Measurement and Management of Proximal Hamstring Tendinopathy

Anthony Michael Nasser
Bachelor of Applied Science (Physiotherapy)
Master of Physiotherapy (Sports)

This thesis is submitted in total fulfilment
of the requirements for the degree of
Doctor of Philosophy

School of Allied Health, Human Services and Sport
La Trobe University
Victoria, Australia

August 2022

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

A	Adequate
ADL	Activities of daily living
ASAS	Assessment of SpondyloArthritis International Society
AUC	Area under curve
AWB	Autologous whole blood
CFA	Confirmatory factor analysis
CI	Confidence interval
COMET	Core Outcome Measures in Effectiveness Trials
COS-GT	Core Outcome Set – Gluteal Tendinopathy
COS-PHT	Core Outcome Set – Proximal Hamstring Tendinopathy
COSMIN	Consensus-based Standards for the selection of health Measurement Instruments
CSI	Corticosteroid injection
CTT	Classical test theory
D	Doubtful
DIF	Differential item functioning
EP	Expert Physiotherapist
ES	Effect size
F	Female
FAI	Femoroacetabular impingement
GRADE	Grading of Recommendations Assessment, Development, and Evaluation
GROC	Global rating of change
GT	Gluteal tendinopathy
HOS	Hip Outcome Score
HOS-Brazil (ADL)	Hip Outcome Score-Brazil (Activities of Daily Living)
HOS-Brazil (Sport)	Hip Outcome Score-Brazil (Sport)
Hypo	Hypothesis testing
I	Insufficient
ICC	Intraclass correlation coefficient
ICON	International scientific tendinopathy consensus

i-HOT	International Hip Outcome Tool
LEFS	Lower Extremity Functional Scale
LOA	Limits of agreement
MeSH	Medical subject headings
MIC	Minimal important change
MHHS	Modified Harris Hip Scale
MRI	Magnetic resonance imaging
MT	Manual therapy
n	Sample size
NPRS	Nirschl Phase Rating Scale
NR	Not Reported
NRS	Numerical Rating Scale
NSAIDs	Non-steroidal anti-inflammatory drugs
NT	Not tested
OMERACT	Outcome Measures in Rheumatoid Arthritis Clinical Trials
PHT	Proximal hamstring tendinopathy
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRP	Platelet rich plasma injection
PROM	Patient-reported outcome measure
PT	Physical therapy
QOL	Quality of life
R^2	Measure of model fit
RCT	Randomised controlled trial
RE	Random effects
Ref	Reference
RMSEA	Root mean square error of approximation
ROB	Risk of bias
RTS	Return to Sport
SDC	Smallest detectable change
SEM	Standard error of measurement
SMD	Standardised mean difference
SPD	Standardised paired differences

SRM	Standard response mean
SWT	Shockwave Therapy
TIDiER	Template for Intervention Description and Replication
TLI	Tucker-Lewis Index
US	Ultrasound
V	Very good
VAS	Visual Analogue Scale
VISA-A	Victorian Institute of Sport Assessment – Achilles
VISA-G	Victorian Institute of Sport Assessment – Gluteal Tendinopathy
VISA-G.BR	Victorian Institute of Sport Assessment - Gluteal Tendinopathy Brazil
VISA-G.DK	Victorian Institute of Sport – Gluteal Tendinopathy Danish
VISA-G.F	Victorian Institute of Sport – Gluteal Tendinopathy France
VISA-G.I	Victorian Institute of Sport Assessment – Gluteal Tendinopathy Italian
VISA-H	Victorian Institute of Sport Assessment – Proximal Hamstring Tendon
VISA-H.Br	Victorian Institute of Sport Assessment – Proximal Hamstring Tendon - Brazil
VISA-H.F	Victorian Institute of Sport Assessment – Proximal Hamstring Tendon - French
VISA-H.Sp	Victorian Institute of Sport Assessment – Proximal Hamstring Tendon - Spanish
VISA-P	Victorian Institute of Sport Assessment – Patella

Abstract

Introduction

Tendinopathy of the proximal hamstring is a cause of ischial pain. The tendinopathy commonly impacts athletic populations and active post-menopausal women and is associated with persistent symptoms that restrict physical activity. A greater understanding of proximal hamstring tendinopathy is needed to develop interventions to improve management. It is also critical that the utility of outcomes measures are understood, to ensure the impact of tendinopathies of the hip are captured in research.

Thesis outline

A systematic review was performed to evaluate the effectiveness of treatments used to manage proximal hamstring tendinopathy. This review established that there was insufficient evidence to make strong recommendations on any treatment. The results also found that there was variability in the outcome measures used to capture the impact of the condition across studies.

A qualitative study was performed that involved interviewing expert physiotherapists on best practice in assessment and management. Experts used a clinical reasoning approach to diagnose proximal hamstring tendinopathy, incorporating information gained from the patient interview and response to a battery of provocation tests. Experts managed the condition using education and progressive exercise.

A Delphi study and two systematic reviews were performed to evaluate outcome measures used to assess tendinopathies of the hip. We found a lack of measures with sufficient properties for capturing the core domains of proximal hamstring and gluteal tendinopathy. A single measure of disability was provisionally recommended: the Victorian Institute of Sport Assessment – Proximal hamstring and Victorian Institute of Sport Assessment - gluteal tendinopathy - for both proximal hamstring and gluteal tendinopathy.

Conclusion

Current best practice management of proximal hamstring tendinopathy involves education and progressive exercise. Further high-quality evidence is required to make strong recommendations on management. Rigorously validated outcome measures for evaluating

common hip tendinopathies are lacking – with an urgent need for further development of high-quality measures.

Statement of Authorship

Except where reference is made in the text of the thesis, this thesis contains no material published elsewhere or extracted in whole or in part from a thesis accepted for the award of any other degree or diploma. No other person's work has been used without due acknowledgment in the main text of the thesis. This thesis has not been submitted for the award of any degree or diploma in any other tertiary institution.

Name: Anthony Michael Nasser Date: 18 August 2022

Publications during candidature

Peer-reviewed papers

Publications included in this thesis

Nasser AM, Vicenzino B, Grimaldi G, Anderson J, Semciw AI. *Proximal hamstring tendinopathy; a systematic review of interventions*. International Journal of Sports Physical Therapy, Volume 16, Issue 2, p288-305, 2021. https://doi.org/10.26603/001c.21250.

Nasser AM, Pizzari T, Grimaldi A, Vicenzino B, Rio E, Semciw AI. *Proximal hamstring tendinopathy; expert physiotherapists' perspectives on diagnosis, management and prevention*. Physical Therapy in Sport, Volume 48, p67-75, 2021. https://doi.org/10.1016/j.ptsp.2020.12.008.

Nasser AM, Vicenzino B, Grimaldi A, Rio E, Pizzari T, Semciw AI. *Core outcome set development for proximal hamstring tendinopathy (COS-PHT): a study protocol*. Physical Therapy Reviews. https://doi.org/10.1080/10833196.2022.2077066.

Nasser AM, Fearon A, Grimaldi A, Vicenzino B, Mellor R, Spencer T, Semciw AI. *Outcome measures in the management of gluteal tendinopathy; a systematic review of their measurement properties*. British Journal of Sports Medicine, Volume, 56, Issue 16, p877-887, 2022. http://dx.doi.org/10.1136/bjsports-2021-104548.

Publications submitted for publication

Nasser AM, Dr Alison Grimaldi, Prof Bill Vicenzino, Dr Ebonie Rio, Mr Aidan Rich, Dr Tania Pizzari, Dr Adam Semciw. *Outcome measures for proximal hamstring tendinopathy; a systematic review of their measurement properties.* Physical Therapy in Sport.

Published abstracts

Nasser AM, Vicenzino B, Grimaldi G, Anderson J, Semciw AI. *Proximal hamstring tendinopathy; A systematic review of interventions*. Vol 21, p44, Supplement 1, S96-S97, November 01, 2018.

