CAD or MI, or both;
2) to analyze the influence of this PEF on secondary prevention of CR;
3) in English or Spanish full text, published in the selected databases; and
4) in a systematic review or meta-analysis modality.

Results: We identified 3902 articles, of which 19 were selected. Sixteen studies were of moderate-high quality (AMSTAR-2). All studies analyzed the effects of PEF on CR variables for adult patients (mean age range: 48 to 79 years). Three studies are systematic reviews, three are meta-analyses, and 13 contain systematic reviews and meta-analyses. Six studies focused their analyses on patients with CAD, 2 studies on patients with MI, and 11 included both pathologies. PEF (i.e. strength training, high-intensity interval training, moderate-intensity continuous training and tai chi) were beneficial on aerobic capacity, left ventricular ejection fraction, recovery heart rate (HR), resting HR, peak HR, muscle strength, functional mobility, body weight and final systolic / diastolic volume. A duration between 7 and 12 weeks and an intensity greater than 90% HR peak – 80% VO2 peak of high-intensity interval training, and a weekly frequency of 5 sessions, a duration between 6 and 12 months, an intensity of 79% VO2 peak, a session time of 45 minutes and an early start (1 to 12 weeks) of moderate-intensity continuous training presented the best results on CR.

Conclusions: PEF (i.e. strength training, high-intensity interval training, moderate-intensity continuous training and tai chi) were beneficial in the CR of patients with CAD or MI, or both. In addition, we found evidence on the best ranges of the FITT principle of high-intensity interval training and moderate-intensity continuous training for the secondary prevention of these pathologies.

Patient or healthcare consumer involvement: Cardiac rehabilitation patients.

Efficacy and safety of risperidone and paliperdone in schizophrenia and bipolar patients

Hodkinson A[1], Heneghan C[2], Mahtani K[2], Panagioti M[1], Kontopantelis E[1]

[1] University of Manchester, UK; [2] University of Oxford, UK

Background: Risperidone and paliperidone are two mainstay anti-psychotic drugs approved for treating schizophrenia in adults and adolescents, and for the short-term treatment of manic or mixed episodes of bipolar disorder. However, over the last decade, the manufacturers have been involved in a rising number of legal cases because they failed to disclose that the drug may cause hormonal imbalances that could lead to breast tissue development ('Gynecomastia') and infertility in boys and girls.

Objectives: As part of the Cochrane Clinical Study Report Working Group, we aim to systematically review all available evidence of the anti-psychotic drugs risperidone and paliperidone to assess whether they increases the risks of 'Gynecomastia', 'Cerebrovascular' and other serious adverse events in patients with schizophrenia or bipolar disorder. The review will use unpublished data sources including clinical study reports (CSRs) at the YODA project and the European Medicines Agency.

Methods: We have access to relevant randomized placebo-controlled trials at YODA, and further requests are being made to identify other sources of unpublished trials. Because of the novelty and size of CSRs, we will subdivide the extraction, appraisal, and analysis of the data into two stages. In stage 1, we assess the reliability and completeness of the identified trial data in a 'scoping review'. This will allow us to identify missing important text or data, and aid us in determining the completeness of the relevant parts of CSRs. Data will only be included in stage 2 (meta-analysis) if they satisfy the following three criteria:

1) Completeness – they provided adequate information on harms according to our bespoke checklist;
2) internal consistency – all parts of the same CSRs are consistent;
3) external consistency – consistency of data as reported in regulatory documents established by cross-checking.

Results: Fifty-five risperidone and 31 paliperidone (palmitate) placebo-RCTs were eligible. We obtained 18 of the risperidone and 22 paliperidone trials from YODA, where we retrieved both the full CSRs and patient-listings data. We have made further requests at the EMA. On average only 16% of adverse events and 23% of serious adverse events were reported in the journal publications compared to the CSRs and patient safety listing. Patient safety narratives were only provided in 10 (25%) of the CSRs, meaning information on the nature, timing and causality were difficult to determine. Redaction were also recorded.

Conclusions: We provide a comprehensive and rigorous assessment of the safety of two major anti-psychotic drugs for treating schizophrenia and bipolar diagnosed patients. The review is the first to involve unpublished data from CSRs and patient-safety listings and will influence current clinical guidance in this high-priority mental health area.

Patient or healthcare consumer involvement: An advisory group will be set up with PRIMER at the University of Manchester to provide advice regarding patient safety concerns and to discuss the overall study findings and implications.

Handwash versus handrub practices for preventing nosocomial infection in hospital intensive care units: a systematic review and meta-analysis

Bello S[1], Bamgboye E[1], Ajayi D[1], Ossai E[2], Aniwada E[3], Salawu M[4], Fawole O[4]

[1] College of Medicine, University of Ibadan, Nigeria; [2] College of Health Sciences, Ebonyi State University, Abakaliki, Nigeria; [3] College of Medicine, University of Nigeria Enugu Campus, Nigeria; [4] College of Medicine, University of Ibadan, Nigeria

Background: Healthcare associated infections (HCAIs) affect the quality of care and are the most frequent adverse consequences of health care worldwide. Hand hygiene is considered to be the most effective tool in HCAI control. Vigorous handwashing for 40 to 60 seconds, or the use of alcohol handrub before and after every patient contact is recommended to prevent transmission of pathogenic organisms from one patient to the other. Compliance is, however, suboptimal and alcohol handrub has been suggested in busy settings like the intensive care units (ICUs) to improve compliance. There is no evidence on the comparative effectiveness between handwash and handrub strategies.

Objective: To assess the effectiveness of handwash versus handrub for preventing nosocomial infection in hospital intensive care units (ICU).

Methods: We conducted a systematic review of studies conducted in ICUs and indexed in PubMed comparing the clinical effectiveness and adverse events between handwash and handrub groups. The primary outcome was nosocomial infection rates. Secondary outcomes included microbial counts on healthcare providers' hands, mortality rates, patient/hospital cost of treatment of HCAIs, length of ICU/hospital stay, and adverse events. At least two authors independently screened studies and extracted data. We conducted meta-analyses of risk ratios (RR), incidence rate ratios (IRR), odds ratios (OR) and mean differences (MD) using the RevMan 5.3 software.

Results: We included seven studies involving a total of 11,663 patients. Five studies (10,981 patients) contributed data to the ICU-acquired nosocomial infection rates. The pooled IRR was 0.71 (95% confidence interval (CI) 0.61 to 0.82; I^2 = 94%) in favour of handrub. On sensitivity analysis, pooled IRR was 0.39 (95% CI 0.32 to 0.48; 4 studies; 8247 patients; I^2 = 0%) in favour of handrub. The pooled OR for mortality was 0.95 (95% CI 0.78 to 1.61; 4 studies; 3475 patients; I^2 = 39%). The pooled MD for length of hospital stay was -0.74 (95% CI -2.83 to 1.34; 3 studies; 741 patients; I^2 = 0%) days, in favour of handrub. The pooled OR for an undesirable skin effect was 0.37 (95% CI 0.23 to 0.60; 3 studies; 1504 patients; I^2 = 0%) in favour of handrub. Overall quality of evidence was low.

Conclusion: Handrub appeared more effective compared to handwash in ICUs.

Haploidentical versus matched unrelated donor stem cell transplantation for patients with acute leukemia: a systematic review and meta-analysis

Nieto W[1]

[1] Instituto de Evaluación de Tecnología en Salud e Investigación, Peru

Background: It is estimated that two-thirds of acute leukemia patients who require stem cell transplantation do not have a matched sibling donor. In these cases, a stem cell matched unrelated donor (MUD-SCT) is recommended. In low-income countries, however, the lack of transplantation lists and difficulties in accessing MUD-SCT delays the transplantation and increases the risk of complications. A potentially effective alternative for MUD-SCT is the Stem Cell haploidentical transplant (haplo-SCT), which has high probability of recruiting a compatible donor.

Objective: To compare the benefits and damages between MUD-SCT and haplo-SCT in acute leukemia patients.

Methods: We performed a systematic review of randomized controlled trials and controlled cohorts published in scientific journals that compared the benefits and damages between haplo-SCT and MUD-SCT in acute leukemia patients (either lymphoblastic or myeloid). This search was performed in two steps: 1) a search of three search engines (PubMed, Scopus and the Cochrane Central Register of Controlled Trials (CENTRAL), and 2) a search of the list of references from the studies included at step 1. Two review authors independently performed study selection and data extraction. We used GRADE methodology to assess the certainty of evidence for each outcome. Protocol is available at 10.6084/m9.figshare.12089457.

Results: We found no randomized controlled trials. We included eight cohort studies (n = 6595 patients with acute leukemia). One of these studies used propensity scores to match the haplo-SCT and MUD-SCT groups, and reported two MUD-SCT groups: MUD 9/10 and MUD 10/10 (Piamontese, 2017). When the meta-analysis was performed, we found a statistically significant difference that favored MUD-SCT at the global survival (hazard ratio (HR) 0.89, 95% confidence interval (CI) 0.8 to 0.99) and disease-free survival (HR 0.85, 95% CI 0.75 to 0.96). Sensitivity analysis excluding the MUD 10/10 result from the Piamontese study did not find significant statistic differences. However, the results of this study suggest that MUD 10/10 had better results than MUD 9/10. Also, we found a statistically significant difference that favored haplo-SCT at the mortality without relapses outcome (HR 0.77, 95% CI 0.65 to 0.92), and no significant statistic differences in the relapse outcome (HR 0.97, 95% CI 0.85 to 1.10).

Conclusion: Haplo-SCT seems to be similar to MUD-SCT 9/10, even if the results based on the Piamontese study suggest that MUD 10/10 had better results.

Additional file: [LLA]

Hepatitis A immunization in South Africa: using Cochrane Reviews as a part of decision-making

Patterson J[1], Irving G[2], Li YQ[3], Jiang Y[3], Mearns H[1], Pope D[4], Muloiwa R[1], Hussey G[1], Kagina B[1]

[1] Institute of Infectious Disease and Molecular Medicine, University of Cape Town, South Africa; [2] Department of Public Health and Primary Care, University of Cambridge, UK; [3] Centre for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, China; [4] Health Inequalities and the Social Determinants of Health, University of Liverpool, UK

Background: Hepatitis A is a vaccine preventable disease caused by the hepatitis A virus (HAV). Currently, South Africa is classified by the World Health Organization (WHO) as a high hepatitis A endemic region where ≥ 90% of children are assumed to be "naturally immunized" following HAV exposure before the age of 10 years old. In high hepatitis A endemic settings, routine vaccination against HAV is not necessary due to high rates of "natural immunization". Recent anecdotal evidence as well as clinical observations, however, suggest a possible shift from high to intermediate HAV endemicity may be occurring in South Africa. Countries with intermediate HAV endemicity and no routine hepatitis A vaccination program have a high risk of experiencing hepatitis A epidemics and high costs associated with care. Currently, there is no routine vaccination program against HAV in South Africa.

Objectives: The aim of this body of work is to generate evidence for decision making on whether a routine vaccination program against HAV should be introduced into the South African Expanded Program on Immunizations. The objectives of this project include gathering context-specific evidence on the epidemiologic features of hepatitis A, clinical characteristics of the disease, hepatitis A vaccine characteristics and cost of case management. The project will also estimate the future epidemiology of hepatitis A and potential epidemiological and economic impacts of routine hepatitis A vaccination in the country.

Methods: The project's overall methods are informed by the principles of evidence-based vaccinology (EBV) for developing vaccine recommendations. The project includes a mixed-methods approach: systematic reviews, a retrospective clinical folder review, costing and epidemiologic-economic modelling. A Cochrane Review entitled "Hepatitis A immunization in persons not previously exposed to hepatitis A" was conducted as a part of this work to assess the beneficial and harmful effects of pre-exposure hepatitis A vaccines (inactivated and live-attenuated) administered to adults and children versus no intervention, placebo, or any other vaccine. The results from the Cochrane Review were used to shape potential vaccination strategies in South Africa and to parameterize the epidemiologic-economic model.

Results: The findings and the dossier are to be shared with the relevant vaccine policy stakeholders in the country such as the National Advisory Group on Immunization (NAGI). To our knowledge, this is the first time a formalized, transparent framework will be applied to the development of a national vaccination policy in the country.

Conclusions: Along with the development of an EBV recommendation on routine use of hepatitis A vaccines in South Africa, this work will explore the applicability of Cochrane Review findings in the EBV process and to highlight potential advantages and disadvantages of conducting Cochrane Reviews in the EBV process.

Patient or healthcare consumer involvement: No patients were directly included in this study.

How precise are we when assessing imprecision? An analysis of Cochrane systematic reviews published during 2019

Llanos P[1], Arcos C[1], Carvajal-Juliá N[2], Urrea G[2], Franco JVA[3], Meza N[4], Madrid E[5], Pérez-Bracchiglione J[5]

[1] School of Medicine, Universidad de Valparaíso, Chile; [2] Hospital Clínico Magallanes "Dr. Lautaro Navarro Avaria", Chile; [3] Research Department – Cochrane Associate Centre – Instituto Universitario Hospital Italiano, Argentina; [4] Interdisciplinary Centre for Health Studies (CIESAL), Cochrane Associate Centre, School of Medicine, Universidad de Valparaíso, Chile; [5] Interdisciplinary Centre for Health Studies (CIESAL), Cochrane Associate Centre, School of Medicine, Universidad de Valparaíso, Chile

Background: Imprecision, understood as the risk of random error, is one of the five dimensions assessed when rating down the certainty of the evidence for a determined outcome. Along with risk of bias, imprecision is the most common domain associated with the GRADE system for estimating overall certainty of evidence (CoE). However, the rationale for assessing imprecision seems to be inconsistently reported among authors of Cochrane Reviews (CRs).

Objectives: To evaluate the reporting of imprecision and the reasons for downgrading certainty of evidence in CRs of interventions published during 2019.

Methods: This cross-sectional study aimed to analyse the rationale behind imprecision assessments by authors of CRs. We included all CRs of interventions published during 2019 with at least one 'Summary of findings' (SoF) table. We excluded non-intervention reviews (such as qualitative or diagnostic test accuracy reviews), empty reviews, overviews and methodological reviews. We extracted information from the first reported SoF table, including footnotes and comments section, for each outcome, and we summarized findings descriptively using absolute and relative frequencies.

Results: We included 499 CRs. In 438 (87.7%) reviews the authors downgraded the CoE of at least one outcome due to imprecision (Figure 1). Among these, 355 (81.0%) CRs explicitly stated a downgraded CoE due to imprecision, whereas 83 CRs (18.9%) provided a rationale for this but did not explicitly mention the domain imprecision. The most common reasons for downgrading the CoE due to imprecision were "few events/patients or small sample size" (291 CRs, 66.4%), "wide confidence intervals" (181 CRs, 41.3%) and "cross the line of no effect" (150 CRs, 34.2%). Only 48 CRs (10.9%) used the concept of optimal information size. In 32 (7.31%) CRs that explicitly downgraded due to imprecision, the authors did not provide a sufficiently clear rationale for their decision.

Conclusions: Imprecision is a common reason for downgrading CoE among CRs. Authors usually justify this assessment arguing the low number of events or patients, and the width of confidence intervals. However, an important proportion of CRs do not justify the reasons for this downgrading and some reviews provided reasons that might not be adequate.

Patient or healthcare consumer involvement: No patients were involved in the development of this research.

Impact of unconventional natural gas or petroleum hydraulic fracturing "fracking" on human health- a systematic review of the literature

Pizarro AB[1], Yucuma D[1]

[1] Department of Epidemiology and Biostatistics, Pontificia Universidad Javeriana, Colombia

Background: The exploitation of unconventional deposits (EUD) is a controversial issue because it faces economic, social and political interests. There is clear uncertainty about the health impacts of the exploitation. However, potential dangers to human health have been raised as serious as cancer, damage to the central nervous system, the endocrine system, the reproductive system and the respiratory system, as well as irritation of the skin and mucosa. In addition to this, there are doubts about the distribution of risk and benefit

Objectives: To perform a systematic review of the literature on effects of fracking (including other unconventional extraction techniques of oil or gas) on human health.

Methods: We searched PubMed, Embase and Scopus. Free terms used were: ("fracking" OR "hydraulic fracturing") OR ((unconventional OR flowback OR shale) AND (gas OR oil)) without language limits, but limited the search to publications during last the 20 years (search date 6 November 2018). Our literature search in Scopus only included journals classified as Medicine, Biochemistry, Genetics and Molecular Biology, Psychology, Immunology and Microbiology, Nursing, or Health Professions. Two members of the research group performed the initial selection of relevant articles, based on title and abstract, and obtained full-text versions for data extraction. The list of references of each study was analyzed manually to apply a snowball strategy. Due to heterogeneity in the methodology of included references, we did not rate quality of evidence with a standardized method.

Results: The initial search retrieved 2420 references, of which we included 166 articles. Additionally, 124 relevant articles were identified after the "snowball" was applied. Out of 232 articles with geographical reference, 153 (75%) came from the US; 21 (9%) from the UK; 17 (7%) Canada, 5 Germany, 4 Australia, 3 Norway, and one each from Albania, Argentina, China, Japan, Netherlands, Poland, South Africa, South Korea, Spain, and Switzerland. We included original research articles, literature reviews, letters to the editor, commentaries and regulatory agency reports.

Conclusions: Current evidence is not enough to show causal association between adverse human health outcomes and fracking. There is, however, moderate evidence that associates air pollutants released during UET with respiratory health effects. Long-term health effects, such as birth defects, cancer, endocrine disruption and reproductive effects were seen as possible. While current scientific evidence leaves questions unanswered about health impacts, there is no evidence that supports the safety of fracking.

Patient or healthcare consumer involvement: We wanted to look for solid, systematic and quality evidence on the health effects of performing this extraction. We must take into account patients with chronic diseases or populations at risk who live in the surrounding areas. With this information in hand, it is easier for decision-makers to know the impact of this technique on public health and environmental health.

Impact of vaccination on antibiotic usage: a systematic review and meta-analysis

Buckley B[1], Bergman H[2], Skidmore B[3], Klemm E[4], Villanueva G[5], Garritty C[6], Paul M[7], Henschke N[8]

[1] Department of Surgery, University of the Philippines Manila, Philippine General Hospital, Manila; Cochrane Response, Philippines; [2] Cochrane Response, Israel; [3] Independent Information Specialist, Ottawa, Ontario, Canada; [4] Wellcome Trust, London, UK; [5] Cochrane Response, UK; [6] Knowledge Synthesis Group, Ottawa Hospital Research Institute, Canada; [7] Institute of Infectious Diseases, Rambam Health Care Campus; Ruth & Bruce Rappaport Faculty of Medicine, Technion – Israel Institute of Technology, Israel; [8] Cochrane Response, Germany

Background: Overuse and inappropriate use of antibiotics (AB) have been identified as major contributing factors in the rise of antibiotic resistance. Immunization has the potential to reduce AB use through reduction of bacterial disease incidence and of symptom-based prescribing for viral or parasitic diseases.

Objectives: This systematic review, commissioned by the Wellcome Trust, aimed to provide a comprehensive and up-to-date assessment of the evidence relating to the effect of vaccines on AB use.

Methods: We conducted this systematic review and meta-analysis in accordance with the Cochrane Handbook and GRADE recommendations. We searched electronic databases for randomized controlled trials (RCTs) and observational studies (published January 1999 to March 2018) comparing vaccines to placebo, no vaccine or another vaccine. The primary outcome of interest was AB use. We screened abstracts and full texts. Two review authors independently extracted and cross-checked data. We assessed risk of bias in observational studies using the Cochrane Risk of Bias in Non-randomized Studies of Interventions tool or the Cochrane Effective Practice and Organisation of Care suggested risk of bias criteria.

Results: We identified 4980 records; assessed 895 full-text reports; and included 96 studies (24 RCTs, 72 observational). Included studies were overwhelmingly from high-income countries. Of 96 included studies, only 6 were from eastern Asia, 2 from South America and 1 from Africa. AB use measurements varied widely, reducing the potential to synthesise results. From RCTs, there was high certainty evidence that intranasal influenza vaccine reduces days of AB use among healthy adults (1 RCT; n = 4253; rate reduction 28.1%, 95% confidence interval (CI) 16.0 to 38.4); moderate certainty evidence that influenza vaccines probably reduce AB use in children aged 6 months to 14 years (3 RCTs; n = 610; ratio of means 0.62, 95% CI 0.54 to 0.70) and that immunization of children aged 3 to 15 years probably reduces community AB use (1 RCT; n = 10,985 person-seasons; risk ratio 0.68, 95% CI 0.58 to 0.83). There was moderate certainty evidence that pneumococcal vaccination probably reduces AB use in children aged six weeks to six years (2 RCTs; n = 47,945; rate ratio 0.93, 95% CI 0.87 to 0.99) and reduces illness episodes requiring ABs in children aged 12 to 35 months (1 RCT; n = 264; rate ratio 0.85, 95% CI 0.75 to 0.97). Other RCT evidence was of low or very low certainty. The majority of the observational studies did not appropriately adjust estimates of AB use for confounding and were considered to be at critical or high risk of bias.

Conclusions: Although vaccination may reduce AB use, the evidence base is poor, particularly in developing regions. There was a large variety of outcome measures used in the different trials which were considered to report "antibiotic use". Future randomized trials assessing the effect of vaccinations should collect and report standardized measures of AB use.

Patient or healthcare consumer involvement: None.

Investigation of the extent of adverse events reports among systematic review: a case study of medical cannabis

Watanabe AH[1], Hong H[1], Lai NM[2], Chaiyakunapruk N[1]

[1] Department of Pharmacotherapy, University of Utah, Salt Lake City, USA; [2] School of Medicine, Taylor's University, Selangor, Malaysia

Background: Several systematic reviews (SRs) have summarized the potential effectiveness of cannabinoids but it is unclear to what extent safety-related outcomes were incorporated. The PRISMA harms checklist was published in 2016 to improve reporting of safety-related outcomes among SRs but its impact on safety evidence reporting remains unclear.

Objectives: To investigate the extent of adverse events reports among SRs.

Methods: We searched PubMed, Embase, CENTRAL, and AMED for SRs investigating the use of cannabis for medical purposes in publications up to June 2019. We included studies if they assessed clinical outcomes for a cannabis-based intervention. The outcome of interest was the inclusion of safety information in different sections of the SR. Two review authors independently extracted data. We assessed study quality using the Cochrane risk-of-bias tool.

Results: We identified 1491 records and included 121 SRs published between 1999 and 2019. A total of 28 (23.1%) SRs included safety information in the title, 103 (85.1%) SRs included safety information in the abstract section, 71 (58.7%) SRs included safety as part of their objective, 83 (68.6%) SRs described safety outcomes in the method section, 112 (92.6%) SRs reported safety-related outcomes in the results section, 109 (90.1%) SRs included safety in the discussion section, and 86 (71.1%) SRs included safety in the conclusion section. The proportion of SRs reporting safety related outcomes in the discussion sections increased after 2016 from 84.4% (95% confidence interval (CI) 75.5% to 93.3%) to 96.5% (95% CI 91.7% to 101.3%, P = 0.03).

Conclusions: This study provides a better understanding of how adverse events were reported among systematic reviews over time. Our findings may highlight areas for improvement for the conduct and reporting of systematic review works to improve patient's safety.

Methodological and reporting quality of meta-analyses in the field of screening

Zhang Q[1], Hou L[2], Wang Q[3], Ge L[4]

[1] School of Nursing, Lanzhou University, China; [2] Evidence-Based Social Science Research Center, School of Public Health, Lanzhou University, China; [3] Department of Social Medicine and Health Management, School of Public Health, Lanzhou University, China; [4] Evidence-Based Social Science Research Center, Department of Social Medicine and Health Management, School of Public Health, Lanzhou University, China

Background: Screening can detect, eliminate, and diagnose diseases early. Effective screening methods can prompt more accurate diagnosis and avoid the adverse effects. Studies of systematic review and meta-analyses on screening are published in the high-quality journals.

Objectives: To evaluate the quality of meta-analyses about screening according to the Assessment of Multiple Systematic Reviews (AMSTAR) and Preferred Reporting Items for Systematic Reviews and Network Meta-Analyses (PRISMA-NMA) guidelines.

Methods: To evaluate the quality of meta-analyses about screening according to the Assessment of Multiple Systematic Reviews (AMSTAR) and Preferred Reporting Items for Systematic Reviews and Network Meta-Analyses (PRISMA-NMA) guidelines.

Results: This study is ongoing, and results will be presented at the Cochrane Colloquium as available.

Mobile applications in the comprehensive management of acute and chronic complications of patients with diabetes: systematic review of the literature

Yucuma D[1], Pizarro AB[1]

[1] Department of Epidemiology and Biostatistics, Pontificia Universidad Javeriana, Colombia

Background: Diabetes Mellitus (DM) is a chronic disease resulting from a deficit in the production of sufficient insulin or ineffective use of this which prevents adequate glucose metabolism. The prevalence of this condition has increased dramatically in recent years, and the development of complications generates high costs to the system, approximately $ 612 billion. However, this can be avoided with adequate glycemic control and early diagnosis. It has been shown that 80% of the diabetic population are in low-income countries where health resources are limited. However, the use of mobile technologies has increased exponentially, and this allows the use of mobile applications to promote self-care in health to be a feasible and cost-effective strategy.

Objectives: To describe the effectiveness and accessibility of mobile application implementation as a therapeutic alternative to prevent complications within the comprehensive management of patients with diabetes.

Methods: In November 2019, we searched PubMed, Embase, ICTRP, Clinical Trials, BVS, DARE, the Cochrane Database of Systematic Reviews and the Cochrane Central Register of Controlled Trials. We included studies with patients diagnosed with diabetes, whose intervention was mobile applications and studies measuring outcomes of efficacy, safety and acceptance for these among patients. Two review authors performed study selection and data extraction, we analyzed manually the list of references of each study to apply a snowball strategy. Due to heterogeneity in included references methodology, we did not rate the quality of evidence with a standardized method.

Results: Our initial search retrieved 1254 references, of which 67 articles were included. Additionally, 16 articles were identified after the 'snowball' methodology was applied. Out of 84 articles with geographical reference, 36 (42%) came from the US; 12 (14.2%) from Canada; 9 (10.7%) China, 5 (6%) France, 4 Netherlands, 3 Germany, 3 South Korea, 2 Ireland, 2 Norway, 2 Australia and one each from China, Italy, Japan, Singapore, Spain, and Switzerland. We included original research articles and literature reviews.

Conclusions: The utilization of mobile applications in the prevention of complications of DM is a useful and accessible tool for patients and health personnel. In the subjects, there was greater empowerment of their disease and significant improvements in self-care activities and metabolic parameters' control. Despite being a relatively new therapeutic alternative, we found satisfactory evidence to support its use in self-care.

Patient or healthcare consumer involvement: This study is of interest to decision-makers and diagnosed patients because it gives us the basis to implement mobile applications for the prevention of complications of DM. Mobile applications allow the clinician to have a better knowledge of the patient's condition and to adapt the treatment. They also have strong potential for patients as a therapeutic aid for self-care.

Additional file: [PRISMA flowchart](#)

Preferred reporting items for systematic review and meta-analysis protocols and reports: extension for children – PRISMA-P-C and PRISMA-C 2020

Offringa M[1], Farid-Kapadia M[1], Askie L[2], Hartling L[3], Contopoulos-Ioannidis D[4], Bhutta ZA[1], Soll R[5], Moher D[6]

[1] Hospital for Sick Children Research Institute, University of Toronto, Canada; [2] NHMRC Clinical Trials Centre, University of Sydney, Australia; [3] Alberta Research Centre for Health Evidence (ARCHE), University of Alberta, Canada; [4] Pediatrics Stanford University School of Medicine, USA; [5] Neonatal-Perinatal Medicine, University of Vermont, USA; [6] Ottawa Hospital Research Institute, Canada

Background: Reporting of child-centric systematic reviews (SRs), i.e. SRs including children, differs from adult SRs in several key aspects such as descriptions of child-tailored interventions, justifiable comparators, valid outcomes and outcome measures, and separate synthesis for targeted pediatric age subgroups. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses – Protocol (PRISMA-P 2015) Statement and its corresponding statement for reports (PRISMA 2009) does not cover the complexities associated with reporting SR in the pediatric population.

Objectives: To develop PRISMA-P-C (Protocol for Children) and PRISMA-C (Report for Children) statements as extensions to the PRISMA-P and PRISMA Statements, respectively, tailored to the specific requirements of reporting protocols and reports of SRs and meta-analyses (MA) that include newborns, children and adolescents or mixed children and adult populations.

Methods: We developed this reporting guidance using the EQUATOR framework for health research reporting guidelines using the following steps: (1) a scoping review identified potential child-centric items; (2) a SR of child-centric SR protocols and reports published in MEDLINE and Embase between 2010 to 2014 evaluated the clarity and transparency in reporting of child-centric modification and extension items and identified areas where reporting could be strengthened; (3) a web-based Delphi survey of researchers with experience in child-centric SRs contributed input on the inclusion and exclusion of potential reporting items; and (4) an international face-to-face consensus meeting of researchers, clinicians, methodologists experienced in conducting child-centric SRs and MA, and journal editors.

Results: The final PRISMA-P-C checklist features three new extension items to PRISMA-P for reporting child-centric SR and MA protocols: two items focus on rationale for a review in children and specification of eligibility criteria, respectively, asking explicit justifications for intervention, comparator and outcome for targeted pediatric age group; a separate data synthesis for adult and targeted pediatric age groups in "mixed" reviews. Wording in four PRISMA-P items is modified to specify review methods in targeted pediatric age subgroups. The PRISMA-C checklist adds seven extension items. These include a description of subgroup analyses for the targeted pediatric age group(s), justifications for intervention, comparator and outcome for these age group(s), justification for combining adult and pediatric data, if done, and appropriateness of outcome measures (i.e. validity, feasibility, reliability and responsiveness) for each of the targeted pediatric age group(s). Wording in eleven PRISMA items is modified. These extensions and modifications are integrated with the recent PRISMA 2020 statement.

Conclusions: PRISMA-P-C and PRISMA-C provide guidelines for reporting protocols and reports of SRs and MA in newborns, children and adolescents including mixed adult and child reviews.

Patient or healthcare consumer involvement: None.

PRIME-IPD: A systematic method to prepare individual participant data for systematic reviews and meta-analyses

Riddle A[1], Dewidar O[1], Ghogomu E[1], Matthew C[1], Hussein A[2], Trawin J[3], Gaffy M[4], Cousens S[5], Bhutta ZA[6], Tugwell P[7], Wells G[2], Welch V[1]

[1] Bruyere Research Institute, University of Ottawa, Canada; [2] Cardiovascular Research Methods, University of Ottawa Heart Institute, Canada; [3] Research Assistant at Women's Health Research Institute, University of British Columbia, Canada; [4] Hospital for Sick Children, University of Toronto, Canada; [5] Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine (LSHTM), UK; [6] Centre for Global Child Health, The Hospital for Sick Children, Canada; [7] Ottawa Hospital Research Institute, Clinical Epidemiology Program, Canada

Background: Individual participant data (IPD) meta-analysis (MA) enables nuanced effect modification analyses and may help standardize outcome measures across studies. However, there are multiple challenges to preparing IPD for MA, such as differences in outcome variables, differences in data collection methods, incomplete data dictionaries. These obstacles and others result in the consumption of extensive amounts of time and resources to prepare IPD for MA.

Objectives: To describe a systematic approach to preparing data for IPD-MA analysis.

Methods: We reviewed relevant guidance from the Cochrane Handbook for Systematic Reviews of Interventions, Get Real-IPD working group, Cochrane Multiple Interventions group and other published literature. We developed a five-step approach to preparing IPD, through iterative consultation with the advisory board for an IPD-MA and systematic review.

Results: The five steps are: Processing, Replication, Imputation and Merging and Estimation (PRIME). The processing step verifies that the variables of interest are available in the original datasets, identifies missing values and relabels all variables of interest across the datasets to have common variable names. The replication step verifies that the processed dataset is consistent with the analyzed dataset in the published papers, using standardized differences. The imputation step involves an algorithm for how to handle datasets with missing values, including multiple imputation and merging of all imputed datasets. The merging step calls for combining all datasets after dealing with missing data. The final estimation step involves calculating any new variables required for analysis, such as categorical variables for gradations of intensity or severity of outcome variables. The outcomes and value of each of these five steps are illustrated using our systematic review and network meta-analysis of deworming for children.

Conclusion: The purpose of this guidance is to standardize the process of preparing data for IPD-MA. This guidance needs to be evaluated by application to other systematic reviews and meta-analyses.

Patient or healthcare consumer involvement: None.

Repetitive peripheral magnetic stimulation for impairment and disability in people after stroke

Sakai K[1], Yasufuku Y[2], Kamo T[3], Ota E[4], Momosaki R[5]

[1] Department of Rehabilitation Medicine, Memorial Hospital, Japan; [2] Department of Physical Therapy, Kyoto Tachibana University, Japan; [3] School of Health Sciences, Japan University of Health Sciences, Japan; [4] Global Health Nursing, Graduate School of Nursing Science, St. Luke's International University, Japan; [5] Department of Rehabilitation Medicine, University School of Medicine University Hospital, Japan

Background: Repetitive peripheral magnetic stimulation (rPMS) is a non-invasive treatment method that can penetrate to deeper structures with painless stimulation to improve motor function in people with physical impairment due to brain or nerve disorders. The effectiveness and safety of this intervention for people after stroke currently remain uncertain. This review update assessed the effects of rPMS in improving activities of daily living and functional ability in people after stroke.

Objectives: To assess the effects of rPMS in improving activities of daily living and functional ability in people after stroke.

Methods: We searched the Cochrane Stroke Group Trials Register; the Cochrane Central Register of Controlled Trials (CENTRAL), in the Cochrane Library; MEDLINE; Embase; the Cumulative Index to Nursing and Allied Health Literature (CINAHL); PsycINFO; the Allied and Complementary Medicine Database (AMED); Occupational Therapy Systematic Evaluation of Evidence (OTseeker); the Physiotherapy Evidence Database (PEDro); ICHUSHI Web; and six ongoing trial registries. We screened reference lists, and we contacted experts in the field. We included randomized controlled trials (RCTs) conducted to assess the therapeutic effect of rPMS for people after stroke. Two review authors independently assessed studies for inclusion. The same review authors assessed methods and risk of bias, undertook data extraction, and used the GRADE approach to assess the quality of evidence.

Results: We included four trials (three RCTs and one cross-over trial) involving 139 participants. We judged the overall risk of bias across trials as low. Only two trials (with 63 and 18 participants, respectively) provided sufficient information to be included in the meta-analysis. We found no clear effect of rPMS on activities of daily living at the end of treatment (mean difference (MD) -3.00, 95% confidence interval (CI) -16.35 to 10.35) and at the end of follow-up (MD -2.00, 95% CI -14.86 to 10.86). We found no statistical difference in improvement of upper limb function at the end of treatment (MD 2.00, 95% CI -4.91 to 8.91) and at the end of follow-up (MD 4.00, 95% CI -2.92 to 10.92). We observed a decrease in spasticity of the elbow at the end of follow-up (MD -0.48, 95% CI -0.93 to -0.03). rPMS treatment was not associated with improved muscle strength of the ankle dorsiflexors at the end of treatment (MD 3.00, 95% CI -2.44 to 8.44). No studies provided information on lower limb function or adverse events, including death. Based on the GRADE approach, we judged the quality of evidence related to the primary outcome as low.

Conclusions: Available trials provided insufficient evidence to permit any conclusions about routine use of rPMS for people after stroke. Additional trials with large sample sizes are needed to provide robust evidence for rPMS after stroke.

Patient or healthcare consumer involvement: None.