Publications outside of this thesis

McCambridge A, Nasser AM, Metha P, Stubbs P, Verhagen A. *METAPHoR: meta-research in physiotherapy trials – has reporting of physiotherapy interventions improved in two decades? Analysis of 140 trials reporting on 225 interventions*. Journal of Orthopaedic Physical Therapy.

Verhagen A, Stubbs PW, Mehta P, Kennedy D, Nasser AM, Quel de Oliveira, Pate JW, Skinner IW, McCambridge A. *Metaphor- meta-research in physiotherapy trials: trends in the reporting of statistical significance and clinical relevance between 2000 and 2018.* British Medical Journal - Open.

Publications submitted for publication outside thesis

Kinsella R, Nasser A, Menz H, Pizzari T, Collins N, Semciw A. *The effect of footwear or foot orthoses on impairments and quality of life in people with hip pain: A systematic review*.

Invited non-peer-reviewed publications

Purdam C, Nasser A, de Vos T, Andersen T, Pizzari T**.** *FC Barcelona Tendon Injuries in Football.* Hamstring tendon injuries. 2021.

Presentations

La Trobe University – Tendinopathy: current state of play (Aug 2019)

La Trobe University – Hip and Groin Pain Symposium (June 2022)

Poster presentations

Sports Kongress, Denmark, 2017

Sports Medicine Australia, 2018

Awards

Best poster - clinical sports medicine (Sports Medicine Australia) 2018

Grants

Sports Medicine Australia early career researcher grant: $2000

University of Technology Sydney early career researcher grant: $10 000

Publications included in this thesis

Nasser AM, Vicenzino B, Grimaldi G, Anderson J, Semciw AI. *Proximal hamstring tendinopathy; a systematic review of interventions.* International Journal of Sports Physical Therapy, Volume 16, Issue 2, p288-305, 2021.

Study one. Incorporated as Chapter two

Contributor	Statement of contribution
Anthony Nasser (Candidate)	Designed the research question (25%) Performed the literature search (80%) Prepared tables and figures and extracted data (70%) Study selection (50%) Data interpretation and analysis (65%) Quality appraisal (50%) Drafted manuscript (100%) Edited manuscript (80%)
Bill Vicenzino	Designed the research question (25%) Data interpretation and analysis (10%) Edited manuscript (5%)
Alison Grimaldi	Designed the research question (25%) Data interpretation and analysis (5%) Edited manuscript (5%)
Jay Anderson	Data interpretation and analysis (5%) Study selection (25%) Quality appraisal (25%) Edited manuscript (5%)
Adam Semciw	Designed the research question (25%) Performed the literature search (20%) Prepared tables and figures and extracted data (30%) Study selection (25%) Data interpretation and analysis (15%) Quality appraisal (25%) Edited manuscript (5%)

Nasser AM, Pizzari T, Grimaldi A, Vicenzino B, Rio E, Semciw AI. *Proximal hamstring tendinopathy; expert physiotherapists' perspectives on diagnosis, management and prevention.* Physical Therapy in Sport, Volume 48, p67-75, 2021.

Study two. Incorporated as Chapter three

Contributor	Statement of contribution
Anthony Nasser (Candidate)	Designed the research question (25%) Recruitment (85%) Data collection (100%) Prepared tables and figures (70%) Data interpretation and analysis (35%) Drafted manuscript (100%) Edited manuscript (60%)
Tania Pizzari	Prepared tables and figures (15%) Data interpretation and analysis (30%) Edited manuscript (10%)
Alison Grimaldi	Designed the research question (25%) Recruitment (5%) Prepared tables and figures and extracted data (5%) Data interpretation and analysis (15%) Edited manuscript (10%)
Bill Vicenzino	Designed the research question (25%) Recruitment (5%) Data interpretation and analysis (5%) Edited manuscript (5%)
Ebonie Rio	Prepared tables and figures (5%) Data interpretation and analysis (10%) Edited manuscript (10%)
Adam Semciw	Designed the research question (25%) Recruitment (5%) Prepared tables and figures (5%) Data interpretation and analysis (5%) Edited manuscript (5%)

Nasser AM, Semciw AI, Grimaldi A, Rio E, Pizzari T, Vicenzino B. *Core outcome set development for proximal hamstring tendinopathy (COS-PHT): a study protocol*. Physical Therapy Reviews, p1-7, 2022.

Study three. Incorporated as Chapter five

Contributor	Statement of contribution
Anthony Nasser (Candidate)	Designed the research question (25%) Prepared tables and figures (85%) Drafted manuscript (100%) Edited manuscript (25%)
Tania Pizzari	Prepared tables and figures (5%) Edited manuscript (15%)
Alison Grimaldi	Designed the research question (25%) Prepared tables and figures (5%) Edited manuscript (15%)
Bill Vicenzino	Designed the research question (25%) Edited manuscript (15%)
Ebonie Rio	Prepared tables and figures (5%) Edited manuscript (15%)
Adam Semciw	Designed the research question (25%) Prepared tables and figures (5%) Edited manuscript (15%)

Nasser AM, Vicenzino B, Rich A, Grimaldi A, Rio E, Pizzari T, Semciw AI. *Outcome measures in the management of proximal hamstring tendinopathy; a systematic review of their measurement properties.* Submitted to Physical Therapy in Sport (under review), 2022.

Study four. Incorporated as Chapter six

Contributor	Statement of contribution
Anthony Nasser (Candidate)	Designed the research question (25%) Performed the literature search (90%) Prepared tables and figures (75%) Study selection (50%) Data interpretation and analysis (40%) Quality appraisal (50%) Drafted manuscript (100%) Edited manuscript (50%)
Aidan Rich	Study selection (40%) Quality appraisal (25%) Data interpretation and analysis (10%)
Tania Pizzari	Data interpretation and analysis (10%) Prepared tables and figures (5%) Edited manuscript (10%)
Alison Grimaldi	Designed the research question (25%) Prepared tables and figures (5%) Data interpretation and analysis (10%) Edited manuscript (10%)
Bill Vicenzino	Data interpretation and analysis (10%) Prepared tables and figures (5%) Edited manuscript (10%)
Ebonie Rio	Designed the research question (25%) Data interpretation and analysis (10%) Prepared tables and figures (5%) Edited manuscript (10%)
Adam Semciw	Designed the research question (25%) Performed the literature search (10%) Study selection (10%)

	Quality appraisal (25%) Data interpretation and analysis (10%) Prepared tables and figures (5%) Edited manuscript (10%)

Nasser AM, Pizzari T, Grimaldi A, Vicenzino B, Rio E, Semciw AI. *Core outcome set for proximal hamstring tendinopathy (COS-PHT); a survey of an international collaboration.*

Study five. Incorporated as Chapter seven

Contributor	Statement of contribution
Anthony Nasser (Candidate)	Designed the research question (20%) Designed survey (50%) Data interpretation and analysis (50%) Drafted manuscript (100%) Edited manuscript (50%)
Tania Pizzari	Designed the research question (15%) Designed survey (10%) Data interpretation and analysis (10%) Edited manuscript (10%)
Alison Grimaldi	Designed the research question (20%) Designed survey (10%) Data interpretation and analysis (10%) Edited manuscript (10%)
Bill Vicenzino	Designed the research question (15%) Designed survey (10%) Data interpretation and analysis (10%) Edited manuscript (10%)
Ebonie Rio	Designed the research question (15%) Designed survey (10%) Data interpretation and analysis (10%) Edited manuscript (10%)
Adam Semciw	Designed the research question (15%) Designed survey (10%) Data interpretation and analysis (10%) Drafted and edited manuscript (10%)

Nasser AM, Fearon A, Vicenzino B, Grimaldi G, Mellor R, Spencer T, Semciw AI. *Outcome measures in the management of gluteal tendinopathy; a systematic review of their measurement properties.* British Journal of Sports Medicine, Volume, 56, Issue 16, p877-887, 2022.