Reporting characteristics of negative pressure wound therapy systematic reviews

Olaiya O[1], Oyesile A[2], Huynh M[3], Alagabi A[4], Mbuagbaw L[5], McRae M[3]

[1] Michael G. DeGroote School of Medicine, McMaster University, Hamilton, Ontario, Canada; [2] University of Western Ontario Faculty of Health Science, London, Ontario, Canada; [3] Division of Plastic Surgery, Department of Surgery, McMaster University, Hamilton, Ontario, Canada; [4] Faculty of Health, York University, Toronto, Ontario, Canada; [5] Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, Ontario, Canada

Background: Negative pressure wound therapy (NPWT) is a medical device used for wound management. Clinical indications for NPWT include chronic and acute wounds, dehisced incisions, diabetic wounds, pressure ulcers, meshed skin grafts, or flaps. Systematic reviews of randomized controlled trials offer the highest level of evidence. Reliable inferences can only be drawn from reviews that are methodologically sound and adequately reported. The quality of previously conducted NPWT systematic reviews is recondite.

Objectives: We aim to compare Cochrane and non-Cochrane reviews and investigate the conduct and reporting characteristics of systematic reviews on the effectiveness of NPWT for treating various wounds.

Methods: We searched MEDLINE, Embase, and the Cochrane Database of Systematic Reviews on 6 January 2020. We included systematic reviews that included patients receiving wound care using NPWT. One review author extracted reporting characteristics, and another independently double-checked them.

Results: We identified 30 systematic reviews for inclusion, after screening 646 title and abstracts and 168 full-text articles. We included five Cochrane Reviews and 25 non-Cochrane reviews. The median year of publication was 2017 (interquartile range (IQR) 2013 to 2019) and 11 studies (36.7%) were published in 2019. Most of the studies were from the Europe (53%) and very few from North America (17%). The included systematic reviews incorporated a median of nine studies (IQR 5 to 17) with a median of 872 patients (IQR 659 to 1689). Very few non-Cochrane systematic reviews (12%, 3 studies) reported registering their protocol. Only one-fifth of the non-Cochrane reviews specified the outcomes that were eligible for inclusion in the review. Nine non-Cochrane reviews (36%) failed to include outcomes related to harm in the review. Each Cochrane systematic review reported publishing a protocol, pre-specifying outcomes, and included adverse event outcomes. Cochrane Reviews demonstrated higher odds of reporting adverse effects (odds ratio (OR) = 33.3, 95% confidence interval (CI) 5.5 to 1355.5). There was no difference in odds of performing meta-analysis between Cochrane and non-Cochrane reviews (OR = 0.390, CI 0.04 to 3.87). Few (16%) non-Cochrane systematic reviews used the GRADE approach to outline the certainty of the evidence. Risk of bias was assessed in 88% of non-Cochrane reviews but was incorporated into the abstract in only 40% of studies.

Conclusions: While outcome data is being pooled and presented in both non-Cochrane and Cochrane Reviews, non-Cochrane reviews on NPWT lack adequate conduct, reporting, and critical appraisal. Novel approaches are required to facilitate better conduct and reporting in non-Cochrane systematic reviews. The results of NPWT non-Cochrane systematic reviews should be used and interpreted with caution by patients or healthcare consumers.

507

SGLT2 inhibitors and diabetic ketoacidosis in patients with type 2 diabetes: systematic review and meta-analysis of randomized controlled trials

Liu J[1]

[1] Chinese Evidence-Based Medicine Center, West China Hospital, Sichuan University, China

Background: The concern continues as to whether Sodium glucose co-transporter-2 (SGLT2) inhibitors can increase the risk of diabetic ketoacidosis (DKA). There is a clear and urgent need for a rigorous evaluation of the risk of DKA.

Objectives: To assess the effects of SGLT2 inhibitors on DKA in patients with type 2 diabetes.

Methods: We searched PubMed, Embase, the Cochrane Central Register of Controlled Trials (CENTRAL) and ClinicalTrials.gov from inception to 13 June 2019 for randomized controlled trials (RCTs) that compared SGLT2 inhibitors with control in patients with type 2 diabetes. Paired review authors independently screened citations, assessed risk of bias and extracted data. We used Peto's method as the primary approach to pool the effect of SGLT2 inhibitors on DKA. Sensitivity analyses with the alternative effect measure (risk ratio) or pooling method (Mantel-Haenszel), the use of continuity correction of 0.5 for zero-event trials, or generalized linear mixed model were conducted. We conducted six pre-planned subgroup analyses to explore heterogeneity. We used the GRADE approach to rate the quality of evidence.

Results: We included 39 RCTs, involving 60,580 patients and 85 DKA events. SGLT2 inhibitors were statistically associated with an increased risk of DKA versus control (SGLT2 inhibitors: 62/34,961 (0.18%) vs. control: 23/25,211 (0.09%), Peto odds ratio (OR) 2.13, 95% confidence interval (CI) 1.38 to 3.27, I^2 = 8%; RD 1.7 more events, 95% CI 0.6 more to 3.4 more events per 1000 over 5 years; high quality evidence) (Figure 1). Sensitivity analyses showed similar results (Table 1). The subgroup analyses by mean age (interaction p = 0.02), length of follow up (interaction p = 0.03) showed larger relative effect among older patients (60 years or over) and those with longer use of SGLT2 inhibitors (over 52 weeks) (Table 2).

Conclusions: High-quality evidence suggests that SGLT2 inhibitors may increase risk of DKA in patients with type 2 diabetes. The apparent differences in treatment effects among patients with different age or follow up were likely, suggesting the advisability of caution in patients with long-term use of SGLT2 inhibitors or older patients.

Patient or healthcare consumer involvement: Patients and healthcare consumer were not involved in setting the research question or the outcome measures, nor were they involved in developing plans for design or implementation of the study. No patients were asked to advise on interpretation or writing up of results. Where possible, results of this meta-analysis will be disseminated to the patient community or individual patients and families through the investigators of this meta-analysis.

Additional files: [Figure 1](); [Table 1](); [Table 2]()

Systematic reviews and meta-analyses on nutrition: a bibliometric analysis

Zhang Q[1], Hou L[2], Wang Q[3], Ge L[4]

[1] School of Nursing, Lanzhou University, China; [2] Evidence-Based Social Science Research Center, School of Public Health, Lanzhou University, China; [3] Department of Social Medicine and Health Management, School of Public Health, Lanzhou University, China; [4] Evidence-Based Social Science Research Center, Department of Social Medicine and Health Management, School of Public Health, Lanzhou University, China

Background: There is increasing evidence to suggest that nutrition can impact both short- and long-term health outcomes. The right nutrition at the right time can help ensure clinically relevant benefits. Systematic reviews and meta-analyses have become more widely accepted by clinicians, researchers and policy makers as a useful tool to critically assess the totality of evidence in nutrition. However, little work has been done to identify the great scientific output in this field. Citation analyses has been regarded as a useful method to evaluate the impact of articles.

Objectives: To identify and analyze the most highly cited systematic reviews and meta-analyses, further analyses the main features of 100 most-cited articles in the field excluding methodology studies.

Methods: We conducted a search to identify all nutritional systematic reviews and meta-analyses ever published in the Web of Science from inception to March 2020. Using the Clarivate Analytics 'Web of Science Core Collection (WoSCC)', we ranked the selected articles in descending order on the basis of their citation counts. Two review authors independently read the abstract of each article on the list. We excluded methodological studies. Finally, we reached a unanimous decision on the list of the top 100 most-cited publications from the included systematic reviews and meta-analyses. We used VOSviewer (Centre for Science and Technology Studies, Leiden University, Leiden, The Netherlands) and CiteSpace (Chaomei Chen, China) to make visualization mapping in this paper.

Results: This study is ongoing, and results will be presented at the Cochrane Colloquium as available.

The effect of active video game training on the overall health in schizophrenia: a systematic review

Yue L[1]

[1] West China Hospital, Sichuan University, China

Background: Active video game is a rapidly developing new research field, and the existing evidence has shown that active video game training can improve the overall health of clinical or non-clinical patients. However, the effect of active video game training on schizophrenia has not been comprehensively evaluated.

Objectives: To systematically evaluate the effects of active video games on the overall health in schizophrenia.

Methods: We searched databases including PubMed, Embase the Cochrane Library, Web of Science CINAHL CBM, CNKI and Wanfang. We included all randomized controlled trials of active video game training in schizophrenia. The quality of articles was appraised by the Cochrane Handbook (5.1.0, and we used RevMan 5.3 software to analyze the data.

Results: We included three studies (90 participants). The results showed that the effect of active video game training and routine training on cognitive function of schizophrenics in 5 to 12 week intervention study shows no significant difference between the two groups, Cognistat score at 5 weeks of intervention (mean difference (MD) 4.7, 95% confidence interval (CI) -1.74 to 11.14, P = 0.15) and MCCB score at 12 weeks of intervention (MD 4.0, 95% CI (-0.23 to 8.23), P = 0.06). The influence of intervention on the cardiopulmonary function of schizophrenic patients in 12 weeks was statistically significant, AF score (MD 2.63, 95% CI 0.22 to 5.04, P = 0.03).

Conclusions: Active Video game training may improve the cardiopulmonary function of schizophrenics to a certain extent, but it has no significant effect on cognitive function. Because the intervention time, frequency and evaluation index are not uniform and the heterogeneity of the studies is high, the reliability of the study may be reduced. More high-quality and large sample studies are needed to further explore the effect of active video game training on the overall health of schizophrenics, especially in cognitive function.

Patient or healthcare consumer involvement: Adults, however defined, with schizophrenia or related disorders, including schizophreniform disorder, schizoaffective disorder and delusional disorder, again, by any means of diagnosis.

The effectiveness of dietary supplements for alopecia: a systematic review and meta-analysis

Li J[1], Li J[1], Li R[1], Yan P[1], Jin Y[1], Li X[1], Chu X[1], Zhang N[1], Yang K[1]

[1] Lanzhou University, China

Background: Alopecia is a common condition among people with a range of causes including nutritional deficiencies.

Objectives: To assess the effects of using dietary supplements on alopecia.

Methods: We conducted a systematic search of relevant articles in the Cochrane Central Register of Controlled Trials, PUBMED, Embase, trial registries, and the reference lists of relevant articles on 30 September 2019. We will include randomized controlled trials (RCTs) focusing on the effectiveness of dietary supplements for alopecia. We used a random-effects model to calculate mean differences (MDs) or standardized mean differences (SMDs) for continuous data, with 95% confidence intervals (CIs). We assessed the risk of bias of included studies using Cochrane's risk-of-bias tool.

Results: We identified 288 papers in the initial search. According to the inclusion and exclusion criteria, we ultimately included eight articles with 588 participants which compared dietary supplements with placebo. The overall quality of the included studies was low or very low for all comparisons. Through the meta-analysis, we found that the terminal hair diameter (MD 0.28, I^2 28%, 95% CI 0.04 to 0.53, $P < 0.05$) and vellus hair count (MD 1.43, I^2 39%, 95% CI 0.32 to 2.54, $P < 0.05$) were significantly increased after dietary supplements treatment compared to the control group. However, compared with the control group, the total hair count, total hair density and terminal hair count of the treatment group did not increase significantly.

Conclusions: Compared with the control group, the total hair count, total hair density and terminal hair diameter of the treatment group did not increase significantly. About self-evaluation of satisfaction and adverse events, we did a systematic review. We don't have enough evidence that dietary supplements are effective and safe for the treatment of alopecia.

The effectiveness of transcranial magnetic stimulation for smoking cessation: a systematic review and meta-analysis of six randomized controlled trials

Guo K[1], Li X[1], Yang K[1]

[1] Lanzhou University, China

Background: Many studies have explored the effectiveness of transcranial magnetic stimulation (TMS) for smoking cessation because of its convenience and operability.

Objectives: To evaluate the effectiveness and safety of TMS for assisting people to quit smoking.

Methods: We searched six databases in December 2019. We included randomized controlled trials (RCTs) of TMS for smoking cessation amongst smokers. We used the Cochrane risk-of-bias tool to assess the risk of bias of the included trials. We used RevMan 5.3 software to perform meta-analysis on the number of daily smoking, the quitting rate, Fagerstrom Test of Nicotine Dependence (FTND) Scores, Positive and Negative Syndrome Scale (PANSS) scores, Montgomery Asberg Depression Rating Scale (MADRS) scores and adverse events. We used the GRADE method to evaluate the quality of evidence.

Results: We included six trials involving 303 smokers. The trials were conducted in China (four trials), Israel and the Czech Republic. All studies were rated as unclear risk of bias. The results of meta-analysis showed that, compared with the sham TMS, TMS significantly decreased the average daily number of cigarettes smoked (6 trials, weighted mean difference (WMD) = -7.35, 95% confidence interval (CI) -8.16 to -6.53), P < 0.05), PANSS score (WMD = -2.85, 95% CI -5.29 to -0.41, P < 0.05), MADRS score (WMD = -3.18, 95% CI -8.22 to -3.34], P < 0.05) and FTND Score (WMD = -0.89, 95% CI -1.30 to -0.48, P < 0.05), and significantly improved the smoking cessation rates (odds ratio (OR) = 3.94, 95% CI 1.09 to 14.19, P < 0.05). The only adverse event reported was a mild headache. As for the quality of evidence, two outcomes (the number of daily smoking and quitting rate) were rated as low, the others were rated as very low.

Conclusions: TMS might provide effective support to people trying to stop smoking. However, this conclusion is based on low-certainty evidence, limited by risk of bias and imprecision. We need higher quality research to explore these conclusions.

The effects of interventions preventing depression in children and adolescents: an overview of systematic reviews

Eidet LM[1], Ludvigsen KH[1], Dahlgren A[1]

[1] Regional Centre for Child and Adolescent Mental Health, Eastern and Southern Norway, Norway

Background: Depression is a relatively common disorder in adolescents, which often lasts into adulthood. Depression in children and adolescents increases the risk of new potential episodes and other mental disorders. In addition, the risk of physical health problems and negative psychosocial outcomes increases. Left untreated, the condition may become chronic. To avoid the burden of the condition and to reduce a potential lifetime of mental health care needs evidence-based prevention strategies is needed.

Objectives: To review and summarize the evidence on effect of interventions preventing depression in children and adolescents by conducting an overview of systematic reviews.

Methods: We searched IN SUM: A database of systematic reviews on effects of child mental health and welfare interventions (www.insum.no). IN SUM includes: the Cochrane Database of Systematic Reviews, Campbell Library, DARE (Database of Abstracts of Reviews of Effects), PsycINFO, MEDLINE, Embase, Evidence Based Mental Health and Web of Science. We also searched the Norwegian Institute for Public Health, the Swedish Agency for Health Technology Assessment and Assessment of Social Services, the Danish Health Authority for Systematic Reviews and the National Institute for Health and Care Excellence (NICE) for evidence-based guidelines. We identified eight relevant systematic reviews concerning children and adolescents (< 18 years) published from 2005 until today investigating the effects of prevention programs for depression in children and adolescents. Two authors independently identified relevant systematic reviews, quality assessed the included reviews using AMSTAR (A MeaSurement Tool to Assess systematic Reviews) and extracted data. All reviews were considered for overlap in PICO. We assessed the certainty of the evidence using GRADE.

Results: We included three systematic reviews. These evaluated prevention programs for children at risk (parents with depression diagnosis), and prevention programs aimed at both children at risk of depression and the general population. Prevention programs for children at risk of developing a depression diagnosis appear to be able to produce a small reduction in the rate of depression diagnosis, and a slight reduction in depression symptoms. Psychological prevention programs targeting children and adolescents in general can also lead to a small reduction in children who develop depression diagnosis and symptoms, as well as increase functioning. The effects appear to diminish over time.

Conclusions: Preventive programs have small but beneficial effects. The clinical importance of these findings should be considered. There is a need for more research on long term effects, to reduce the progression of depression to a long-lasting condition.

Patient or healthcare consumer involvement: This overview is developed with the objective to provide the public and practitioners with reliable summarized evidence of prevention of depression in children and youth. This will enable shared decision making and informed prevention choices.

The effects of interventions to reduce trauma symptoms and PTSD in children and adolescents: a review of systematic reviews

Borren I[1], Eidet LM[1], Bræin M[2], Dahlgren A[1]

[1] Regional Center for Child and Adolescent Mental Health (RBUP East and South), Norway; [2] Regional Resource Center on Violence, Traumatic Stress and Suicide Prevention (RVTS East), Norway

Background: Most people experience a traumatic event in childhood, and one in 5 have experienced violence, bullying, rape or childhood sexual abuse. In Norway, about half of those referred to special health services have experienced at least one potentially traumatizing event, and a substantial share of individuals with a trauma history develop long-term negative consequences. A range of treatment methods are available, and synthesized evidence of their efficiency is warranted in order to offer children the best possible treatment and improve quality of life for this group.

Objectives: The aim of this review of systematic reviews was to synthesize available international high-quality evidence of the effect of various interventions for trauma and post-traumatic stress disorder (PTSD) in children and adolescents (3 to 18 years old).

Methods: We searched for relevant reviews in IN SUM, an extensive database of systematic reviews evaluating interventions for mental health and welfare in children and adolescents, based on eight research databases. We also handsearched systematic reviews developed by NICE, Socialstyrelsen and Sundhedsstyrelsen. We used AMSTAR (A MeaSurement Tool to Assess systematic Reviews) to assess the quality of the systematic reviews, and the GRADE criteria to consider the certainty of the evidence. We defined traumatic events as neglect, abuse, sexual abuse, violence, child genital mutilation, natural/manmade disasters, parental bereavement, accidents, explosions, terror attacks, having experienced war, or being a refugee or asylum seeker. We included reviews that reported treatment outcomes for trauma symptoms or PTSD.

Results: We screened 68 recent publications (published 2014 to 2018) for inclusion, and of these 14 met the stringent inclusion and exclusion criteria and were selected for inclusion. We found 39 different types of interventions, many of which had several treatment comparisons. Trauma-focused cognitive behavioral therapy was found to be effective for children having experienced a wide range of traumas. For some of the traumatized groups, psychoeducation, attachment-based treatments, selective serotonin reuptake inhibitors ((SSRI) combined with cognitive processing therapy), eye movement desensitization and reprocessing (EMDR), school-based interventions or support programs could improve symptoms. Improvement up to 6 years post treatment was found in some groups. Evidence was for some treatments limited by low quality (GRADE) or small sample sizes.

Conclusions: Effective treatments are available for trauma symptoms or post-traumatic stress disorder (PTSD) in children and adolescents. Trauma-focused cognitive behavioral therapy was overall found to be an efficient treatment for trauma symptoms and PTSD of various origins, but other treatments also yielded positive effects.

Patient or healthcare consumer involvement: The current review was performed as part of a project that aims to improve the public and providers' access to reliable evidence on treatments of trauma and PTSD.

What is the experience of family carers of patient with dementia living in community of South Korea?

Seo H[1], Kim SM[2], Park JM[3]

[1] Chungnam National University, South Korea; [2] Donggang University, South Korea; [3] Nambu University, South Korea

Background: Korea became an aged society in 2018 when its elderly population (people over the age of 65) reached 14.3% of the total population. According to the 2018 Korea Dementia Observatory Report, in 2017, among the elderly population, the number of dementia patients was estimated at 705,473 (male: 254,676, female 450,797), and the prevalence of dementia was 10%. In 2018, the 16.2 million family and other unpaid caregivers of people with Alzheimer's or other dementias provided an estimated 18.5 billion hours of unpaid care.

Objectives: To develop new understanding on the experience of family carers of patient with dementia living in the community in South Korea. We will identify studies using four electronic databases. Two review authors will independently select studies, extract data, and perform the quality assessment. We will use a narrative synthesis approach and provide a structured summary.

Methods: We will perform a qualitative evidence synthesis of relevant qualitative research exploring the experience, coping, and attitude of family carers who are looking after a patient with dementia living in the community in South Korea. We will search the following four electronic bibliographic databases: MEDLINE (Ovid), Embase (Ovid), CINAHL, and PsycINFO. There will be no date restrictions. We will include any peer-reviewed primary study using recognized qualitative research methods of both data collection and analysis. We will exclude studies that target the caregivers of dementia patients who are institutionalized in healthcare facilities. Two review authors will independently select studies, extract data, and apply the Critical Appraisal Skills Programme (CASP) quality assessment tool. To synthesize the data, we will use a narrative synthesis approach that involves developing a theoretical model, conducting a preliminary synthesis, exploring relations in the data, and providing a structured summary. At least two review authors will independently apply the thematic framework to extracted data. We will use the GRADE Confidence in the Evidence from Reviews of Qualitative Research tool to evaluate confidence in the synthesis findings.

Results: We anticipate this review will provide a robust understanding of how family or unpaid caregivers might experience and cope with caregiving.

Conclusions: Understanding such experience is critical to determining and implementing the policy strategies aimed at reducing the burden of caregiving, depression, and stress of family caregivers.

Patient or healthcare consumer involvement: None

What is the nature and magnitude of inappropriate healthcare services in Canada? A systematic review

Squires JE[1], Cho-Young D[1], Demery-Varin M[2], Aloisio LD[1], Greenough M[2], Lavis J[3], Bell R[4], Bornstein S[5], Brien S[6], Décary S[7], Dobrow M[8], Estabrooks CA[9], Graham ID[1], Grinspun D[10], Hillmer M[11], Horsley T[12], Hu J[2], Katz A[13], Krause C[14], Légaré F[15], Levinson W[16], Levy A[17], Mancuso M[18], Maybee A[19], Morgan S[20], Nadalin-Penno L[2], Neuner A[21], Rader T[22], Roberts J[23], Teare G[24], Tepper J[25], Widmeyer D[26], Wilson M[27], Grimshaw J[1]

[1] Clinical Epidemiology Program, Ottawa Hospital Research Institute, Canada; [2] Faculty of Health Sciences, School of Nursing, University of Ottawa, Canada; [3] Department of Health Research Methods, Evidence, and Impact, McMaster University, Canada; [4] Ministry of Health and Long-Term Care, Canada; [5] Department of Political Science and Faculty of Medicine, Memorial University, Canada; [6] Public Reports, Health Quality Ontario, Canada; [7] Department of Family Medicine, Faculty of Medicine, Université Laval, Canada; [8] Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Canada; [9] Faculty of Nursing, University of Alberta, Canada; [10] Registered Nurses' Association of Ontario, Canada; [11] Management Division, Ontario Ministry of Health and Long-Term Care, Canada; [12] Royal College of Physicians and Surgeons of Canada, Canada; [13] Manitoba Centre for Health Policy, Department of Community Health Sciences and Family Medicine, University of Manitoba, Canada; [14] British Columbia Patient Safety and Quality Council, Canada; [15] Health and Social Services Systems, Knowledge Translation and Implementation Core, Quebec SPOR Support Unit, Canada; [16] Institute of Health Policy, Management, and Evaluation, Department of Medicine, University of Toronto, Canada; [17] Department of Community Health and Epidemiology, Dalhousie University, Canada; [18] New Brunswick Health Council, Canada; [19] Patients Advisors Network, Canada; [20] Faculty of Medicine, School of Population and Public Health, University of British Columbia, Canada; [21] Health Quality Council of Alberta, Canada; [22] Canadian Agency for Drugs and Technologies in Health, Canada; [23] Patient Collaborator, Canada; [24] Alberta Health Services, Alberta Cancer Prevention Legacy Fund, Population, Public and Indigenous Health, Canada; [25] Department of Family and Community Medicine, Institute of Health Policy, Management and Evaluation, Canada; [26] Manitoba Institute for Patient Safety, Canada; [27] Department of Health Research Methods, Evidence and Impact, McMaster University, Canada

Background: A substantial proportion of health care delivered globally is inappropriate as evidenced by harmful and/or ineffective practices being overused and effective practices being underused. This inappropriate health care leads to negative patient experiences, poor health outcomes, and inefficient use of scarce healthcare resources. The purpose of this study was to conduct a systematic review of inappropriate healthcare practices in Canada.

Objectives: 1) Systematically search and critically review published and grey literature on inappropriate healthcare practices in Canada. 2) Describe the nature and magnitude of inappropriate healthcare in Canada.

Methods: We searched multiple online databases and grey literature sources to identify quantitative studies reporting objective or subjective measurements of inappropriate healthcare practices in Canada. We limited inclusion to studies from January 2007 to September 2019 that reported on large or diverse populations. Two authors independently screened, assessed quality and extracted the data. Study findings were synthesized narratively by overuse and underuse in three care categorizations: tests, treatments, and procedures.

Results: We included 138 studies. The majority of studies were conducted in the acute or specialty sector (n = 74, 54%), followed by primary care (n = 19, 14%). Other sectors represented less frequently included: long-term care, home and community, public health, and rehabilitation. While all Canadian regions were represented, most studies reported data from Ontario (n = 85, 62%) or Alberta (n = 46, 33%). Similar proportions of studies examined underuse (n = 78, 56%) and overuse (n = 72, 52%) of clinical practices. There was wide variation between studies in the magnitude of inappropriate care reported. Underuse of:

- tests ranged from 0.1% (computed tomography imaging for abdominal pain) to 100% (subjective global assessment tool for assessing malnutrition);
- treatments ranged from 1.1% (anti-hyperglycemic medications for Diabetes) to 100% (probiotics before colorectal surgery); and
- procedures ranged from 9.4% (endoscopy for colorectal cancer screening) to 98.1% (carotid endarterectomy/stenting for transient ischemic attack/stroke patients).

Overuse of:

- tests ranged from 0.09% (carotid imaging) to 92.7% (breast cancer imaging);
- treatments ranged from 0.06% (opioid use for dental pain) to 86.8% (polypharmacy (≥ 10 medications) among older adults); and
- procedures ranged from 10.8% (angiography) to 22% (caesarean delivery).

Conclusions: Through this review, we addressed a critical gap in the literature by producing the first-ever evidence-based Canadian compendium of inappropriate healthcare practices. Our findings can be used to advance quality improvement programs and to support agencies dedicated to quality and patient safety in Canada.

Patient or healthcare consumer involvement: Several stakeholders, including patients and representatives from quality improvement organizations across Canada were centrally involved in designing and carrying out this systematic review.

OTHER TOPICS

A review on systems perspectives in systematic reviews

Hong QN[1], Bangpan M[1], Stansfield C[1], Kneale D[1], O'Mara-Eves A[1], van Grootel L[2], Thomas J[1]

[1] EPPI-Centre, Institute of Education, University College London, UK; [2] Tilburg University, The Netherlands

Background: Reviewing complex interventions is challenging because they include a large number of elements that can interact dynamically in a non-linear manner. Also, complex interventions are often nested within a wider system that cannot be fully understood by only using linear causal models and by examining their components in isolation. To address this complexity, it has been advocated to use systems perspectives in systematic reviews. However, it is not clear to what extent systems perspectives provide different analytical possibilities and how to apply them in systematic reviews.

Objective: To explore how systems perspectives have been applied in systematic reviews.

Methods: We performed a mapping review. We searched seven databases (May 2019): MEDLINE, Embase, PsycINFO, CINAHL, Social Sciences Citation Index, Public Health Database, and ABI/INFORM. The search strategy combined key terms on 'systems perspectives' and 'systematic reviews'. We supplemented the search with focussed searching of three online search engines: Google Scholar, Microsoft Academic, and BASE (Bielefeld Academic Search Engine). Two review authors independently screened 5% of the records to clarify the selection criteria, and one review author screened the remaining records. Descriptive synthesis served to summarize the reviews' characteristics (year, country, topic) and data on systems perspectives (framework, systems methods, software, and systems results). We categorised the data on the reasons for using systems perspectives and challenges encountered when using systems perspectives.

Results: We screened 3028 records. We assessed 436 full-text papers for eligibility and retained 112 papers (representing 111 reviews). In general, two categories of papers were identified. First, several reviews mentioned using 'systems lens' (n = 86). A systems framework or theory (e.g. Bronfenbrenner's ecological systems theory, von Bertalanffy's general system theory) was used to frame the topic, generate hypotheses, guide the selection of studies, structure the analysis of the literature and/or interpret the results of the review. Second, a smaller number of reviews (n = 25) used systems methods to analyse the findings from included studies (e.g. systems dynamic modelling, soft systems approach) and/or developed systems models (e.g. causal loop diagrams, systems maps).

Conclusions: This review found examples of application of systems perspectives in reviews. Using systems perspectives offers a holistic way of thinking to better understand complex interventions. They allow for a deeper and broader understanding of the interrelationships, synergies, and feedback-loop interactions between elements within the boundary of a system. These approaches can enhance the relevance of findings from systematic reviews by facilitating understanding of complex situations and producing useful information for patients, stakeholders and decision-makers.

A survey of the registration of COVID-19 clinical trials

Ma Y[1], Zhang J[2], Wang J[2], Yang N[1], Zhang H[1], Chen Y[1]

[1] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Cochrane China Network; Chinese GRADE Center, China; [2] School of Public Health, Lanzhou University;Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Cochrane China Network; Chinese GRADE Center, China

Background: On 11 March 2020, the World Health Organization (WHO) said that the COVID-19 outbreak can be characterized as a pandemic as the SARS-COV-2 virus spreads increasingly worldwide. Following the outbreak of COVID-19, China and even the world have started a number of COVID-19 clinical trials in order to verify the effectiveness and safety of various diagnosis and treatment measures.

Objectives: To survey the registration of COVID-19 clinical trials and analyse their characteristics.

Methods: We will systematically search the following four clinical trial registry platforms up to 25 March 2020: the World Health Organization Clinical Trials Registry Platform (http://www.who.int/ictrp/en/), the International Standard Randomized Controlled Trial Number(ISRCTN) Register (https://www.isrctn.com/), the US National Institutes of Health Trials Register (https://clinicaltrials.gov/) and the Chinese Clinical Trial Registry (http://www.chictr.org.cn/index.aspx) to identify all COVID-19 clinical trials. After selection, we will extract the basic characteristics and perform descriptive statistics: the study setting and funding, sample size, study design, intervention, indication, primary outcome and so on. If the trial is focusing on the effects of drugs, we will also identify the type of drugs. If the trial is a randomized controlled trial, we will also identify the type of blinding. The search, selection and extraction will be conducted by two researchers independently and the discrepancies will be resolved by discussion and consensus to a third researcher.

Results: This survey is ongoing and results will be presented at the Cochrane Colloquium as available.

Conclusions: This survey is ongoing and conclusions will be presented at the Cochrane Colloquium as available.

Patient or healthcare consumer involvement: None.

Application of the GRADE method in public health: a bibliometric study of the 87 systematic reviews in the Cochrane Library

Guo K[1], Li X[1], Yang K[1]

[1] Lanzhou University, China

Background: The GRADE method has been used in many fields and its application in public health can be more conducive to its development.

Objectives: To analyse the main features of the 87 systematic reviews about public health in the Cochrane Library published from 2000 to February 2020.

Methods: We analysed 87 systematic reviews about public health in the Cochrane Library, and analysed the outcomes of each review and the application of the GRADE method. We looked at the number of outcomes evaluated by the GRADE method, and analysed the GRADE classification (high, moderate, low, very low) and promotion or degradation factors of GRADE. We also extracted other data including AMSTAR (A MeaSurement Tool to Assess systematic Reviews) scores for each review and information on year of publication, contributing authors, institutions and countries, study design of included studies, topic of the article and keywords. VOS viewer and Bibliographic Items Co-occurrence Matrix Builder (BICOMB) were used for plotting or statistics of authors and keywords.

Results: The 87 articles of public health in Cochrane Library were published from 2000 to 2020, which were mainly published in 2019 (n = 13), 2016 (n = 11) and 2018 (n = 8). Among 396 authors, the greatest number of articles was associated with three individuals namely Ker K (n = 6), Verbeek J (n = 5) and Pena J (n = 5). Most articles were published in the United Kingdom (n = 33), Australia (n = 14) and the United States (n = 10). The most published institution is London School of Hygiene and Tropical Medicine (n = 6). All reviews included mainly randomized controlled trials (RCTs), others were non-RCTs, interrupted time series and controlled before-after studies. Among 479 unique keywords, infant or child (n = 49), RCTs (n = 46), adult (n = 25), adolescent (n = 19) and female or pregnancy (n = 18) were the most frequently used. Forty-one of the 87 reviews used the GRADE method. The total number of outcomes in all articles was 578, outcomes evaluated by the GRADE method was 498, including high level (n = 23), moderate level (n = 100), low level (n = 178), very low level (n = 197). Most articles have an AMSTAR score of around 9.

Conclusions: This study has revealed that reviews in the field of public health are constantly increasing, especially in recent years (year from 2016 to 2019). The application of the GRADE method to evaluate different outcomes has become more widespread over time.

Assessing the methodological and reporting quality of clinical systematic reviews and meta-analyses published in the leading sports medicine journals

Chu X[1], Lu C[1], Li Y[1], Lu T[1], Yang K[1]

[1] Lanzhou University, China

Background: Systematic reviews and meta-analyses (SR/MAs) provide a comprehensive summary of research studies and are used to assess clinical evidence, form policy and construct guidelines. Although the prevalence of SR/MAs about sports medicine has increased in recent years, significant heterogeneity may exist among these literatures, leading to an uncertain ability to adopt these findings to clinical practice. Clinical practice, guidelines and checklists such as Assessing the Methodological Quality of Systematic Reviews (AMSTAR) and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) can be used to enable authors to publish comprehensive and unbiased data.

Objectives: To systematically review methodological and reporting quality of SR/MAs of interventions published in two high-impact factor sports medicine journals, including *British Journal of Sports Medicine* (BJSM) and *Sports Medicine*.

Methods: We will search SR/MAs published in *BJSM* and *Sports Medicine* from 1 January 2015 to 31 December 2019 through Web of Science Core Collection (WoSCC). The search strategy was as follows: "SO = (British Journal of Sports Medicine) OR SO = (sports medicine)". After identification, two review authors will independently screen the records based on title, abstract and full-text. We will create an electronic database in which study type, author name, publication date, journal, topic, review type (SR/MAs/both), conclusion and contents related to AMSTAR-2 and PRISMA will be recorded. Two review authors will use the AMSTAR-2 and PRISMA checklists to assess and analyse the methodological and reporting quality of included studies. We will use linear regression analysis to evaluate the correlation between basic characteristics, methodological quality and reporting quality. Evidence mapping will present the efficacy of interventions in SR/MAs. All data will be performed and analyzed using Excel 2019 and SPSS version 25.0.

Results: We retrieved 2800 documents from WoSCC. Final results will be available by the time of the Cochrane Colloquium.

Conclusions: Findings of this project will provide information about current methodological and reporting quality of SR/MAs of sports medicine. The relevant researchers should ensure the scientific quality and standardization of SR/MAs and report them according to the PRISMA statement. The synthesized evidence will be presented simply and concisely for informed decision-making.

Patient or healthcare consumer involvement: No

Blood transfusions for treating acute chest syndrome in people with sickle cell disease

Dolatkhah R[1], Dastgiri S[2]

[1] Liver and Gastrointestinal Diseases Research Center, Tabriz University of Medical Sciences, Tabriz, Iran; [2] Tabriz Health Services Management Research Center, Tabriz University of Medical Sciences, Tabriz, Iran

Background: Sickle cell disease is an inherited autosomal recessive blood condition and is one of the most prevalent genetic blood diseases worldwide. Acute chest syndrome is a frequent complication of sickle cell disease, as well as a major cause of morbidity and the greatest single cause of mortality in children with sickle cell disease. Standard treatment may include intravenous hydration, oxygen as treatment for hypoxia, antibiotics to treat the infectious cause and blood transfusions. This is an update of a Cochrane Review first published in 2010 and updated in 2016.

Objectives: To assess the effectiveness of blood transfusions, simple and exchange, for treating acute chest syndrome by comparing improvement in symptoms and clinical outcomes against standard care.