Study six. Incorporated as Chapter eight

Contributor	Statement of contribution
Anthony Nasser (Candidate)	Designed the research question (5%) Performed the literature search (75%) Study selection (25%) Quality appraisal (50%) Data interpretation and analysis (65%) Prepared tables and figures (75%) Drafted manuscript (100%) Edited manuscript (35%)
Angie Fearon	Designed the research question (25%) Performed the literature search (20%) Data interpretation and analysis (10%) Prepared tables and figures (5%) Edited manuscript (10%)
Bill Vicenzino	Designed the research question (30%) Data interpretation and analysis (5%) Prepared tables and figures (5%) Edited manuscript (15%)
Alison Grimaldi	Designed the research question (25%) Data interpretation and analysis (10%) Study selection (25%) Prepared tables and figures (5%) Edited manuscript (10%)
Rebecca Mellor	Designed the research question (10%) Study selection (25%) Data interpretation and analysis (20%) Prepared tables and figures (5%) Quality appraisal (25%)

	Edited manuscript (10%)
Trevor Spencer	Study selection (25%) Quality appraisal (25%) Edited manuscript (10%)
Adam Semciw	Designed the research question (5%) Performed the literature search (5%) Data interpretation and analysis (10%) Prepared tables and figures (5%) Edited manuscript (10%)

Concise statement of research question

The principal aims of this thesis were to investigate the treatment options used to manage proximal hamstring tendinopathy and to evaluate the outcome measures used to evaluate common tendinopathies of the hip and pelvis.

Significant research has been conducted to investigate tendinopathy in the upper and lower limbs such as the Achilles tendon, extensor tendons of the forearm, rotator cuff of the shoulder and patella tendon. In contrast, little attention has been given to tendinopathies of the hip and pelvis, in particular proximal hamstring tendinopathy.

This thesis will (i) summarise the current evidence for the efficacy of different treatments on proximal hamstring tendinopathy (study one); (ii) describe how expert physiotherapists assess, manage and prevent proximal hamstring tendinopathy (study two); and (iii) appraise outcome measures used to evaluate outcomes in common hip tendinopathies (studies three to six).

This thesis consists of the following studies:

- Study one (Chapter two): Proximal hamstring tendinopathy; a systematic review of interventions.
- Study two (Chapter three): Proximal hamstring tendinopathy; expert physiotherapists' perspectives on diagnosis, management and prevention.
- Study three (Chapter five): Core outcome set development for proximal hamstring tendinopathy (COS-PHT): a study protocol.
- Study four (Chapter six): Outcome measures in the management of proximal hamstring tendinopathy; a systematic review of their measurement properties.
- Study five (Chapter seven): Core outcome set for proximal hamstring tendinopathy (COS-PHT); a survey of an international collaboration.
- Study six (Chapter eight): Outcome measures in the management of gluteal tendinopathy; a systematic review of their measurement properties.

Contribution of others to the thesis

I would like to acknowledge the team of investigators who contributed to this thesis.

Chief investigators

- Mr Anthony Nasser (Candidate)
- Professor Bill Vicenzino
- Dr Alison Grimaldi
- Associate Professor Tania Pizzari
- Dr Ebonie Rio
- Associate Professor Adam Ivan Semciw

In addition, the following people contributed to research in this thesis:

- Mr Jay Anderson
- Mr Aidan Rich
- Mr Trevor Spencer
- Associate Professor Angie Fearon
- Dr Rebecca Mellor

Acknowledgments

To my supervisors: Associate Professor Adam Semciw, Dr Alison Grimaldi, Professor Bill Vicenzino, Associate Professor Tania Pizzari and Dr Ebonie Rio – thank you for your support. Working with you all has been an incredibly enjoyable and rewarding experience. I look forward to collaborating with you all in the future.

Adam, I started this PhD as a novice researcher. Your kind and relaxed disposition created a great working environment from day one, as we slowly chipped away at each study. It will feel odd not having our fortnightly meetings. I learnt a lot of qualities from you as a researcher - in particular, how respectful of other researchers' work you are. I am forever thankful for the hours you have put into making my PhD journey an enjoyable experience.

Ali, thank you for giving me the nudge to have a chat with Adam back in 2016. I have learnt so much from you over the years. Your generosity with your time and supportive nature both as my boss at PhysioTec and as my supervisor is something that I am incredibly thankful for. Your clinical expertise and constant thirst for knowledge are a real inspiration.

Bill, completing the Master of Sports program in 2014 set me on the track to start this PhD. I was incredibly fortunate to have your expertise to tap into - firstly as a lecturer and then as a supervisor. It has been an absolute privilege to work alongside you, and I have enjoyed the jokes shared along the way.

Tania, thank you for agreeing to join the team. The study I'm most proud of in this thesis is the qualitative study, which I enjoyed working on with you. I found your very honest and direct feedback whilst honing presentations incredibly valuable. You helped me see the big picture of research and felt you were really in my corner.

Ebonie, the passion you have for tendon research is inspiring - hearing you present your tendon research back in Phuket in 2013 motivated me to get more out of myself as a physio. Your clinical motivation for research is ever-present and has really resonated with me over the past few years. I have enjoyed working with you and hope to continue to in the future.

To the co-authors who contributed to the studies in this thesis, thank you for your contribution and comments on each paper.

To Oscar my Labrador, you've spent a lot of time over the past five years hanging out with me curled up under my desk, whilst I was hunched over my computer. Your company has made this process much more pleasant. Now, that this is submitted we'll go for longer walks, I promise. To my family - Mum, Dad, Nell and Nick thank you for the support. Mum, you've always been there for me - to bounce ideas off or talk through any challenging situations. Your encouragement right from the start helped me find the courage to enrol. Dad, you have the best work ethic of anyone I know. You've set an amazing example in getting the best out of yourself and believing in yourself. Nell, moving up to Queensland to study with you was a great move for both of us – thank you for helping to make the transition up to the Sunshine state easy and always encouraging me. Nick, you've always been there to give me perspective – I'm looking forward to spending more time together, fishing. Jen, thank you for the encouragement and the support in getting this thesis over the line. I can't wait to marry you next year.

Statement of parts of the thesis submitted to qualify for the award of another degree

None

Financial support

This work was supported by an Australian Government Research Training Program Scholarship and Sports Medicine Australia

Australian and New Zealand Standard Research Classifications (ANZSRC)

Physiotherapy (100%)

Keywords

Tendinopathy, buttock pain, hip pain, gluteal tendinopathy, hamstring tendinopathy, clinimetrics and outcome measure

Research involving human or animal subjects

Ethical approval for research studies in this thesis was obtained through multiple applications to the La Trobe University Human Ethics Committee and the University of Queensland ethics committee.

Project Title:	Proximal Hamstring Tendinopathy: current practices in assessment and management and prevention
Approval Number:	2018001158
Committee:	University of Queensland Health and Behavioral Sciences, Low & Negligible Risk Ethics Sub-Committee
Granting Agency:	None

Copy of the ethics approval letter is included as Appendix A

Project Title:	Development of a core outcome set for proximal hamstring tendinopathy (COS-PHT): a Delphi study of healthcare professionals and patients
Approval Number:	HEC21210
Committee:	La Trobe University Human Ethics Committee
Granting Agency:	None

Copy of the ethics approval letter is included as Appendix B

1 Chapter 1: Background

1.1 Tendinitis to tendinopathy: evolution and background

The evolution in the understanding of tendinopathy has seen changes in the terminology used to describe the condition. The term tendinitis, indicating an inflammatory aetiology, was replaced by 'tendinosis' when histological studies demonstrated that degenerative processes were more commonly associated with tendon pain (1, 2). As a true diagnosis of 'tendinosis' requires histopathology, the less specific umbrella term 'tendinopathy' came into common use (Figure 1.1) (2, 3). Tendinopathy is the now preferred nomenclature used to describe persistent tendon pain and loss of function related to mechanical loading (2). More explicit descriptors are applied to pathology involving full-thickness tendon ruptures and pathology in surrounding tissue with different disease processes, such as paratendinitis (4).

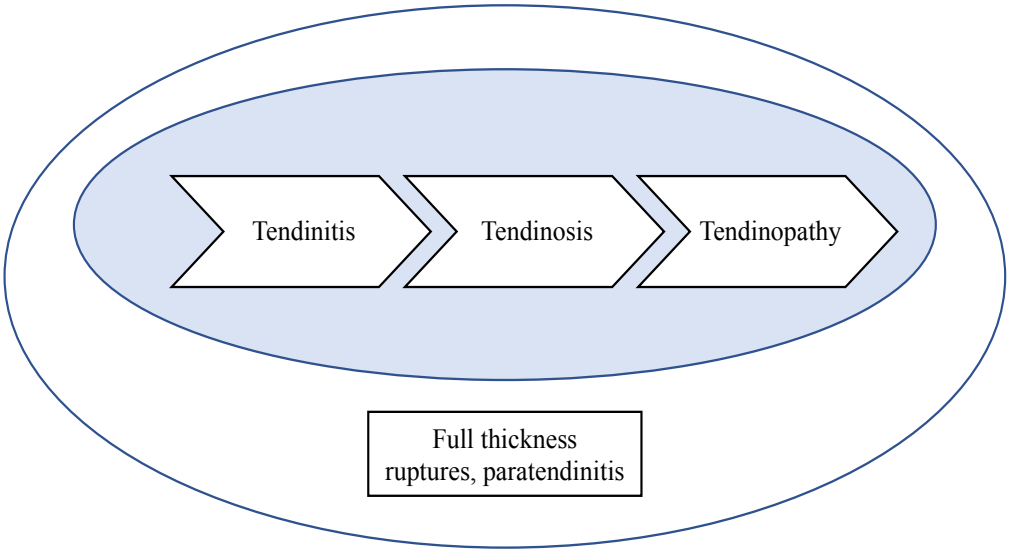


Figure 1.1 The evolution in terminology of tendon pain

1.2 Epidemiology

1.2.1 Epidemiology of tendinopathy

Although there are more than 600 muscular-tendon units in the body, tendinopathy only affects a small number in specific locations (5). In the upper limb, the most common tendinopathies are rotator cuff-related shoulder pain and lateral elbow tendinopathy (5). In the lower limb, gluteal tendinopathy, plantar heel pain (also termed plantar fasciopathy), Achilles

tendinopathy, and patellar tendinopathy (also known as jumper's knee) are the most common (in a population attending general practice) (6). These conditions share the feature of pain at the site of the tendon with mechanical loading (7).

Proximal hamstring tendinopathy was first described as "the hamstring syndrome" by Puranen in 1988 (8). Since this first published case series, the terminology used to describe the condition has included: hamstring tendinitis, high hamstring tendinopathy, and hamstring origin tendinopathy (9). Clinical terminology for various tendinopathy sites was agreed upon in a recent international consensus (2). However, nomenclature for several tendinopathies (gluteal tendon, tibialis posterior tendon) including the proximal hamstring tendon was not addressed (2). In the literature, proximal hamstring tendinopathy is now the most frequently used term.

It has been estimated that 30% of musculoskeletal consultations in general practice are related to tendinopathy (10). The incidence and prevalence of lower limb tendinopathy were 11.83 and 10.52 per 1000 person-years, respectively in a general practice population (Figure 1.2) (6). In this sample, just 29% of individuals described a relationship with sports, demonstrating the widespread impact of tendinopathy (6).

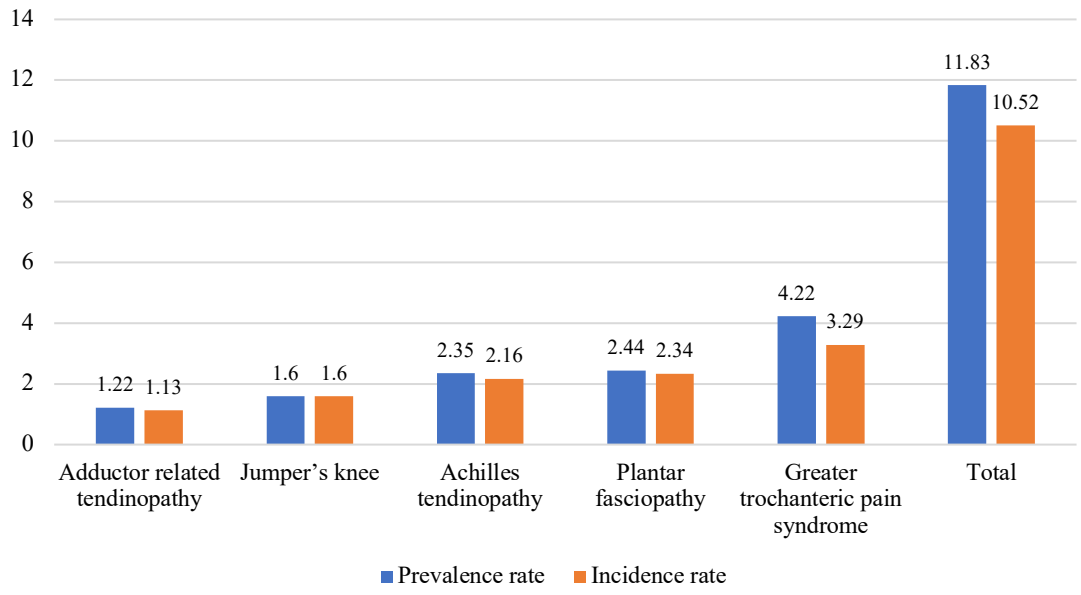


Figure 1.2 Prevalence rate and incidence rate of tendinopathy per 1000 person-years Dutch General Practice: n = 10 651, tendinopathy patients n = 126 [6]

1.2.2 Epidemiology of proximal hamstring tendinopathy

There is data from a single longitudinal study on the incidence of proximal hamstring tendinopathy (11). This study reported the incidence of proximal hamstring tendinopathy in soccer was low (11). Specifically, each season there were 1.5 (95% CI 0.5-3.2) cases in professional soccer and 0.4 (95% CI 0.1-1.0) cases per 100 athletes in youth soccer (11). In women's soccer, there were 1.2 (95% CI 0.2-3.4) cases and 0 (95% CI 0.0-0.6) cases in professional and youth soccer respectively (per 100 athletes) (11).

1.3 Tendon structure - from healthy to tendinopathic

1.3.1 Normal tendon structure and function

Tendons serve as an important mediator of force transmission between bone and muscle (10, 12). For this reason, tendons have been termed 'mechanical bridges' (13). Tendons are also capable of absorbing large external forces (13). The structural makeup of the tendon brings with it the ability to store and release energy. This allows people to perform activities such as running and jumping (5).

At a macroscopic level, healthy tendons are organised structures composed of mainly type I collagen (14). Collagen forms over half of a tendon's dry weight (60–85%) (12). Tendon strength is related to its thickness and collagen content (15). The structural composition of tendon is not uniform throughout its length. For example, tendon structure varies at the myotendinous junction, enthesis and where tendons are compressed (12).

The basic cellular components of the tendon are called tenocytes (5), which are aligned along the length of the collagen fibrils (16). Equilibrium is maintained by tenocytes, which control protein synthesis. These cells are capable of producing components for tendon function, such as type I collagen (10).

1.3.2 Response to loading

Tenocytes respond actively to mechanical load (e.g. exercise) (16, 17). Loading of tendon results in increased synthesis of collagen protein. For example, following exercise, there is an increase in collagen synthesis and degradation (12, 16). Collagen formation peaks at around 24 hours post activity and stays above the original steady-state level for about three days (16).

In the 24-36 hours following exercise, this response results in the reduction of collagen, which is followed by an increase in net collagen synthesis 36-72 hours following an activity (16).