Methods: We searched The Cochrane Cystic Fibrosis and Genetic Disorders Group's Haemoglobinopathies Trials Register, which comprises references identified from comprehensive electronic database searches and handsearching of relevant journals and abstract books of conference proceedings. Randomized controlled trials and quasi-randomized controlled trials comparing either simple or exchange transfusion versus standard care (no transfusion) in people with sickle cell disease suffering from acute chest syndrome. Both authors independently selected trials and assessed the risk of bias. No data could be extracted.

Results: One trial was eligible for inclusion in the review. While in the multicentre trial 237 people were enrolled (169 SCC, 42 SC, 15 Sβ0-thalassaemia, 11Sβ+-thalassaemia); the majority were recruited to an observational arm and only ten participants met the inclusion criteria for randomization. Of these, four were randomized to the transfusion arm and received a single transfusion of 7 to 13 mL/kg packed red blood cells, and six were randomized to standard care. None of the four participants who received packed red blood cells developed acute chest syndrome, while 33% (two participants) developed acute chest syndrome in the standard care arm. No data for any pre-defined outcomes were available.

Conclusions: We found only one very small randomized controlled trial; this is not enough to make any reliable conclusion to support the use of blood transfusion. While there appears to be some indication that chronic blood transfusion may play a role in reducing the incidence of acute chest syndrome in people with sickle cell disease and while offering transfusions may be a widely accepted clinical practice, there is currently no reliable evidence to support or refute the perceived benefits of these as treatment options; very limited information about any of the potential harms associated with these interventions or indeed guidance that can be used to aid clinical decision making. Clinicians should therefore base any treatment decisions on a combination of their clinical experience, individual circumstances and the unique characteristics and preferences of adequately informed people with sickle cell disease who are suffering.

Characteristics of meta-analyses and included trials associated with data contribution to individual participant data meta-analyses of randomized controlled trials

Azar M[1], Benedetti A[1], Riehm KE[1], Imran M[1], Krishnan A[1], Chiovitti M[2], Sanchez T[2], Shrier I[2], Thombs BD[2]

[1] McGill University, Canada; [2] Lady Davis Institute for Medical Research, Jewish General Hospital, Canada

Background: Increasingly, members of the scientific community and stakeholders expect transparency in the conduct and reporting of randomized controlled trials (RCTs). Measures have been implemented to attempt to increase the accessibility and availability of trial data. The individual participant data meta-analysis (IPDMA) is a type of study design which relies on the sharing of data to synthesize raw data from primary studies. In practice, data sharing is largely left at the discretion of study authors.

Objectives: To determine the proportion of RCTs that contributed data to IPDMAs and explore factors associated with data sharing.

Methods: We identified IPDMAs with ≥ 10 eligible RCTs, a documented systematic review of the literature, published references for all eligible RCTs indicating which provided data by searching MEDLINE, Embase, CINAHL and the Cochrane Library from 1May 2015 to 13 February 2017. From each IPDMA, we ascertained if there was a published protocol or a PROSPERO registration, country of the corresponding author, participant population medical condition, and type of intervention assessed. For all eligible RCTs within each IPDMA, we ascertained if the RCT had contributed data, 2015 Thomson Reuters impact factor of the journal where the RCT was published, RCT publication year, RCT funding source and presence of author financial conflict of interest, and the number of participants from the RCT included in the IPDMA. We used mixed effect logistic regression to identify factors associated with data contribution at the IPDMA and at the trial level.

Results: Of 774 eligible RCTs from 35 included IPDMAs, 517 (67%, 95% confidence interval (CI) 63% to 70%) contributed data. Compared to RCTs from journals with low impact factors (0 to 2.4), RCTs from journals with higher impact factors were more likely to contribute data: impact factor 5.0 to 9.9, odds ratio (OR) 2.6, 95% CI 1.37 to 4.86; impact factor 10.0 to 19.9, OR 5.7, 95% CI 3.0 to 10.8; impact factor > 20.0, OR 4.6, 95% CI 1.9 to 11.4. RCTs from the United Kingdom were more likely to contribute data than those from the United States (reference; OR 2.4, 95% CI 1.3 to 4.6). There was an increase in OR per publication year (OR 1.05, 95% CI 1.02 to 1.09).

Conclusions: Country where RCTs are conducted, impact factor of the journal where RCTs are published, and RCT publication year were associated with data contribution in IPDMAs with ≥ 10 eligible RCTs.

Patient or healthcare consumer involvement: Data sharing promotes transparent verification and replication of trial results, ensures that important trial findings are reported, reduces waste in research by avoiding unnecessary repetition of efforts, guides the planning of future trials, and may serve to reduce the frequency and impact of non-publication and selective reporting of trial results. This study provides insight into potential factors associated with data sharing that may guide future interventions and practices to increase open data sharing.

Clinical characteristics of children with COVID-19: a systematic review

Wang Z[1], Zhou Q[2], Wang C[3], Shi Q[2], Lu S[4], Xun Y[1], Estill J[5], Liu E[3], Chen Y[6]

[1] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, China; [2] The First School of Clinical Medicine, Lanzhou University, China; [3] National Clinical Research Center for Child Health and Disorders, Children's Hospital of Chongqing Medical University, China; [4] Sichuan Provincial People's Hospital; School of Medicine, University of Electronic Science and Technology of China, China; [5] Institute of Global Health, University of Geneva; Institute of Mathematical Statistics and Actuarial Science, University of Bern, Switzerland; [6] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, Lanzhou, China

Background: Most guidelines on COVID-19 published so far include recommendations for patients regardless of age. Clinicians need a more accurate understanding of the clinical characteristics of children with COVID-19.

Objectives: To identify the main clinical characteristics of children with COVID-19.

Methods: We searched studies reporting clinical characteristics in children with COVID-19 published until 31 March 2020 in electronic databases, clinical trial registration platforms, Google Scholar and the preprint servers. We screened the literature, extracted the data and evaluated the risk of bias of the included studies. We combined the dichotomous outcomes like symptoms and CT imaging, using single-arm meta-analysis of rate, and for continuous outcomes, like laboratory results, we combined some of the main outcomes, for the studies with at least nine patients, using single-arm meta-analysis of continuous variables. As we expected clinical and methodological heterogeneity in the study design, characteristics of participants, interventions and outcome measures, we used random-effects models. We defined heterogeneity as $P < 0.10$ and $I^2 > 50\%$. We also evaluated the quality of main evidence using the GRADE tool.

Results: Our search retrieved 49 studies, including 25 case reports and 23 case series, with a total of 1667 patients. Our meta-analysis showed that most children with COVID-19 have mild symptoms. Eighty-three per cent of the children were within family clusters of cases, and 19% had no symptoms. The main symptoms in children were fever (48%, 95% confidence interval (CI) 39% to 56%) and cough (39%, 95% CI 30% to 48%). Thirty per cent (95% CI 18% to 42%) of children had both cough and fever. At least 7% had digestive symptoms. The lymphocyte count was below normal level in only 15% (95% CI 4% to 26%) of children which is different from adult patients. Sixty-six per cent (95% CI 55% to 77%) of children had abnormal findings in CT imaging.

Conclusions: Children with COVID-19 are likely to have only mild symptoms, and many children are completely asymptomatic. Fever and cough are the main symptoms of COVID-19 in children. Vomiting and diarrhea occurs less frequently in children. Ground-glass opacity is the most common CT imaging of children. Whereas adults tend to have elevated lymphocyte count at the beginning of the disease, in children the lymphocytes were usually within the normal range. As the characteristics of COVID-19 differ between adults and children in multiple ways, specific criteria for the diagnosis and treatment of COVID-19 in children are urgently needed.

Patient or healthcare consumer involvement: None

Cochrane Methods Support Unit: how can we help?

Moore T[1], Anglemyer A[2], Dwan K[3]

[1] Cochrane Editorial and Methods Department Population Health Sciences Department, Bristol Medical School, University of Bristol, UK; [2] Cochrane Editorial and Methods Department: Preventive and Social Medicine, University of Otago, New Zealand; [3] Cochrane Editorial and Methods Department, UK

Background: Cochrane is recognized internationally as the benchmark for high-quality information about the effectiveness of health care and aims to produce high-quality, relevant, up-to-date systematic reviews. Methods used in Cochrane Reviews are becoming more complex and diverse, as evidenced in the newly updated Cochrane Handbook. Methods once considered novel are quickly becoming mainstream (e.g. network meta-analysis (NMA) and inclusion of non-randomized studies of interventions (NRSI)), and new methods continue to emerge (Risk of bias 2 (RoB2) tool and methods for Systematic reviews WIthout Meta-analysis (SWiM). Evidence has shown that methods in systematic reviews are not always implemented appropriately. Cochrane must strive to maintain its reputation by supporting authors and review teams to use complex and novel methods accurately, while continuing to support those seeking guidance for more standard methods. The Cochrane Methods Support Unit (MSU) was set up in October 2019 to improve the consistency and methodological quality of Cochrane Reviews through support, peer review and training for both new, more complex methods and standard methods. The MSU team comprises three staff: a statistician, an epidemiologist and a methodologist. They work collaboratively with Cochrane Review Group (CRG) Networks by providing advice and support on requests from Associate Editors, Network Support Fellows or Network Senior Editors, and in response to queries escalated from the Community Support Team, Copy Edit Support or from the Editor in Chief directly.

Objectives: To summarise the type, frequency and origin of requests for support from the Cochrane MSU.

Results: In our first five months we have responded in total to 98 queries from all eight CRG Networks. This included comments on 28 reviews and 34 protocols. Requests originated from 23 review groups (44%). Methods commented on included: NMA (n = 34); RoB2 (n = 17); NRSI (n = 8); standard intervention reviews (n = 14); living systematic reviews (n = 2); Risk-of-Bias 1 tool (n = 1); individual patient data (n = 1) and systematic reviews without meta-analysis (n = 1). In addition, we have answered 31 statistical and methodology queries from the networks, Cochrane Fast Track and Cochrane Response. Methods tackled in these queries ranged from NMA (n = 6), standard reviews (n = 14), use of funnel plots (n = 1), split body designs (n = 1), inclusion of cluster and cross-over RCTs (n = 3) and unit of analysis issues (n = 2).

Conclusions: The evolving nature of systematic review methods requires parallel evolution of knowledge within Cochrane, which is evident in the number and range of requests received. The Unit provides a broad range of support to the Cochrane community and is ideally placed to identify common errors and training needs to improve the exemplary use of both standard and more complex methods.

Patient or healthcare consumer involvement: None

Collaboration between Cochrane groups and specialty journals to improve the efficiency and quality of systematic review publication

Menard A[1], Pilkington K[2], Wieland LS[3]

[1] BMC Complementary Medicine and Therapies, UK; [2] University of Portsmouth School of Health and Care Professions, UK; [3] University of Maryland School of Medicine, USA

Background: Many Cochrane groups have relationships with specialty journals in their area of research and practice. Some of these Cochrane groups have established mechanisms to support the publication of high-quality systematic reviews in specialty journals related to their scope.

Objectives: To describe the evolving roles of a Cochrane group currently affiliated with a BMC specialty journal, and the results of this collaboration.

Methods: We will describe the procedures for drawing up an affiliation agreement between a Cochrane group and a journal and deciding upon the scope of Cochrane support to editorial staff and peer reviewers, using as exemplar the relationship between Cochrane Complementary Medicine and *BMC Complementary Medicine and Therapies*. We will describe the Cochrane group participation in establishing reporting requirements, screening submissions, carrying out peer review, and educating editorial and peer review staff on systematic review reporting and methodology, and how the contributions of the Cochrane group to editorial processes evolved over time.

Results: The publication of reporting requirements and screening of new submissions required limited Cochrane staff input but had a correspondingly limited effect on the editorial process. Carrying out full peer reviews for systematic reviews was highly effective in identifying problems with manuscripts but was limited by the volume of submissions and the capacity of the Cochrane group. Efforts to supplement both these approaches by educating editorial and peer review staff on systematic review methodology are ongoing. We will present documented outcomes of all approaches with regard to turnaround times, rejection rates, and quality of published systematic reviews.

Conclusions: The successes and limitations of different approaches to supporting the publication of high-quality non-Cochrane systematic reviews may serve as guidance for other Cochrane groups who wish to formalize relationships with specialty journals. The methods used to achieve these goals should be tailored to the needs of the journal and the capacity of the Cochrane group staff, and the effects of these collaborations should be monitored to establish their effectiveness and optimize their benefits to the editorial process.

Patient or healthcare consumer involvement: None

Cross-sectional study on the quality evaluation criteria of mobile health apps

Wang Q[1], Lai H[2], Guo T[2], Xiao Y[2], Ge L[1]

[1] Evidence-Based Social Science Research Center, School of Public Health, Lanzhou University, China; [2] Department of Social Science and Health Management, School of Public Health, Lanzhou University, China

Background: With the rapid development of wireless communication technology and the popularity of mobile phones, mobile medicine provides a platform for patients to receive personalized medical services and real-time and convenient communication in recent years. Mobile medicine has been widely used in our daily life. Using mobile medicine technology to help people's health has become common. The number of applications of mobile medicine has grown all around the world, but the criterion of quality has not been evaluated.

Objectives: To investigate the tools or checklists of quality evaluation for mobile medicine apps.

Methods: We searched PubMed, Embase, MEDLINE, Chinese National Knowledge Infrastructure (CNKI) and the Cochrane Library to collect systematic reviews and meta-analysis about mobile medicine as fully as possible. We included "mobile medicine", "telemedicine", "e-health", "criterion", "quality evaluation" and extracted the first author, year of publication, name of the application, the standard of quality evaluation and so on. Four review authors independently screened literature and extracted data using EndNote software. Any conflicts in the results are resolved after careful discussion. Statistical analyses will be conducted with SPSS 24.0.

Results: This study is ongoing, and results will be presented at the Cochrane Colloquium as available.

Conclusions: This study is ongoing, and results will be presented at the Cochrane Colloquium as available.

Patient or healthcare consumer involvement: Not applicable.

Current practice in systematic reviews for structuring the PICO for synthesis and the use of synthesis methods other than meta-analysis

Cumpston M[1], McKenzie J[2], Thomas J[3], Brennan S[2]

[1] Monash University and University of Newcastle, Australia; [2] Monash University, Australia; [3] EPPI-Centre, University College London, UK

Background: The synthesis component of systematic reviews is often narrowly considered to include only statistical methods, primarily meta-analysis. However, synthesis is a process, beginning with 1) defining the groupings of populations, interventions and outcome to be compared within the review (the 'PICO for synthesis'), 2) examining the characteristics of the available studies, and 3) applying appropriate synthesis methods from among multiple options. This study examines two intertwined aspects of synthesis (1 and 3) that commonly challenge authors and end users of systematic reviews. First, the 'PICO for synthesis' involves decisions about which studies and outcome data will be combined in each analysis – decisions that affect the review's findings. Second, meta-analysis is not used in around a third of systematic reviews, requiring other methods of summary and synthesis. Where the PICO for synthesis is not clearly defined, or synthesis methods are not applied optimally, this can reduce transparency, replicability and end users' ability to interpret the review's findings. Guidance on both aspects is included in the 2019 Cochrane Handbook for Systematic Reviews of Interventions, but further work is required to explore the feasibility and impact of applying the guidance in practice.

Objectives: To identify and describe current practice in systematic reviews in relation to structuring the PICO for synthesis and methods for synthesis when meta-analysis is not used.

Methods: We randomly sampled systematic reviews of public health and health systems interventions indexed in 2018 in the Health Evidence and Health Systems Evidence databases. Eligible reviews were systematic reviews of primary studies, assessing the quantitative effects of health interventions, including at least two included studies, and published in English. We aimed to select a sample size of at least 100, including reviews both with and without meta-analysis. Two authors independently screened studies for inclusion. One author extracted data on the eligibility criteria (PICO for the review); approaches to grouping populations, interventions and outcomes for synthesis (PICO for synthesis); and the summary and synthesis methods used (e.g. tabulation, visual displays, text-based description and statistical synthesis methods such as combining P values, vote counting based on direction of effect and meta-analysis). A second author undertook independent data extraction for a subsample of reviews. Descriptive statistics were used to summarise the findings.

Results: We retrieved 865 unique records published in 2018 from the databases. We randomly sampled and screened the title and abstract of 166 reviews. We screened the full text of 151 reviews. Full results will be presented at the Colloquium.

Conclusions: This study presents a detailed description of current practice in specifying PICO for synthesis and the use of summary and synthesis methods other than meta-analysis.

Patient or healthcare consumer involvement: None in this study.

Current trends in the treatment of patients with post-stroke unilateral spatial neglect: a scoping review

Umeonwuka C[1], Roos R[1], Ntsiea V[1]

[1] Department of Physiotherapy, University of the Witwatersrand, South Africa

Background: Unilateral spatial neglect (USN) is a leading cause of disability and handicap in stroke survivors affecting functional recovery.

Objectives: To explore the current treatment approaches for patients with post-stroke unilateral spatial neglect.

Methods: We performed a three-step search strategy using the Johanna Briggs Institute's guideline. We searched PubMed, CINAHL, The Cochrane Central Register of Controlled Trial (CENTRAL), SCOPUS, PROSPERO, Johanna Briggs Institute Database of Systematic Reviews and Implementation Reports, Sport Discus and Google Scholar databases. The searches were limited to publications from 1 January 2008 to 31 September 2019. We excluded non-English articles and extracted data using a study-specific charting table. Findings were categorized and descriptively presented, with tables and figures.

Results: We identified 81 studies exploring 29 intervention for post-stroke unilateral spatial neglect symptom amelioration. Studies using randomized controlled trial study design was in preponderance. Intervention categories were prism adaptation and visual scanning, mental practice and mirror therapy, electrical stimulation and robotics, combination therapy, pharmacological therapy and others. Studies combining more than one treatment was in preponderance

Conclusions: A plethora of intervention studies has been explored to ameliorate neglect symptoms post-stroke. Both positive and negative results were obtained. Interventions incorporating more than one treatment show promise as a unilateral spatial neglect treatment strategy.