Collagen requires a period of restoration following exercise (12). Repeated loading with insufficient rest, can lead to net collagen degradation (12). This might be a key factor in overload resulting in tendinopathy. Consequentially, to maintain tendon health there is a fine balance between loading (collagen breakdown) and recovery (collagen recovery).

1.4 Patho-aetiological models of tendon pathology – why tendon fails to heal

It is largely unknown why healthy tendon fails to adapt and tendinopathy develops (5). Over the past decades, several theories have been proposed, which are discussed below.

1.4.1 Mechanical Theory

The mechanical theory maintains that repeated loading results in tensile micro-strain, which with inadequate time to repair, may result in tendinopathy (18). Clinical evidence in support of this theory is the frequency of tendinopathy in athletic populations who place high demands on their tendons. For example, patella tendinopathy is more common in individuals who perform high levels of physical activity, such as volleyball (19), with the likelihood of injury increasing with greater training volume (20). Furthermore, former elite running athletes, who presumably have experienced a high level of Achilles loading, experience over 10 times the rate of Achilles tendinopathy (40%) before the age of 45, compared to non-athlete controls (3%) (21).

There are limitations to the mechanical theory. Several risk factors and prognostic factors do not fit neatly into the mechanical model. For example, tendinopathy is common in sedentary populations which is thought to be influenced by factors other than mechanical load. This includes systemic factors such as inflammatory arthropathies or diabetes mellitus (22), as well as risk factors such as alcohol consumption and associations of severity with psychosocial factors (23, 24).

1.4.2 Inflammatory Theory

The inflammatory theory advocates that tendinopathy arises predominantly from an inflammatory process. This model went out of favour in the 1990s and early 2000s due to the publication of studies that showed degenerative changes in tendons resulting from increased demand with inadequate time for repair (25). There has been a slight resurgence in the last decade due to the publication of studies reporting the presence of low concentrations of inflammatory mediators in chronic tendinopathy (26). However, it is argued that the presence of inflammatory cells in chronic tendinopathy (e.g. inflammatory cytokines) does not necessarily mean that inflammation is the primary driver of pathology (27). Further research is required to understand the role inflammatory cytokines play in tendinopathy as they can have different functions at differing concentrations and different actions in different environments (27).

1.4.3 Apoptosis Theory

Apoptosis is a normal physiological process that controls the cell population by removing damaged or infected cells and is described as programmed cell death (28). Uncontrolled apoptosis may be pathogenic, with excessive apoptosis seen in several conditions, such as osteoarthritis and rheumatoid arthritis (5). This theory argues that increased apoptosis negatively alters the rate of collagen synthesis and repair (29). This may lead to weaker tendon tissue and eventually increase the risk for tendinopathy or rupture (28).

1.4.4 Continuum Model

The continuum model, proposed in 2009, was created to explain the variability in the way patients present with tendon pain, the capability of structural recovery and the structural factors that limit the return to pain-free function (30). The model has three stages: i) reactive tendinopathy ii) tendon dysrepair and iii) degenerative tendinopathy (30).

As the name continuum suggests, there is a connection between stages. The reactive stage describes an acute adaptation to tendon overload (30). This causes the tendon to thicken to reduce stress on the tendon (30). If the load on the tendon is reduced the tendon can return to its normal state. The dysrepair phase represents the attempt of the tendon to heal (30). The degenerative stage is typified by areas of cell death (30). There is little capacity for reversibility of the areas of degeneration (30).

The continuum model was revisited in 2015 (27). The update emphasised an important clinical presentation which is 'reactive-on-degenerative tendinopathy'. This presentation refers to where the structurally normal portion of the tendon (to conventional imaging modalities) goes in and out of a reactive state (27).

1.4.5 Tendon compression

In 2003, tendon compression was introduced as an important factor in the development of tendinopathy (31). This was further popularised in 2012, through a seminal publication in the British Journal of Sports Medicine (32). Before this publication, Benjamin et al. described the concept of an enthesis organ which explained how tendons react to compression (33). The enthesis organ is a group of tissues that function to reduce the load concentration at the enthesis (where the tendon attaches to bone) (33). This organ includes a bony prominence and a bursa that is adjacent to the enthesis (33). The enthesis adapts to increased compressive loads by changing its composition, more specifically by increasing fibrocartilage, proteoglycans and type II collagen content (34). The function of the enthesis organ is to reduce the compression that occurs between the tendon and the bone, reduce the tensile strain on the tendon insertion and provide a mechanical advantage (33).

The impact compressive load has on tendons has been investigated in animal studies. Soslowsky *et al.* investigated the impact that compressive, tensile and combined compressive and tensile load on the supraspinatus tendon of rats (35, 36). The authors found that compression alone had little impact on the tendon. Whilst tensile load was detrimental to the tendon, compression combined with tensile load was the most damaging (35, 36). Further evidence is required to understand the impact of compressive load on tendinopathy, as most of the research is based on animal or anatomical studies.

Contemporary management of tendinopathy requires the consideration of not only tensile but also compressive loads (32). The consideration of tendon compression is also an important factor in exercise selection and education on activity modification. Another clinical deduction from this model is that stretching is best avoided, due to the compressive element (32). This is important as stretching has historically been considered a panacea for various musculoskeletal conditions. Several randomised controlled trials on gluteal tendinopathy have incorporated education on limiting compressive loading as part of an intervention, in both

activities of daily living and exercise selection, with promising results (37-39). Further research is required to understand the role management of compressive loads has in the treatment of proximal hamstring tendinopathy.

1.5 Risk factors in tendinopathy

Several health conditions that have a systemic impact have been proposed to decrease the threshold for tendon pain and prolong recovery (40). Systematic reviews have shown a strong relationship between diabetes mellitus and tendinopathy (41), as well as a relationship between higher serum cholesterol and tendon health (42). Other risk factors include obesity and statin use which have been associated with predisposition to tendinopathy and may affect its course (40). Certain medications, such as corticosteroids and fluoroquinolone antibiotics have been associated with tendinopathy (5).

A link has also been identified between tendinopathy and menopausal status. Before menopause, tendinopathy is less common in women, whereas post-menopause the incidence of tendinopathy is similar between sexes (43). Reduction of estrogen due to menopause is believed to influence tendon homeostasis, with menopause resulting in a rapid reduction in estrogen in the first six months, that continues for about three years (43).

No studies have explored risk factors in developing proximal hamstring tendinopathy. Future research should engage with health care professionals who regularly see patients with proximal hamstring tendinopathy to identify potential risk factors worth exploring.

1.6 The burden of disease

The economic cost of proximal hamstring tendinopathy to patients and the health care system is unknown. It is understood that hip pain can have a substantial impact on health and quality of life (44). The effect on quality of life and disability of other tendinopathies around the hip, such as gluteal tendinopathy, is significant. One study found that disability and quality of life levels of gluteal tendinopathy were comparable to the end-stage of hip osteoarthritis (45). The impact of tendinopathy on participation in activities that involve higher hamstring loads is significant (46). Examples include activities such as running, dancing, yoga and walking (particularly walking ascents). The condition also impacts sedentary activities due to

symptoms with sitting. This often will affect work and recreational activities (e.g. cycling, watching television, dining).

Proximal hamstring tendinopathy may also have a significant economic impact. People with proximal hamstring tendinopathy may seek opinions from various health professionals, and be referred for costly imaging services (e.g. magnetic resonance imaging). Patients may also spend money on expensive treatments such as platelet-rich plasma injection, surgery or ongoing consultations from healthcare practitioners (e.g. physiotherapists, chiropractors). Further to this, the condition may impact individuals for years, with a case series reporting the average length of symptoms of patients enrolled to undergo surgery was 23 months (47).

1.7 Pain in tendinopathy

The understanding of the science of pain in tendinopathy has developed significantly in the last two decades. Nociception appears to arise at the site of the tendon affected following an intensive bout of activity (7). The pain typically reduces while exercising, however, with sustained activity, symptoms increase and function declines (e.g. running performance) (4). Symptoms may also be present at rest. For example, sitting pain is a key symptom in people with proximal hamstring tendinopathy (46, 48).

Pain is a fundamental characteristic of tendinopathy, though the connection between tendon structure and pain is complex because tendon pathology can exist in the absence of pain (49). Imaging findings showing tendinosis may be an incidental finding and unrelated to a clinical presentation; imaging findings are not sufficient to diagnose tendinopathy (4).

1.8 Anatomy and function of the hamstring

1.8.1 Anatomy of the proximal hamstring complex

The 'hamstring' includes four muscles on the posterior aspect of the thigh (biceps femoris short head and long head, semimembranosus and semitendinosus) (50). The short head of biceps femoris spans a single joint, whereas the other three hamstring muscles cross two joints. The long head of biceps femoris and semitendinosus arise from a common tendon that originates from the inferomedial aspect of the ischial tuberosity (Figure 1.3) (50). The long head of biceps femoris and semitendinosus split at an average of 9.9cm (+/-1.5cm) from their proximal origin (51). The origin of semimembranosus is deep and lateral to the common tendon (50). The

word semimembranosus is derived from the Latin words semi meaning “half” and membrane meaning “skin”, with the name describing the structure of the flat, broad, membranous tendon (52).

1.8.2 Innervation

The sciatic nerve comes in proximity to the hamstring origin. On average the sciatic nerve is only 1.2 +/-0.2cm from the lateral aspect of the ischial tuberosity (in cadaveric dissections) (51). The tibial portion of the sciatic nerve innervates the long head of biceps femoris, semitendinosus and semimembranosus (52). Whereas the common peroneal portion of the sciatic nerve innervates the short head of biceps femoris (52)

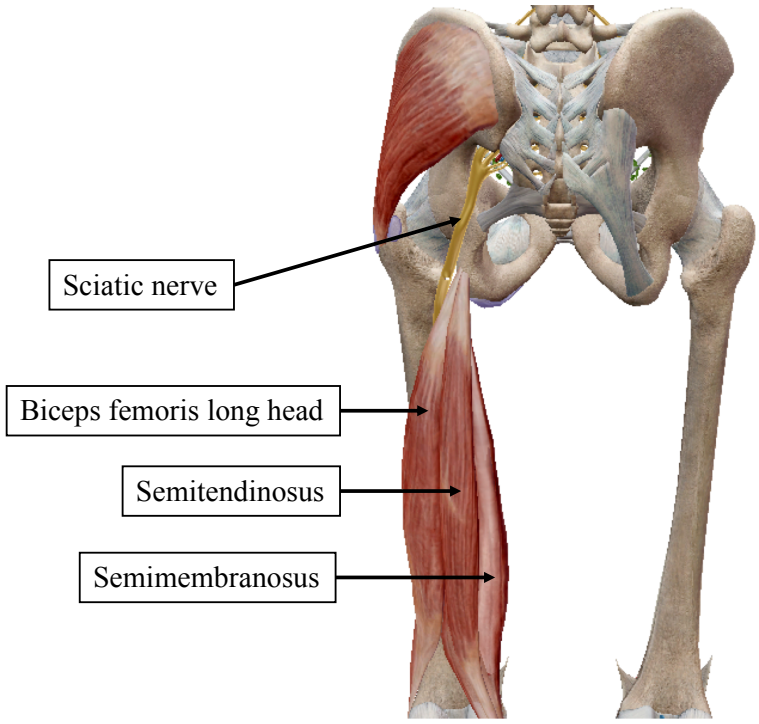


Figure 1.3 Anatomy of the proximal hamstring region and surrounding neural structures

Primal pictures ©

1.8.3 Hamstring function

The hamstring muscle group is capable of flexing the knee joint and extending the hip, with the biceps femoris also able to assist with external rotation (52). The hamstring muscles endure significant loads in the stance and swing phase of running (53), with the load on the hamstrings highest during late swing (54). In this phase, the hamstrings absorb kinetic energy and

complete negative work (55). In late swing the hamstrings are vulnerable to muscle injury (55). This is believed to be because of the high load experienced whilst the hamstrings are lengthening (54).

Whether the hamstrings are lengthening in late swing is the cause of debate. Some researchers contend that the increased distance between the origin and insertion seen in biomechanical studies, which are interpreted traditionally as muscle lengthening, occurs due to the effect of muscle slack (56). This theory proposes that the hamstrings are instead functioning in an isometric fashion (56). This debate has led to differences in opinion on the ideal method for rehabilitation and prevention of hamstring strain injuries. The hamstrings are also active throughout the stance phase where they contribute to the propulsive ground force impulse that accelerates the body (55).

1.8.4 Tendons affected in proximal hamstring tendinopathy

The part of the proximal hamstring tendon complex that is affected by proximal hamstring tendinopathy was reported in two retrospective surgical studies (Figure 1.4) (47, 57). Lempainen et al. reported that in all cases (n =103) the semimembranosus was affected (57). The results of a smaller case series by Benazzo et al. (n =17) were more variable (bicep femoris in nine (53%), semimembranosus five (29%), semitendinosus one (6%) undetectable in two (12%)) (47). More research is required to assist in evaluating if there is a vulnerable region due to the inconsistency in these studies. A better understanding of the area affected by proximal hamstring tendinopathy may have implications on surgical interventions, exercise selection and the location of injection therapies.

1.9 Diagnosis of proximal hamstring tendinopathy

No gold standard exists for tendinopathy diagnosis. A marriage of information from the patient history and clinical tests is used to achieve diagnosis (7). Imaging is useful to assist with differential diagnoses (7).

1.9.1 Clinical history

In proximal hamstring tendinopathy the onset of symptoms is typically insidious, arising in the 24 hours following an increase in activity levels, especially in the frequency of activities that involve deep ranges of hip flexion (46). The area of pain reported is localised to the hamstring

origin at the ischial tuberosity, which can be seen in Figure 1.5 (58). Symptoms are commonly reported as a deep, dull ache (46). Symptoms may extend posteriorly down the thigh, but not as far as the knee (57). Research is needed to better define the presentation of patients with proximal hamstring tendinopathy. This will assist clinicians with diagnosis.

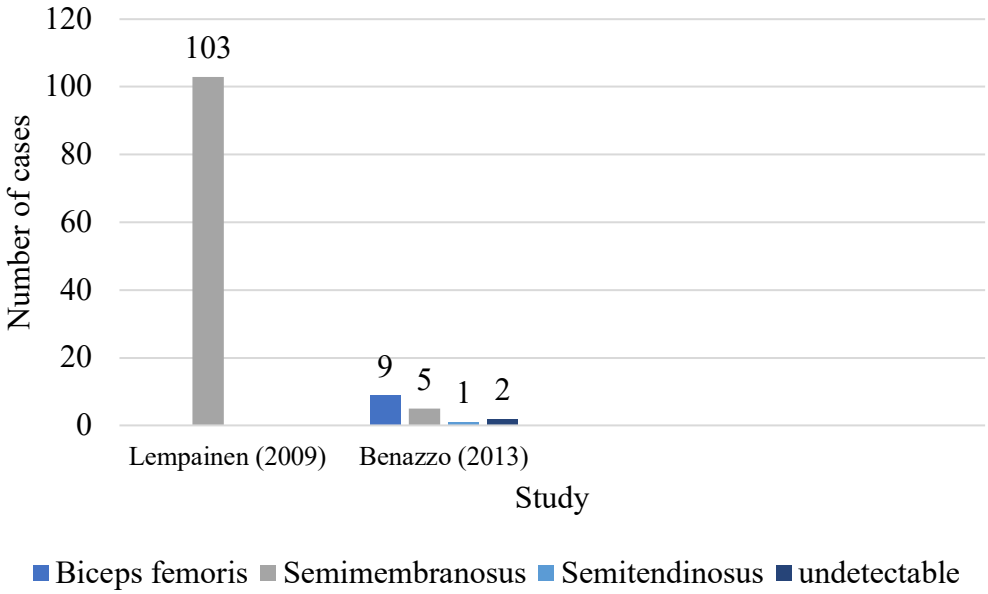


Figure 1.4 Location of proximal hamstring tendinopathy in two surgical case series

1.9.2 Behaviour of symptoms

Inactivity may be a pre-curser to increased symptoms. For example, symptoms are typically worse at the start of a run (5). Symptoms typically ease after a period of activity, but can worsen again towards the end of the activity as the muscle-tendon unit fatigues (59). Activities that involve deep hip flexion range of motion, while concurrently loading the hamstring (Figure 1.6), such as walking or running up a hill are often reported as particularly painful. This may be related to the combination of significant compressive and tensile loads (46, 58). Symptoms are worse when running at faster speeds or accelerating (58).