Patient or healthcare consumer involvement: Results from this review will inform health care practitioners on smarter evidence-based approaches to care for patients.

Deciding when to make a Cochrane Review update the final iteration

Barron Millar E[1], Stoniute A[1], Still M[1], Wallace S[1]

[1] Newcastle University, UK

Background and Objectives: The current iteration of the Cochrane Handbook for Systematic Reviews of Interventions advocates for updating a review when the review question is still relevant, new data or methods available, and incorporating these new data is likely to impact upon the overall findings of the previously published version of the review. It is also important that a review update is designed to answer current questions about policy and practice, but also takes into account areas of importance for the patient population which the review is relevant to.

Methods: Between January and April 2020 we undertook a definitive final update of a Cochrane Review examining the use of anticholinergic drugs versus placebo for overactive bladder syndrome in adults. The review group comprised clinicians who had contributed to previous iterations of the review, the Cochrane Incontinence Editorial base team and Evidence Synthesis Specialists who were new to Cochrane methodology. The updated Cochrane Review was produced using the methods outlined in the 2019 version of the Cochrane Handbook, however subgroup analyses were planned around a published core set of outcome measures developed in collaboration with patient groups (ICHOM standard set for overactive bladder).

Results: Between an unpublished update in 2016 and the 2020 update, the number of randomized controlled trials available to meta-analyse had increased by a third. As a result, cross-over trials and cluster randomized controlled trials were removed from the analysis. Subgroup analyses were by type of anticholinergic for patient perception of cure or improvement, urgency episodes in 24 hours, withdrawal due to adverse events, and adverse events.

Conclusions: When the addition of a large amount of new data to a meta-analysis does not change the overall conclusions of a Cochrane Review, authors can have increased confidence in the longevity of the results and conclusions of the definitive update. Although in other speciality areas the closure of a review has generated concern that it will be detrimental to further research in that area, in fact the results of subgroup analyses can be used to spark new questions for additional research into outcomes of interest to patient populations

Patient or healthcare consumer involvement: The outcome selection was based on a previously published core outcome set for overactive bladder syndrome, which was developed with patient input.

Developing scoring system to inform the reliability of evidence about magnitude of invasive pneumococcal disease in Indian children

Jaiswal N[1], Singh M[1], Singh M[1], Chauhan A[1], Agarwal A[1], Pradhan P[1]

[1] Postgraduate Institute of Medical Education and Research, India

Background: Recently, there has been a shift from consensus guidelines to evidence-informed guidelines in India. The need for reliability evaluation of the evidence bodies is increasing. But there is no clear methodology regarding the reliability of evidence on infectious disease prevalence, unlike the evidence for intervention and diagnosis. We conducted this study to address the questions raised on the prevalence of invasive pneumococcal disease among Indian children by developing a tool to assess the reliability of Indian evidence.

Objectives: To access the reliability of the evidence about magnitude of invasive pneumococcal disease among Indian children

Methods: We used a three-step approach to develop a tool to assess the reliability of the evidence for invasive pneumococcal disease in India. A qualitative study comprised three focus group discussions, including expert clinicians, researchers having experience of conducting systematic reviews on infectious diseases in order to identify relevant constructs to be included in the modified Delphi survey (N = 2). We conducted the survey to develop a framework, build consensus and finally apply the same to the exiting systematic reviews of prevalence of IPD in Indian children. The experts for Delphi were chosen using the convenient sampling methods and were advised to follow the principles of GRADE approach for interventions.

Results: After three rounds of focused group discussions with six subject experts and two rounds of modified Delphi method with relevant experts, we developed a reliability assessment tool with a 10-point scoring system for evidence on magnitude of disease from cross-sectional descriptive studies. The experts followed the GRADE principles and the resultant assessment tool had five domains: survey design, risk-of-bias assessment, indirectness of the evidence variations and adjustments for confounding with a highest score of 10. The evidence body for magnitude of invasive pneumococcal disease in Indian children was subject to the assessment tool and scored low.

Conclusions: The framework developed for descriptive studies on IPD showed that the evidence on Indian children is unreliable and needs more studies with rigorous methodology.

Experiences with Cochrane KT mentoring scheme in the Polish context

Zajac J[1], Bała M[1]

[1] Chair of Epidemiology and Preventive Medicine; Department of Hygiene and Dietetics; Systematic Reviews Unit Jagiellonian University Medical College, Poland

Background: In September 2019 Cochrane launched a project called "Cochrane KT Mentoring Programme". This mentoring pilot scheme was based on pairing people who were less experienced in planning or running knowledge translation activities with those who were experienced at delivering knowledge translation projects. The project aimed to develop and build skills and confidence within the Cochrane community.

Objectives: To describe experiences with the Cochrane KT mentoring scheme in the Polish context.

Methods: Mentoring was based on monthly meetings with the mentor. Meetings included in-depth discussion about actions that were already done, comparisons with the mentor environment and experience and taking actions that were not tested yet.

Results: Individual consultations helped to clearly assess the capacity of the Polish Cochrane team. Brainstorming technique helped to analyze the pros and cons of existing activities and helped to think "outside the box". Regular online meetings encouraged to test new (not checked yet) strategies or recheck old ones that at the beginning were not working and led to new ideas (e.g. talks with editors of local journals or attending meetings with journalists). As a result of monthly meetings, we were able to develop co-operation with The College of Family Physicians in Poland and put on their website information about Cochrane and links to Polish Cochrane websites. We also started a co-operation with a National Chamber of Physicians, which publishes a bulletin sent to all physicians registered in Poland (online and printed version). We also approached the Promotion Office of the Jagiellonian University, where we are affiliated. This activity aims to publish on their website information about Cochrane and other materials that promote evidence-based health care for professional and lay audiences. We started a co-operation with pharmacists which led to workshops about systematic reviews and occasional promotion of our Facebook posts. Talks with the mentor encouraged us to take part in actions for journalists that led to a radio broadcast and talks with the Jagiellonian University Promotion Office.

Conclusions: The scheme helped to improve existing strategies and opened new ideas. Probably the individualization of meetings and careful adjustment to the real capacity of the Polish Cochrane branch resulted in the rapid development of dissemination strategies in a short time.

Patient or healthcare consumer involvement: via co-operation with patients' fan pages.

Facilitators, barriers, and strategies for health-system guidance implementation: a protocol for a critical interpretive synthesis

Wang Q[1], Chu XJ[2], Zhang JY[3], Ma YF[2], Chen YL[2], Yang KH[2], Lavis J[1]

[1] Department of Health Research Methods, Evidence & Impact (HEI), Faculty of Health Sciences, McMaster University; McMaster Health Forum, McMaster University, Canada; [2] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Cochrane China Network, China; [3] School of Public Health, Lanzhou University; Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Cochrane China Network, China

Background: The health systems cover three types of arrangements, that is, governance arrangements, financial arrangements, and delivery arrangements, and the implementation considerations for supporting these arrangements. Some evidence showed that weak health systems could hinder the implementation of effective interventions. As an important input, health system guidance (HSG) could help address the challenges of weak health systems (i.e. strengthen health systems) and provide evidence to support policymaking about health systems. However, there is no framework or tool for HSG implementation.

Objectives: To develop a theoretical framework about the facilitators to, barriers to, and strategies of HSG implementation at different levels (individual, organizational, community and system), further to develop HSG implementation tool, then to assist the use and uptake of global HSG at national or sub-national levels.

Methods: We will conduct a critical interpretive synthesis (CIS) approach to develop a theoretical framework, which aims to identify the facilitating factors and barriers during HSG implementation at the four levels, and the corresponding strategies, and their relationships and connections. We will identify the related literature by searching four types of data sources: electronic databases, organizational websites and conference abstracts, the reference of included studies, and contacting experts. Two review authors will independently assess eligibility for relevance and conduct the quality appraisal. There are no restrictions in the time frame, context, study design or language. We will extract and synthesize key findings using frameworks related to policy development and implementation, health system contextual factors and behavioral changes to develop a framework about HSG implementation. Lastly, we will use the framework to mock support the use of several HSG developed by the World Health Organization.

Results: The HSG implementation framework will cover facilitators, barriers, and strategies at four different levels, i.e. system (including political and health system), organizational, community and individual. For example, the financial incentives might be the strategy to promote the HSG implementation at the organizational level. Based on the mock results, we will summarize the strengths and weaknesses, to refine the HSG implementation framework. We will present the initial results at the conference.

Conclusions: The theoretical framework will help HSG users (such as policymakers) to better understand facilitating factors, barriers and strategies about HSG implementation processes at the individual, organizational, community and system level. Also, the HSG implementation framework will be used to further guide the development of implementation tool.

Patient or healthcare consumer involvement: One of the expected components of the HSG implementation framework will involve the facilitators, barriers, and strategies at the individual level, which will include the consumer (i.e. patients and the public).

Gender diversity analysis in Cochrane Reviews

Zambrano-Achig P[1], Masson-Palacios M[1], Montesinos-Guevara C[1], Viteri-García A[1]

[1] Universidad UTE; Centro Asociado Cochrane de Ecuador; Centro de Investigación de Salud Pública y Epidemiología Clínica, Ecuador

Background: Despite the increasing representation of women in science, they still encounter discrimination, limited opportunities to succeed and bias when compared with men. Gender diversity in science refers to the balance between gender at work, including the activities researchers are involved at and the time and dedication they spend in those tasks. Several organizations understand the importance of gender diversity to bring ideas, beliefs and perspectives from women, non-binary people and men into the team. Therefore, the aim of this work is to analyse the gender composition of Cochrane research groups in order to assess Cochrane's performance in terms of gender diversity. In this analysis, gender diversity only refers to the composition of a research team based on binary gender, male and female.

Objectives: To systematically analyse binary gender composition in Cochrane research groups in order to assess Cochrane's performance in relation to gender diversity.

Methods: We systematically searched in the Cochrane Library for systematic reviews published between 21 February 2019 and 19 February 2020. Three review authors (PZ, MM, CM) extracted data based on authors' names. Extracted data included: authors' gender, authorship (first author and last author), and number of women and men who published in each Cochrane group. Two review authors independently extracted and assessed data.

Results: We retrieved and included 577 published systematic reviews. According to data from authors' name, 51% of team members are men and 49% are women. In terms of first author, 55% are women and 45% are men, while 40% of women and 60% of men published as last author. From the 52 included research groups, most of them include roughly similar number of women and men as authors. However, nine groups showed differences in gender diversity: The percentage of women authors were higher in the Public Health (76%); Drugs and Alcohol (74%); Skin (69%); Multiple Sclerosis and Rare Diseases of the CNS (62%); and Methodology Review groups (63%), while the percentage of men authors were higher in the Injuries (77%), Hepato-Biliary (67%), Urology (72%) and Heart (66%) groups.

Conclusions: This assessment shows that Cochrane composition of research groups is balanced in terms of binary gender diversity. However, some groups show differences in terms of authorship since most women in Cochrane research teams publish as first authors while most men publish as last authors.

Patient or healthcare consumer involvement: Not applicable.

Group authorships in Cochrane had low compliance with Cochrane recommendations

Andersen MZ[1], Fonnes S[1], Andresen K[1], Rosenberg J[1]

[1] Cochrane Colorectal Group, Denmark

Background: Recently, group authorships have become more common. Group authorship describes a situation where the name of a group of people is included in the byline of an article. Historically, however, group authorships have been associated with citation errors and difficulties identifying who could be regarded as an author. Cochrane is a collaboration that publishes high quality systematic reviews and meta-analyses and transparency in authorship should be high. Group authorships in Cochrane have not previously been examined. This study aimed to describe group authorships in the Cochrane Database of Systematic Reviews.

Methods: This study was reported according to STROBE (STrengthening the Reporting of OBservational studies in Epidemiology) guidelines. We screened 8396 reviews from the Cochrane Database of Systematic Reviews for group authorships. We extracted and analysed data from group authorships.

Results: A total of 41 reviews with group authorships were included. Almost half of group authorships, 19/41 (46%), were published from 2015 to 2019. Median publication time (protocol to review) of group authorships was 3.1 years. Of all group authorships, 39% met ICMJE's first authorship criterion, 41% met the second, and 12% met the third criterion. All authors met the three authorship criteria in only two studies.

Conclusion: A low prevalence of group authorships existed in Cochrane Reviews. Reviews with group authorships took median three years to publish, and very few group authorships in Cochrane complied with the ICMJE authorship criteria. Group authorships may lessen the transparency of authorship and thus credit in Cochrane Reviews.

Half of Cochrane Reviews were published more than two years after the protocol

Andersen MZ[1], Gülen S[1], Fonnes S[1], Andresen K[1], Rosenberg J[1]

[1] Cochrane Colorectal Group, Denmark

Background: Cochrane Reviews are known for their rigorous methodology and high quality, and authors must follow specific guidelines to ensure this. Cochrane Reviews can also require a longer time to complete than other systematic reviews. It is important that Cochrane Reviews are up to date as they have a significant influence on clinical guidelines.

Objectives: This study aimed to examine the time from publication of a protocol for a Cochrane Review to publication of the actual Cochrane Review for the entire Cochrane Database of Systematic Reviews.

Methods: This study was reported in line with the STROBE (STrengthening the Reporting of OBservational studies in Epidemiology) guideline. Cochrane Reviews from the Cochrane Database of Systematic Reviews published between 1995 and 2019 were assessed. We extracted characteristics of the reviews and calculated time from publication of protocol to publication of review. These times were grouped for relevant characteristics and visualized through charts and tables to illustrate trends.

Results: Of the total 8201 reviews in the database, we included 6764. The median publication time was two years (range 0 days to 21.7 years). Reviews that were published more than five years after the protocol made up 11% of all included reviews, while 19% of reviews were published within a year. The median publication time for the individual Cochrane Review Groups ranged from 15 to 39 months.

Conclusion: Half of Cochrane Reviews were published later than Cochrane's aim of two years. Furthermore, the Cochrane Review Groups' times from publication of protocol to publication of review varied widely.

Patient or healthcare consumer involvement: This study is important as it highlights an issue with the length of time from consumers' need for an answer to a review question until the availability of the results.

How to disseminate and implement guidelines: a systematic review of guidance documents

Wang Z[1], Shi Q[2], Zhou Q[2], Xun Y[1], Zhao S[3], Chen Y[4]

[1] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China; [2] The First School of Clinical Medicine, Lanzhou University, Lanzhou, China; [3] School of Public Health, Lanzhou University, Lanzhou, China; [4] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, Lanzhou, China

Background: A crucial part of the guideline development process is reaching the target audience in a timely manner and ensuring the guidelines are better used by them. Some guidance documents for guideline development have reported the content of dissemination and implementation. But the process of dissemination and implementation in guidelines differs across guidance documents. Guideline developers want to know the minimum number of steps required to disseminate and implement a quality clinical practice guide. But there is no unified standard.

Objectives: To assess how guidance documents for developing clinical practice guidelines (CPGs) address the dissemination and implementation in clinical guidelines.

Methods: We collected the guidance documents for developing clinical practice guidelines by searching PubMed and Google search engine (Alphabet). We also searched the reference lists of all eligible documents and relevant literatures for additional materials not captured by the aforementioned searches. We included documents that provided guidance on the entire development process of practice guidelines and mentioned dissemination and implementation. We excluded documents that were written by individuals, were outdated versions that had been subsequently updated, or were focused on specific aspects of guideline development (such as updating, systematic reviews, or the GRADE process). Two researchers independently screened records and extracted data. We extracted the title of guidance documents, publication date, development organization, etc. as basic information. For the information on dissemination and implementation, the primary framework was based on the form from the World Health Organization Handbook for Guideline Development (2nd edition), and the information outside the primary framework were supplemented in an iterative way. The information related to dissemination and implementation in the current guidance documents were finally comprehensively summarized and graded according to the frequency of the report.

Results: We retrieved 89 guidance documents, of which 60 reported dissemination and implementation. The detailed information of dissemination and implementation will be presented at the meeting.

Conclusions: The results will be presented at the meeting.

Patient or healthcare consumer involvement: None.

Justifications for labelling Cochrane systematic reviews as stable were diverse and not always clear

Babic A[1], Poklepovic Pericic T[2], Pieper D[3], Puljak L[4]

[1] Institute of Emergency Medicine in Split-Dalmatia County, Split, Croatia; [2] Department of Research in Biomedicine and Health, University of Split School of Medicine, Split, Croatia; [3] Institute for Research in Operative Medicine, Faculty of Health, School of Medicine, Witten/Herdecke University, Cologne, Germany; [4] Center for Evidence-Based Medicine and Health Care, Catholic University of Croatia, Croatia

Background: It is challenging to keep systematic reviews (SR) current, and it is expected that SRs will be continuously updated. For some of its reviews, Cochrane has declared that they are "stable", i.e. not in a need of updating.

Objectives: The aim of this study was to analyze publicly available justifications for stabilizing a Cochrane Review.

Methods: We conducted a methodological (research on research) study. On 28 October 2019 we searched Archie, Cochrane's central system for managing documents and contact details and for tracking the stages of editorial and publishing process, to retrieve list of Cochrane Reviews with publication flag "Version is stable". From the 'What's new' section of the stable Cochrane Reviews in Cochrane Library, we extracted justification for stabilization and categorized them.

Results: We included 545 Cochrane Reviews labelled in Archie as stable. The most common five reasons for stabilization was that 'last search did not identify any potentially relevant studies likely to change conclusions' (N = 99; 18%), followed by 'research area no longer active' (N = 86; 16%), 'review is or will be superseded' (N = 41; 7.5%), 'evidence is conclusive" (N = 35; 6.4%) and 'intervention no longer in general use' (N = 34; 6.2%). In 30 (5.5%) reviews the explanation for stabilization was unclear, and it was not possible to categorize the reason. There were 28 (5.1%) reviews that were stabilized because the review was withdrawn. In 27 (4.9%) review's explanation indicated that they were stabilized because it is unlikely that there will be any new studies for inclusion, in 24 (4.4%) it was indicated that no new studies likely to change the conclusions were expected, while in 22 (4%) it was indicated that a new search within two years is not likely to identify any potentially relevant studies likely to change the conclusions. For the 269 (49%) Cochrane Reviews, we considered that the justification for stabilization was not sufficiently transparent.

Conclusions: Our findings suggest that Cochrane Reviews would benefit from a more uniform and transparent approach to publicly available information about why a Cochrane Review has been declared stable, i.e. not in need of updating. This would be of major importance not only for Cochrane Reviews, but also because the same approach could then be applied to non-Cochrane reviews and the overall standardization in priority research topics.

Patient or healthcare consumer involvement: This was a research methodology study and it did not include patient or healthcare consumers.

Nutritional recommendations for type 2 diabetes: a cross-sectional survey

Hou L[1], Wang Q[1], He J[1], Qin T[1], Cao C[1], Ge L[1]

[1] Evidence-Based Social Science Research Center, School of Public Health, Lanzhou University, China

Background: Several organizations have developed clinical practice guidelines (CPGs) to guide patients with type 2 diabetes for nutrition management. However, these CPG recommendations may be inconsistent, and little is known about their quality.

Objectives: To systematically review the consistency of globally available nutritional recommendations for managing type 2 diabetes and to assess their methodological and reporting quality.

Methods: We searched PubMed, China Biology Medicine disc (CBM), and four main guideline websites for type 2 diabetes CPGs and nutritional recommendations. Four researchers independently assessed their methodological and reporting quality using the Appraisal of Guidelines for Research and Evaluation (AGREE II) instrument and Essential Reporting Items for Practice Guidelines in Healthcare (RIGHT) checklist and extracted nutritional recommendations on managing type 2 diabetes.

Results: Fifteen CPGs involved 65 nutritional recommendations regarding six topics. Finally, eight CPGs (six broad and two nutrition specific guidelines) were classified as recommended for clinical practice. The 15 CPGs adhered to less than 60% of RIGHT checklist items. The AGREE II instrument and the RIGHT checklist should be endorsed and used by CPG developers to ensure higher quality and adequate use of their products. Seven CPGs support lifestyle modification to achieve modest weight loss and energy balance. The general recommendations regarding macronutrient intake for type 2 diabetes range: carbohydrates 45% to 60% total energy, fat: 25% to 35% total energy, protein: 15% to 20% total energy. However, the ideal macronutrient distribution for the management of diabetes may vary, depending on the quality of the various macronutrients, the goals of the dietary treatment regimen and the individual's values and preferences. Three CPGs recommend integration of food and dietary pattern-based approaches such as dietary patterns emphasizing dietary pulses (e.g. beans, peas, chickpeas, and lentils), fruit and vegetables and nuts. Specific functional foods are not recommended by two CPGs. Eight CPGs recommended vitamin or mineral supplementation are necessary unless malnourished. Based on three recommendations given in six CPGs, a cultural context may influence recommendations on drinking alcohol.

Conclusions: The AGREE instrument and the RIGHT reporting checklist should be endorsed and used by CPG developers to ensure higher quality and adequate use of their products. The ideal amount of nutrition for the management of diabetes may vary, depending on the goals of the dietary treatment regimen, the cultural context and the individual's values and preferences.

Patient or healthcare consumer involvement: Not applicable

Pathways between evidence of overuse and de-implementation: a collaboration between Cochrane Sustainable Healthcare and Choosing Wisely

Johansson M[1], Born K[2], Muscat Meng D[3], Levinson W[2]

[1] Cochrane Sustainable Healthcare and Cochrane Sweden, Sweden; [2] Choosing Wisely International, Canada; [3] Cochrane Sustainable Healthcare and Nordic Cochrane Centre, Denmark

Background: Cochrane Sustainable Healthcare is a new Cochrane field focused on tackling medical excess. Choosing Wisely is an international, clinician-led campaign that seeks to advance a dialogue on reducing overuse by developing evidence-based recommendations of unnecessary tests, treatments and procedures. Choosing Wisely campaigns are present or planned in 25 countries globally. Campaigns are shifting from awareness of overuse towards de-implementation of unnecessary medical care identified by recommendations. However, there is a need for evidence to support de-implementation efforts. The objectives of these two groups align and a collaboration offers potential for synergistic effects to promote evidence-informed approaches to reduce overuse.

Objectives: To analyze potential mismatches between Cochrane Reviews and Choosing Wisely recommendations, and through this identify how these two organizations can contribute to each other's collective goals of reducing medical excess and overuse.

Methods: In accordance with a pre-specified protocol, two authors independently assessed the evidence underpinning the Choosing Wisely International top 10 recommendations as well as the available Cochrane Reviews relevant for each of these recommendations. We analyzed mismatches between the recommendations and the reviews. We also identified potential gaps in the Cochrane Reviews. For example; differences in questions asked or outcomes reported compared to those needed to support recommendations about de-implementation. The analysis included both evidence-based approaches to de-implementation, as well as evidence on measuring outcomes and impact of de-implementation.

Results: We aim to arrive at a framework for the pathway between recommendations and evidence of overuse to evidence-based de-implementation approaches. This framework can help guide the work and collaboration between Cochrane Sustainable Healthcare and Choosing Wisely, and identify and prioritize knowledge gaps for de-implementation research.

Conclusions: The primary goals of Cochrane Sustainable Healthcare are to enhance the relevance of primary research and evidence synthesis to tackle medical excess, and to increase the use of that evidence to enable more sustainable healthcare for individual patients and societies. Crucial in this is a close collaboration with stakeholders. The collaboration with Choosing Wisely offers a model of how partnership with aligned organizations can advance collective goals, in this case improving knowledge on de-implementation.

Prevalence of comorbidities in coronavirus disease 2019 (COVID-19): a systematic review and meta-analysis

Xun Y[1], Zhou Q[2], Wang Z[1], Shi Q[2], Chen Y[1]

[1] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, Lanzhou, China; [2] The First School of Clinical Medicine, Lanzhou University, Lanzhou, China

Background: Subjects particularly vulnerable to severe disease may be those with pre-existing medical conditions such as diabetes, cardiovascular diseases, renal failure, obesity, and immunodeficiency due to their relatively weak immunity. Evaluating the prevalence of these chronic conditions is fundamental to mitigate COVID-19 complications. So far, there have been no relevant systematic reviews in this area.

Objectives: To provide a systematic evaluation and a detailed estimate of the prevalence of comorbidities in cases of severe COVID-19. This assessment may aid the public health sector while developing policies for surveillance, preparedness, and response to COVID-19 and its severe outcomes.

Methods: We searched PubMed, Embase, Web of Science, the Cochrane Library, China National Knowledge Infrastructure, Wanfang and China Biology Medicine disc to 6 April 2020 using the search terms (MeSH) "SARS-CoV-2", "COVID-19", "Diabetes, Hypertension, Cardiovascular diseases" OR "Obesity". We limited the search to English and Chinese language articles describing the epidemiological, demographic, and clinical features of COVID-19 cases and reporting the prevalence of a number of chronic diseases in infected adults (age ≥ 18 years). We excluded reports published as review articles, letters, case studies, editorials, conference abstracts, vaccination trials, family-based studies, and articles without abstracts. We extracted data on the prevalence of comorbidities including diabetes, hypertension, cardiovascular disease (CVD)/coronary artery disease (CAD), and obesity, together with clinical symptoms such as cough, fever, shortness of breath, and sore throat, from the identified studies. The primary outcome measure was the prevalence of comorbidities in cases of severe COVID-19. We used Review Manager version 5.3 software to perform a meta-analysis of proportions (with 95% confidence intervals (CI)) for the clinical symptoms and for each of the selected comorbidities. We used a random-effects model since we assumed that the relationship between the comorbidities and severe COVID-19 varies across populations. We examined the presence of heterogeneity among the identified studies (Cochran's Q) and the extent of heterogeneity (I^2) as described previously. We used forest plots to illustrate the prevalence of comorbidities in severe COVID-19 from the selected studies and to inspect the heterogeneity of the individual findings.

Results and conclusions: This study is ongoing, and results will be presented at the evidence summit as available.

Patient or healthcare consumer involvement: None.

Prognostic factors for VTE and bleeding in hospitalized medical patients: a systematic review and meta-analysis

Darzi AJ[1], Karam SG[1], Charide R[2], Ikobaltzeta IE[1], Cushman M[3], Gould MK[4], Mbuagbaw L[5], Spencer FA[6], Spyropoulos AC[7], Streiff MB[8], Woller S[9], Zakai NA[3], Germini F[10], Rigoni M[11], Agarwal A[12], Morsi RZ[2], Iorio A[10], Akl EA[2], Schünemann HJ[10]

[1] Department of Health Research Methods, Evidence and Impact; Michael G. DeGroote Cochrane Canada and McMaster GRADE Centres; McMaster University, Hamilton, Canada; [2] AUB GRADE Center, Clinical Research Institute; American University of Beirut, Beirut, Lebanon; [3] Department of Medicine and Department of Pathology and Laboratory Medicine, Larner College of Medicine; University of Vermont, Burlington, USA; [4] Department of Research and Evaluation; Kaiser Permanente Southern California, Pasadena, USA; [5] Department of Health Research Methods, Evidence and Impact; Michael G. DeGroote Cochrane Canada and McMaster GRADE Centres; McMaster University and Biostatistics Unit, Father Sean O'Sullivan Research Centre; St Joseph's Healthcare; Hamilton, Canada; [6] Department of Medicine, McMaster University; Hamilton, Canada; [7] The Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Department of Medicine, Northwell Health at Lenox Hill Hospital, Manhasset, USA; [8] Division of Hematology, Departments of Medicine and Pathology; The Johns Hopkins University School of Medicine, Baltimore, USA; [9] Department of Medicine, Intermountain Medical Center; University of Utah School of Medicine, Salt Lake City, USA; [10] Department of Health Research Methods, Evidence and Impact; Department of Medicine; Michael G. DeGroote Cochrane Canada and McMaster GRADE Centres; McMaster University, Hamilton, Canada; [11] Healthcare Research and Innovation Program, Health Technology Assessment Unit; Fondazione Bruno Kessler, Trento, Italy; [12] Faculty of Medicine, University of Toronto, Toronto, Canada

Background: There may be many predictors of venous thromboembolism (VTE) and bleeding in hospitalized medical patients, but until now, systematic reviews and assessments of the certainty of the evidence have not been published.

Objectives: To identify prognostic factors for VTE and bleeding in hospitalized medical patients.

Methods: We searched MEDLINE and Embase from inception through May 2018. We considered studies that identified potential prognostic factors for VTE and bleeding in hospitalized adult medical patients. Review authors extracted data in duplicate and independently and assessed the certainty of the evidence using the GRADE approach.

Results: Of 69,410 citations, we included 17 studies in our analysis: 14 that reported on VTE and 3 that reported on bleeding. For VTE, moderate certainty evidence showed a probable association with older age, elevated CRP, D-dimer, fibrinogen levels, tachycardia, thrombocytosis, leukocytosis, fever, leg edema, lower Barthel Index score, immobility, paresis, previous history of VTE, thrombophilia, malignancy, critical illness and infections. For bleeding, moderate certainty evidence shows a probable association with older age, sex, anemia, obesity, low hemoglobin, gastroduodenal ulcers, rehospitalization, critical illness, thrombocytopenia, blood dyscrasias, hepatic disease, renal failure, antithrombotic medication and central venous catheter (CVC). Elevated CRP, a lower Barthel Index, history of malignancy and tachycardia are not included in most VTE risk assessment models (RAMs).

Conclusions: This study informs risk prediction in the management of hospitalized medical patients for VTE and bleeding; it also informs guidelines for VTE prevention and future research.

Recommendations for the development, piloting and use of data extraction sheets in systematic reviews: a review of methodological guidance

Büchter R[1], Gast A[1], Pieper D[1]

[1] Witten/Herdecke University, Germany

Background: Data extraction forms link systematic reviews with the primary research and provide the foundation for appraising, analysing, summarizing and interpreting the body of evidence. This makes their development, pilot testing and application a crucial part of the systematic reviews process. Several studies have shown that data extraction errors are frequent in systematic reviews, especially regarding outcome data. Despite this, data extraction methods receive relatively little attention in the literature.

Objectives: To review the guidance that is available to systematic review authors for the development, pilot testing and application of data extraction sheets.

Methods: We reviewed four types of sources:

1) methodological handbooks of major systematic review organizations (SRO);
2) textbooks on conducting systematic reviews;
3) methods documents from health technology assessment (HTA) agencies and
4) published journal articles on the use of data extraction sheet in systematic reviews.

We retrieved documents in February 2019. We extracted recommendations on the development, pilot testing and application of extraction forms. Items were chosen based on iterative reading of relevant guidance until saturation was reached and personal experience in conducting systematic reviews. One author extracted the data and a second author checked it for accuracy. We will summarize the results of our findings descriptively.

Results: We analysed 4 SRO handbooks, 11 textbooks and 6 HTA documents. We are conducting database searches for journal articles. Preliminary results show that the most common recommendations on form development is that review authors should plan in advance which data to extract; develop or adapt an extraction form custom to their review question; provide instructions on use and make sure to link multiple reports of the same study. While piloting the sheet is often recommended, little information is provided on how this should be done. Regarding the data extraction process, the most frequent recommendation is that data should be extracted by two review authors (mostly independently) and that procedures to deal with disagreements should be in place. Few sources made recommendations on the expertise of the review authors involved, training and reliability assessments.

Conclusions: Overall, our preliminary results suggest a lack of comprehensiveness and consistency of recommendations in many of the reviewed documents. This may be particularly problematic for less experienced review authors. Limitations of our method are the scoping nature of the review and that we did not analyse the internal documents of health technology agencies.

Patient or healthcare consumer involvement: Because this is a descriptive methodological analysis, we did not involve patients or healthcare consumers.

Risk prediction models for in-hospital mortality in acute aortic dissection: a systematic review

Ren Y[1], Huang S[1], Li Q[1], Liu C[1], Li L[1], Tan J[1], Wang W[1], Zou K[1], Sun X[1]

[1] Chinese Evidence-Based Medicine Center, West China Hospital, Sichuan University, China

Background: The potential benefits of management for acute aortic dissection (AAD) depend on the accuracy of prognosis assessment. A variety of prediction models for in-hospital mortality in AAD have been reported in the past decades. These models are distinctly composed of a single or combination of biomarkers, demographic information, and clinical presentations, and showed diverse performances.

Objectives: Previous studies identified several predictors and prognostic models for in-hospital mortality in AAD. Our objective was to identify studies evaluating these predictors and prediction models and to illustrate their performance in predicting in-hospital mortality in AAD.

Methods: We searched for studies in PubMed and Embase until July 2019. Two review authors independently screened records for inclusion, assessed risk of bias, and extracted data. We collected the following information from each eligible study: general study characteristics, predictors used, study population, performance of the model, and likelihood of use in practice.

Results: We identified 9526 reports and included 17. Performance measures were poorly reported as only three studies reported both discrimination value and calibration value. For prediction model, the prediction model using International Registry of Acute Aortic Dissection (IRAD) from multinational data reported good calibration, while EuroSCORE II prediction models did not show good discriminative ability and good calibration. For biomarkers used in the prediction model, discriminatory power varied from 0.58 to 0.95; D-dimer, NLR, and CRP predictors were the most popular biomarkers for predicting in-hospital mortality in AAD. The risk of bias in the domains of participants, predictors, and outcome were rated as low for most studies, but the risk of bias in the domain of sample size and missing data and statistical analysis were rated as high or unclear for most studies.

Conclusions: In-hospital mortality risk prediction in AAD has been modelled, but many of these models are methodologically weak and biomarkers used in the prediction model have highly variable performance across different populations. A new model derived from IRAD from multinational data adding the relevant biomarkers may be required for improved prognostic performance.

Scientometric analysis of highly cited papers in general practice research: 2010 to 2020

Wang Y[1], Wang J[2], Pan B[2], Ge L[1], Chen Y[1]

[1] School of Public Health, Lanzhou University, China, [2] Gansu Provincial Hospital, Lanzhou, China

Background: In the wave of promoting medical-sanitary system reform and attaching importance to medical quality, general practice has become a trend of medical and health development. There has been a rapid increase in research on general practice worldwide since 1999. Although some scholars have researched the status of general practice, there are no published papers systematically analyzed the characteristic of highly cited papers in general practice research, including research hotspots, frontiers, and future challenges.

Objectives: To analyze the characteristics of highly cited papers in general practice research, including research hotspots, frontiers, and prospective.

Methods: We collected the data for this study from the Web of Science core database using a comprehensive electronic strategy from 2010 to 2020. The filtering results are based on the selection of "highly cited in the field" and the paper types are limited to the academic article (Article) and the review (Review). The search strategy was developed in January 2020. There are no limitations on language. We will use Excel 2016 to analyze the overall situation and the number of papers published over time. To assist the analysis and to display the data visually, we will use VOSviewer1.6.10 software to analyze the relationship of the high-frequency keywords, highly productive countries and research institutions and we will use the CitespaceV software to identify co-occurrence keywords and co-cited references, capture keywords, and references with strong citation bursts. We will present the cluster analysis and social network maps for keywords, institutions, and countries.

Results: This study is ongoing and results will be presented at the Cochrane Colloquium as available.

Conclusions: This study is ongoing and results will be presented at the Cochrane Colloquium as available.

Patient or healthcare consumer involvement: Not applicable.

Side effects of treatment of COVID-19 with drugs or medical devices: systematic review and case report

Xun Y[1], Zhou Q[2], Wang Z[1], Shi Q[2], Chen Y[1]

[1] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; WHO Collaborating Centre for Guideline Implementation and Knowledge Translation; Chinese GRADE Center; Cochrane China Network, Lanzhou, China; [2] The First School of Clinical Medicine, Lanzhou University, Lanzhou, China

Background: The treatment of COVID-19 has been in the exploration stage. The initial use of drugs and medical device adjuvant therapy inevitably also produced some side effects. However, it is unclear to what extent the side effects appear.

Objectives: To systematically analyse side effects of the treatment of COVID-19 with drugs or medical devices.