Activities that place only a small amount of load on the hamstring unit, but involve sustained compression of the hamstring insertion, such as sitting whilst at a computer or driving are typically painful (46). In contrast, activities that involve minimal compressive or tensile load on the hamstring, such as lying supine (in minimal hip flexion) or when standing, are rarely painful (46)

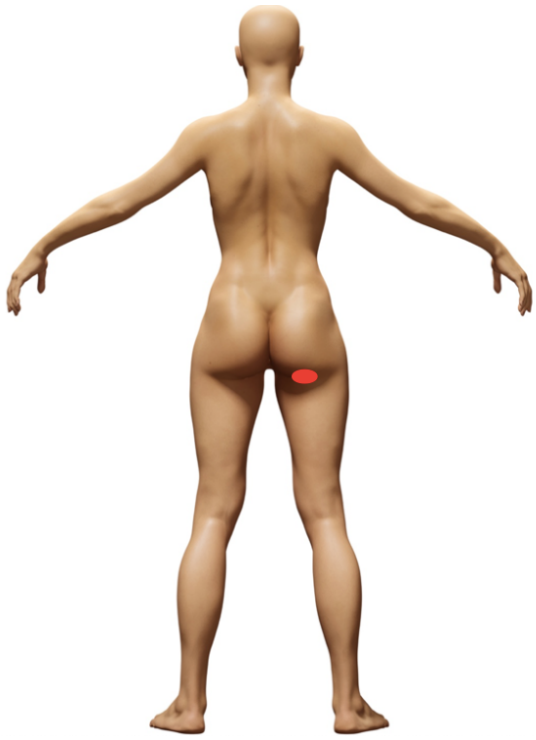


Figure 1.5 Location of symptoms in proximal hamstring tendinopathy [60]

1.9.3 Clinical tests

The diagnostic accuracy of three pain provocation tests: the Puranen-Orava, bent-knee stretch test, and modified bent-knee test have been reported in the literature (61). The tests have been assessed in a single study of 92 professional athletes (mean age 23), with and without proximal hamstring tendinopathy (n=46 in each group). Athletes were involved in a variety of sports including sprinting (n=21), long-distance running (n =9), long jump (n =7), hurdles (n =9), soccer (n = 21) and rugby (n=25). In this study, all athletes received a clinical examination, which included the three aforementioned tests, palpation of the ischial tuberosity area, followed by history taking (61). Athletes with a history of possible proximal hamstring tendinopathy (self-reported history of pain over the lower gluteal area) and positive clinical examination received an MRI (magnetic resonance imaging) (61).

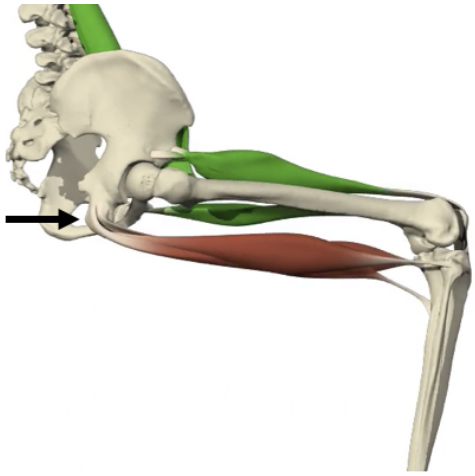


Figure 1.6 Hamstring tendon compression at the ischium tuberosity in hip flexion

Primal pictures ©

Using MRI coupled with clinical diagnosis as a reference standard, all three tests demonstrated high sensitivity and specificity (61). Clinical diagnosis included: reported pain of ≥4cm in the lower gluteal region on a 10cm VAS scale and tenderness in the ischial tuberosity area. The diagnosis was confirmed by MRI.

The sensitivity, specificity and likelihood ratios of the three tests are reported in Table 1.1. The modified bent-knee stretch, which involves stretching the hamstring at velocity was the most sensitive and specific (61). Despite the high levels of sensitivity reported, it has been suggested that these tests may not be provocative enough to reproduce symptoms in patients who are not irritable, and may lead to false negatives (46).

The diagnostic accuracy of another three pain provocation tests was reported in a single, retrospective study with a mixed cohort of proximal hamstring tendon injuries (e.g. partial tears, tendinopathy, avulsions and ruptures) (62). The study included 42 patients (mean age of 50.3 years) and included the Active-90 test, Active-30 test and long stride walking test hamstring. The activity level of the population was not described (e.g. sport). The Active-90 and Active-30 tests involve resisted isometric knee flexion for 5 seconds in inner (90 degrees) and outer range (30 degrees) of knee flexion respectively, whilst the examiner simultaneously palpates the conjoined tendon and semimembranosus lateral to the ischium (62). The long stride heel strike test involves a patient walking a short distance (6-7 metres), with a positive result re-provoking symptoms at the ischial tuberosity (or just lateral) at heel strike (62). In

this study, the reference standard was MRI with or without diagnostic injection (e.g. MRI was performed for all patients, with diagnostic injection then performed for patients where the diagnosis was uncertain). The sensitivity and specificity values are summarized in Table 1.1. The authors suggested that combining the Active-30 and Active-90 was most useful in diagnosing a tendon tear in the proximal hamstring, and a negative result of both tests had utility in ruling out the condition (as diagnosed by MRI plus or minus diagnostic injection) (62).

These two studies suggest that several tests have a high sensitivity and specificity in the populations tested. A significant limitation of these studies is the use of MRI or diagnostic injection as a reference standard. Future research that compares pain provocation tests to a complete clinical workup is warranted, although this too has its limitations.

1.9.4 Palpation

Historically, palpation has been used, along with imaging, as a key tool in the diagnosis of tendinopathy. The value of palpation varies in different lower limb tendinopathies. For example, palpation has limited clinical value in the diagnosis of Achilles tendinopathy (negative likelihood ratio of 0.48 (95% CI 0.29-0.80), positive likelihood ratio 3.15 (95% CI 1.6.1, 6.18)) (63). In contrast, the absence of pain with palpation of the gluteus medius/minimus insertion at the greater trochanter is valuable in ruling out gluteal tendinopathy in patients with signs of tendon abnormality on MRI (64). The diagnostic utility of palpation as a pain provocation test for proximal hamstring tendinopathy has not been assessed, although it is commonly used in the diagnostic workup in case series (8, 47, 65). Palpation was purposely left out of the recent ICON consensus guidelines that outlined what should be reported in inclusion criteria as developed by