Methods: We searched PubMed, Embase, Web of Science, the Cochrane Library, China National Knowledge Infrastructure, Wanfang and China Biology Medicine disc to 6 April 2020 using the search terms (MeSH) "SARS-CoV-2", "COVID-19", "2019-nCoV" and "2019-novel coronavirus", etc. We will include case reports or case series which report information on side effects of treatment of COVID-19 with drugs or medical devices. We will extract and analyse data on the characteristics of the study, interventions, the information on side effects or adverse reactions.

Results and conclusions: This study is ongoing and results will be presented at the evidence summit as available.

Patient or healthcare consumer involvement: None

The current situation in Cochrane Reviews of traditional Chinese medicine

Wang Q[1], Lai H[2], Guo T[2], Xiao Y[2], Ge L[1]

[1] Evidence-Based Social Science Research Center, School of Public Health, Lanzhou University, Lanzhou, China; [2] Department of Social Science and Health Management, School of Public Health, Lanzhou University, Lanzhou, China

Background: The researchers of Chinese Medicine have conducted an increasing number of systematic reviews in the field of Chinese medicine with themes of drugs and diseases currently. These include some high-quality research that has received attention at home and abroad, which plays a positive role in Chinese medicine reaching out to the world. Cochrane has already published a number of reviews of Traditional Chinese Medicine; however, we do not yet know the status of those reviews.

Objectives: To carry out a cross-sectional study on the current publication status Cochrane Reviews in the field of Traditional Chinese Medicine.

Methods: One author found all the Cochrane Reviews. All authors were divided into two groups and two authors in each group independently screened titles and abstracts of Cochrane Reviews. We included all of the Chinese medicine (Chinese herbal medicine and Chinese patent medicine), acupuncture and massage. We excluded retracted articles. We also excluded the updated and duplicated articles in full-text screening. We extracted the number of trials included, the type of Chinese medicine, the control interventions, quality evaluation and so on. Two authors extracted data independently and the other member checked. We resolved any conflicts in the results after careful discussion. We will conduct statistical analyses with SPSS 24.0.

Results: We included 311 articles in the preliminary screening, 68 were excluded from the full-text screening. We analyzed 243 articles in total, among which 79 (32.51%) were acupuncture. There are 90 (37.04%) articles using GRADE to assess the quality of evidence.

Conclusions: This study is ongoing and we will present all results at the Cochrane Colloquium as available.

Patient or healthcare consumer involvement: Not applicable.

The reporting and methodological quality of meta-analysis related to interventions published in the leading surgery journals: overview and evidence mapping

Ke L[1], Lu C[2], Guo Y[3], Lu T[2], Shen R[1], Yang K[2], Hua Y[1]

[1] Department of Liver Surgery, The First Affiliated Hospital, Sun Yat-Sen University, China; [2] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, China; [3] Organ Transplant Center, The First Affiliated Hospital, Sun Yat-Sen University, China

Background: Meta-analysis (MA) is regarded as the best evidence resource and recommended to guide clinical practice and decision making. But inadequate reporting and low methodological quality influences the reliability and validity of MA.

Objectives: To assess the reporting and methodological quality of paired MAs related to interventions published in two surgery journals with high impact factors: *Annals of Surgery* (Ann Surg) and *British Journal of Surgery* (Br J Surg).

Methods: We searched studies published in Ann Surg and Br J Surg on 15 January 2020 through PubMed database from 1 January 2015 to 31 December 2019. The search terms were as follows: "Ann Surg [Journal]" OR "Br J Surg [Journal]". We downloaded all records from PubMed and imported them into Endnote X9 software. Two authors independently read each title, abstract, full-text and selected paired MAs. We will extract and cross-check data from included paired MAs. The characteristics of the data included: name of first author, number of authors, year of publication, journal, funding, sample size, number of included primary studies, number of included RCTs, type of diseases, details of intervention and control, outcomes, effect size, confidence interval, conclusion and contents related to AMSTAR-2 and PRISMA. Two authors will independently assess the reporting and methodological quality of included paired MAs using PRISMA and AMSTAR-2 tools. Any disagreements will be resolved through discussion. We will use descriptive statistics (frequency, percentage) and radar map to present results. We will also use univariable and multivariable linear regression analysis to explore whether basic characteristics (such as number of authors, year of publication, journal, funding, sample size, etc.) are associated with the reporting quality and methodological quality. We will use Spearman rank correlation coefficient to evaluate the correlation between methodological quality and reporting quality. The evidence mapping will show the plausible benefits or harms of surgical interventions. We will perform all data analyses using Excel 2019 (Microsoft, WA, USA) and SPSS version 25.0 (IBM, Armonk, New York, USA). We will define statistical significance as two-sided $P < 0.05$.

Results and conclusions: We retrieved a total of 4404 records from PubMed. We included 118 paired MAs related to surgical interventions for further analysis. The data extraction and quality assessment are ongoing. We will submit the results of this study to a peer-reviewed journal for publication.

Patient or healthcare consumer involvement: Not applicable.

The research status of eHealth literacy

Wang Q[1], Lai H[2], Guo T[2], Xiao Y[2], Ge L[1]

[1] Evidence-Based Social Science Research Center, School of Public Health, Lanzhou University, China; [2] Department of Social Science and Health Management, School of Public Health, Lanzhou University, China

Background: Health literacy has become a global issue, and patients and individuals have access to health information and educational services by using information technology. The role of electronic media in health promotion has become increasingly important in recent years. Using information technology for health requires eHealth literacy, including the ability to read, use computers, search for information, understand health information, and put it into context. At present, researches on eHealth literacy at home and abroad are gradually maturing and the amount of research is increasing, but the status quo is still unclear.

Objectives: To understand the research status of electronic health literacy at home and abroad, and provide a theoretical basis for further research on residents' electronic health literacy level.

Methods: We systematically searched PubMed, Embase, MEDLINE, the Cochrane Library, CNKI, CBM, VIP, and WAN FANG database. We included "eHealth literacy", "E-health literacy", "electronic health literacy" to search and extracted the first author, year of publication, name of the country, the study population, name of scale, entries of Containing, scores, and so on. Four reviewers independently screened literature and extracted data using EndNote software. Any conflicts in the results are discussed after careful discussion. Statistical analyses will be conducted with SPSS 24.0.

Results: This study is ongoing, and all results will be presented at the Cochrane Colloquium as available.

Conclusions: This study is ongoing, and all results will be presented at the Cochrane Colloquium as available.

Patient or healthcare consumer involvement: Not applicable.

The Use of GRADE in Cochrane Reviews of traditional Chinese medicine

Wang Q[1], Lai H[2], Guo T[2], Xiao Y[2], Ge L[1]

[1] Evidence-Based Social Science Research Center, School of Public Health, Lanzhou University, Lanzhou, China; [2] Department of Social Science and Health Management, School of Public Health, Lanzhou University, Lanzhou, China

Background: The researchers of Chinese Medicine have conducted an increasing number of systematic reviews in the field of Chinese medicine with themes of drugs and diseases currently. These include some high-quality research that has received attention at home and abroad, which plays a positive role in Chinese medicine reaching out to the world. GRADE is widely used in literature quality assessment, but the application status in the Cochrane Reviews of traditional Chinese medicine is unclear.

Objectives: To carry out a cross-sectional study on the application status of GRADE in Cochrane Reviews of traditional Chinese medicine.

Methods: One author found all the Cochrane Reviews. All authors were divided into two groups and two authors in each group independently screened titles and abstracts of Cochrane Reviews. We included all of the Chinese medicine (Chinese herbal medicine and Chinese patent medicine), acupuncture and massage. We excluded retracted articles. We also excluded the updated and duplicated articles in full-text screening. We extracted the comparison group, the number of studies, the number of participants, downgraded entry of GRADE and the reason for downgrading and so on. Two authors extracted data independently and the other author checked. We resolved any conflicts in the results after careful discussion. We will conduct statistical analyses with SPSS 24.0.

Results: We included 243 articles and 90 (37.04%) articles using GRADE to assess the quality of evidence. We extracted 649 outcomes of traditional Chinese medicine treatment, including 500 (77.04%) were downgraded due to the risk of bias, 451 (69.49%) were downgraded due to imprecision, and 96 (14.79%) were downgraded due to inconsistency. Indirectness and publication bias accounted for 5.55% and 2.16%.

Conclusions: This study is ongoing and we will present all results at the Cochrane Colloquium as available.

Patient or healthcare consumer involvement: Not applicable.

Variation of work participation outcomes and measurement methods in trials

Ravinskaya M[1], Verbeek J[1], Langendam M[1], Verstappen S[2], Madan I[3], Kunz R[4], Hulshof C[1], Hoving J[1]

[1] Amsterdam UMC, The Netherlands; [2] University of Manchester, UK; [3] Guy's and St Thomas' NHS Trust, UK; [4] University of Basel, Switzerland

Background: Work participation is an important outcome in research which includes people of working age and in particular occupational health research. However, there are no standardized definitions of measurements of common work outcomes such as return to work or sickness absence. As a first step in developing a core outcome set for work participation we undertook a systemic review with the aim of establishing: 1) Which outcomes, definitions of outcomes and outcome measurement instruments are used in trials of interventions that aim to promote work participation either directly or indirectly, in workers with a health problem? 2) What rationale, theory, perspective or framework for selection of outcomes and measurement instruments are reported in these trials?

Methods: We searched MEDLINE, Embase, PsycINFO and the Cochrane Library for randomized controlled trials published between 1 January 2014 and 21 May 2019. Trials were included on any type of intervention which measured work participation outcomes, with participants of the working age, who were either currently employed or seeking competitive employment. We screened 10,222 abstracts and 819 full text articles. A total of 500 articles were eligible for inclusion.

Results: We found a large amount of heterogeneity in outcome reporting. Authors rarely reported on why certain outcome measurement methods were chosen. Results show that outcomes are distributed around four main outcome categories: employment status; employability; absence from paid work and at-work productivity loss.

Conclusion: The results of our review will be used as input for developing an internationally agreed, standardized core outcome set for work participation.

Variations in approaches to mixed methods review: a case study

Hong QN[1], Rees R[1], Sutcliffe K[1], Thomas J[1]

[1] EPPI-Centre, Institute of Education, University College London, UK

Background: Mixed methods reviews have been advocated to provide a complete and rich understanding of complex phenomena to facilitate decision-making. However, conducting this type of review is challenging due to the diversity of included studies. A better understanding is needed on how and why to conduct this type of review, and how to integrate the different components.

Objective: To examine a body of reviews in order to understand variations and similarities in approaches to mixed methods reviews.

Methods: We conducted a case study to describe the mixed methods review process used at the Department of Health and Social Care Reviews Facility in England. The methods consisted of document analysis. The data extraction and analysis focused on the: steps followed, questions addressed, reasons for conducting a mixed methods review, types of evidence and sources used, and integration strategies used. The analysis used existing frameworks from the literature on mixed methods research.

Results: We identified a total of 31 reviews published between 1999 and 2019. Different types of questions were found that addressed stakeholders' views, intervention processes, and/or intervention effectiveness. The mixed methods questions aimed at exploring intervention effectiveness or appropriateness, identifying critical intervention features, quantifying the effect of critical intervention features, and making recommendations about future research. Twelve reasons for performing mixed methods reviews were found: 1) completeness, 2) contextual understanding, 3) credibility, 4) different research questions, 5) diversity of views, 6) enhancement, 7) explanation, 8) process, 9) triangulation, 10) utility, 11) framework development, and 12) promising interventions identification.

The reviews used five main sources of evidence: 1) formal evidence from primary studies, 2) informal evidence, 3) policy documents, 4) systematic reviews, and 5) consultations with stakeholders. The consultations with stakeholders used different methods (e.g. workshops, interviews, surveys) and aimed at understanding the views of stakeholders to inform the analysis and interpretation of the review findings. Different integration strategies for comparing findings, connecting phases and/or assimilating data were used to achieve these aims.

Conclusions: We identified significant variation across the body of mixed methods reviews examined. The review process was bespoke and driven by the questions, needs and concerns of the stakeholders as well as available evidence, resources and time. The reviews we analysed covered different types of questions (e.g. what, how), evidence (e.g. views, effectiveness, process), sources (e.g. primary studies, stakeholders' consultation), perspectives (e.g. patients, clinicians), and synthesis methods (e.g. meta-analysis, thematic synthesis). This study also suggests broadening the conceptualization of mixed methods reviews to take into account a variety of sources and types of evidence.

Views towards publishing protocols for systematic reviews as peer-reviewed articles: an international survey of authors of non-Cochrane systematic reviews

Rombey T[1], Puljak L[2], Allers K[3], Ruano J[4], Pieper D[1]

[1] Institute for Research in Operative Medicine, Witten/Herdecke University, Germany; [2] Center for Evidence-Based Medicine and Health Care, Catholic University of Croatia, Croatia; [3] Department of Health Services Research, Carl von Ossietzky University Oldenburg, Germany; [4] Department of Dermatology, Reina Sofía University Hospital, Spain

Background: To ensure that systematic reviews (SRs) are truly systematic it is essential to define their methods a priori, e.g. by registering them in PROSPERO, the international prospective register for SRs. As PROSPERO records are not peer-reviewed or indexed in bibliographic databases, it might be useful to publish a protocol for the SR as a peer-reviewed article in addition to registering the SR in PROSPERO. For Cochrane Reviews, publishing a protocol is mandatory, but for most non-Cochrane SRs it is optional.

Objectives: To explore views of authors of non-Cochrane SRs registered in PROSPERO towards publishing SR protocols as peer-reviewed articles.

Methods: We invited the contact persons of all PROSPERO records for non-Cochrane SRs registered in 2018 (N = 12,531) to participate in an anonymous five-minute online survey that was administered through SurveyMonkey. The main question addressed SR authors' views towards publishing SR protocols as peer-reviewed articles. Data were analyzed descriptively.

Results: In total, 4223/12,531 (33.7%) invitees responded, of which 3739/4223 (88.5%) completed the survey. Almost half of the respondents had published or planned to publish a protocol for the SR described in their PROSPERO record as a peer-reviewed article (1811/4,054; 44.7%). The remaining 2243/4054 (55.3%) respondents stated that there was no published protocol for their SR and they have not tried or plan to publish one. Of those, 66.4% (1456/2192) stated that there was an unpublished protocol that they followed during the conduct of their SR. The respondents had inconsistent views towards publishing protocols as peer-reviewed articles; most agreed that external feedback from peer-reviewers increases SR quality (2899/3739; 77.5%) but at the same time agreed that publishing a protocol in a peer-reviewed journal is not necessary if the SR is registered in PROSPERO (2399/3739; 64.2%). Respondents' views towards acceptable manuscript processing times for SR protocols were consistent but far below actual manuscript processing times for SR protocols.

Conclusions: Although PROSPERO records are not peer-reviewed, many SR authors seem to consider registration in PROSPERO sufficient. Hence, awareness about the benefits of additionally publishing a SR protocol as a peer-reviewed article should be raised. Our findings might be useful to various stakeholders of SRs; for example, they could support the decision-making of funding agencies on formally requiring published protocols, of SR authors on publishing a protocol as a peer-reviewed article, or of publishers on transforming the way manuscripts for protocols are being processed.

Patient or healthcare consumer involvement: As this was a research-on-research study, which had no health-related outcome, we did not involve patients or healthcare consumers. However, as this was a survey of SR authors, we actively involved the people that our research was about and is intended ultimately to benefit.

What is a meta-epidemiological study? Analysis of published literature showed heterogeneous study designs and definitions

Puljak L[1], Lovric Makaric Z[2], Buljan I[3], Pieper D[4]

[1] Center for Evidence-Based Medicine and Health Care, Catholic University of Croatia, Croatia; [2] Department of Epidemiology, Croatian National Institute of Public Health, Croatia; [3] Department of Research in Biomedicine and Health, University of Split School of Medicine, Croatia; [4] Institute for Research in Operative Medicine (IFOM), Witten/Herdecke University, Germany

Background: The term "meta-epidemiology" is relatively recent; it first appeared in literature in 1997, and in 2002, a methodological guidance was published for "meta-epidemiological" studies that evaluate the effect of trial characteristics on effect sizes. However, it has been reported that the methodology of "meta-epidemiological" studies is not standardized and that terminology used for such studies varies. Therefore, it appears that there is an ambiguity in the understanding and use of the term meta-epidemiological study in the research community.

Objectives: To analyze descriptors and definitions of meta-epidemiological studies in published literature, as well as study design of articles that were self-described as meta-epidemiological studies.

Methods: This was a primary methodological (research-on-research) study in which the units of analysis were published manuscripts. We searched MEDLINE and Embase on 6 August 2019. We extracted definitions of meta-epidemiological studies. From studies self-identified as meta-epidemiological, we extracted their aim, description of the study designs, statistical methods used, unit of analysis, whether they had made their study protocol publicly available (and where), and whether they mentioned that they used any reporting guideline/checklist to report their study.

Results: We included 175 information sources in the analysis. Definitions of meta-epidemiological studies varied, and some studies used the term meta-epidemiological study to describe methodological research-on-research studies. There were 127 (73%) full-text journal articles that authors self-identified with an expression indicating that this was "meta-epidemiological" study or that "meta-epidemiological" analysis was performed, although with varied terminology. Definitions and descriptions of meta-epidemiological studies were found in 40 (25%) records. Some of them defined them broadly as research that examines influence of trial/study characteristics on effect estimates, while for some it appeared that the definition described completely different designs (e.g. surveys). Less than a half of those studies (n = 54, 42.9%) used the two-step meta-epidemiological approach in data analysis. Among studies self-labelled as meta-epidemiological, 9.4% reported registration in PROSPERO, and 11% indicated they reported the study in line with PRISMA.

Conclusion: Authors of published literature use heterogeneous definitions and descriptors for meta-epidemiological studies. Methodological research-on-research studies are also labelled as meta-epidemiological. The research community would benefit from consensus about the definition of a meta-epidemiological study.

Patient or healthcare consumer involvement: This was a research methodology study and it did not include patient or healthcare consumers.