Table 1.1 Differential diagnosis of buttock pain

Diagnosis	Common features	Clinical test	Sensitivity (95% CI)	Specificity (95% CI)	Likelihood ratio positive (95% CI)	Likelihood ratio negative (95% CI)
Extra-articular						
Proximal hamstring tendinopathy	Insidious onset, increase in hamstring loading, particularly in activities involving hip flexion, localised ischial pain with hamstring loading and sitting [46]	Puranen-Orava test [61]	0.76 (0.61, 0.87)	0.82 (0.68, 0.92)	4.2	0.29
		Bent knee stretch test [61]	0.84 (0.71, 0.93)	0.87 (0.73, 0.95)	6.5	0.18
		Modified bent knee stretch test [61]	0.89 (0.76, 0.96)	0.91 (0.79, 0.97)	10.2	0.12
		Long stride heel strike test [62]	0.55 (0.37-0.72)	0.73 (0.48-0.89)	2.08 (0.84-5.15)	0.61 (0.36-1.02)
		Active 30 deg hamstring test + Active 90 deg knee flexion test [62]	0.84 (0.66-0.93)	0.97 (0.76-0.99)	26.86 (1.75-413.02)	0.17 (0.07-0.39)
Hamstring strain injury	Acute onset, high speed sprinting typically affects the biceps femoris long head, primarily affects biceps femoris long head (secondary semitendinosus)[66]	Composite clinical assessment (passive straight leg raise, active knee extension, resisted isometric knee flexion) [67]	0.95 (0.83, 0.99)	0.03 (0.00, 0.22)	0.97 (0.88, 1.08)	1.9 (0.2, 16.0)
		Taking shoe off test [67]	1.00 (0.97, 1.00)	1.00 (0.97, 1.00)	280.0 (17.6, 4454.6)	0.00 (0.00, 0.06)
Hamstring strain injury -*stretch type*	Acute onset, overstretch injury (maximal hip flexion and knee extension), preferentially affects semimembranosus [68, 69]					
Acute hamstring origin rupture or acute partial tear of hamstring origin	Acute onset – sudden severe pain report a popping sensation, ecchymosis down the posterior thigh, overstretch injury (maximal hip flexion and knee extension e.g. split), may involve 1,2 or 3 tendons, present with stiff leg gait pattern, palpation of local mass of retracted muscle (rupture) [70]	Active 30 deg hamstring test + Active 90 deg knee flexion test [62]	0.84 (0.66-0.93)	0.97 (0.76-0.99)	26.86 (1.75-413.02)	0.17 (0.07-0.39)
		Supine plank test [71]	NR	NR	NR	NR
Adductor tendinopathy (ischocondylar portion)	Insidious onset, pain slightly medial to ischial tuberosity	NR	NR	NR	NR	NR
Intra-articular/articular						
Hip joint chondral labral pathology (e.g.		Flexion adduction internal rotation test [73]	0.96 (0.91-0.99)	0.11 (0.06-0.20)	1.079 (0.99-1.17)	0.364 (0.12-1.08)

femoroacetabular impingement)	Insidious onset, posterior symptoms may present with concurrently with anterior hip pain, posterior symptoms exacerbated towards end of range hip movements [72]	Maximal squat test [73]	0.75 (0.57-0.89)	0.41 (0.27 to 0.57)	1.278 (0.93-1.75)	0.605 (0.30-1.21)
Ischiofemoral impingement	Insidious, distal pain that is lateral to the ischium, pain over area of quadratus femoris, poor hip abductor function, previous total hip replacement	Ischiofemoral impingement test (side lying take hip in to adduction whilst maintaining full hip extension) [74]	0.82 (0.56-0.95)	0.85 (0.54-0.97)	5.35 (1.47-19.52)	0.21 (0.07-0.60)
		Long stride walking test [74]	0.94 (0.69-0.99)	0.85 (0.54-0.97)	6.12 (1.70-22.01)	0.07 (0.01-0.48)
Avulsion fracture ischial tuberosity	Acute onset, adolescence, "hurdlers fracture", overstretch injury [7]	Resisted isometric knee flexion Passive straight leg raise	NR	NR	NR	NR
Apophysitis hamstring origin	Insidious onset, adolescence, repetitive high hamstring loads in hip flexion with knee extended [70]		NR	NR	NR	NR
Stress fractures of the ischial ramus or pubic ramus	Insidious onset, history of increase in workload, common in running-based sports, female gender, amenorrhea, osteopenia/osteoporosis, worsening of symptoms with exercise (e.g. no warm-up effect) [75]	Magnetic resonance imaging Localised pain on palpation[9]	NR	NR	NR	NR
Somatic referred pain						
Lumbar spine	Typically insidious onset, diffuse area of pain, hard to localise, accompanied with/without lower back pain, pain in the posterior thigh and/or lower leg aggravated by movements involving the lumbar spine, neurological exam	Active movement of lumbar spine Palpation of spinal segments (PAIVMs) Lumbar quadrant test	NA	NA	NA	NA
Sacroiliac joint	Acute or insidious onset, symptoms in over the sacroiliac joint or in the region of posterior iliac crest and the gluteal fold. Pain may refer to the posterior thigh. Pain does not tend to refer to lumbar spine Common aggravating factors involve load transference e.g. sitting, standing, sit to stand, rolling, single leg stance, hopping/jumping, during/post pregnancy.	Cluster of: Distraction, compression, thigh thrust, FABER, Gaenslen's test, active straight leg raise in supine, sacral thrust [76]	0.83; 95% CI: 0.62, 0.93)	0.59; 95% CI: 0.36, 0.79)	2.13 (95% CI: 1.2, 3.9)	0.33 (95% CI: 0.11, 0.72)
Neural						
Sciatic nerve entrapment	Localised or diffuse pain radiating into posterior thigh	Straight leg raise [62]	0.15 (0.05-0.33)	0.95 (0.68-1.00)	3.20 (0.18-56.92)	0.90 (0.73-1.10)
		Active piriformis test [62]	0.78 (0.58-0.90)	0.80 (0.49-0.94)	3.90 (1.11-13.77)	0.27 (0.12-0.63)

		Seated piriformis stretch test [62]	0.52 (0.33-0.71)	0.90 (0.60-0.98)	0.53 (0.33-0.85)	0.53 (0.33-0.85)
Peripheral nerve entrapment: inferior cluneal nerve	History of a fall onto the bottom or pain with sitting on hard surfaces, area of symptoms typically localised to inferior part of buttocks	NR	NR	NR	NR	NR
Peripheral nerve entrapment: pudendal nerve	Pain, numbness, and dysfunction in the genitalia, rectum and urinary tract History of trauma to hip region or pelvis (e.g. childbirth, surgery, fracture), history of prolonged compression e.g. cycling Symptoms may include: sexual dysfunction, sphincter dysfunction, fecal incontinence and urinary hesitancy [77]	Nates inclusion criteria [77] • Pain in distribution of pudendal nerve • Pain predominantly in sitting • The patient does not get up with pain at night • No sensory loss • Relief with block to pudendal nerve	NR	NR	NR	NR
Other						
Spondyloarthropathy	Symptoms may include: insidious onset, prolonged morning stiffness (e.g. greater than 30 minutes), alternating buttock pain, nail changes, relief with NSAIDs, waking at night because of back pain in the second half of the night and improvement with exercise, but not rest [78]	ASAS criteria – 'iPAIN' for spondyloarthritis [79] At least 4 of the 5 criteria below: • Insidious onset of symptoms • Pain at night with improvement upon getting up • Onset < 40 years • Improvement of symptoms with exercise and lack of improvement with rest	74.4 (68.1-80.8)	39.5 (33.0-46.1)	NR	NR

ASAS = Assessment of SpondyloArthritis international Society, NSAIDs = non-steroidal anti-inflammatory drugs

patients and clinicians in a recent consensus (80). Research is required to understand how clinicians use and interpret diagnostic tests, such as palpation, in their clinical examination.

1.9.5 Successive progressive loading tests

Progressive loading tests have been suggested to have use in diagnosing proximal hamstring tendinopathy (5). The utility has not been tested in diagnostic accuracy studies (46). Progressive loading tests should reveal an increase in symptoms with sequential activities that place increasing amounts of tensile and compressive loads on the proximal hamstring musculotendon-unit (46). For example, in proximal hamstring tendinopathy, minimal symptoms may be elicited with resisted isometric knee flexion in prone (e.g. hip positioned in neutral hip flexion range of motion). Symptoms should increase when repeating resisted isometric knee flexion in supine, with the hip flexed to 90 degrees. During successive tests, symptoms should remain localised to the proximal hamstring origin and increase in intensity with greater load (tensile or compressive) (46).

1.9.6 Physical examination deficits

Information on physical examination deficits has been inconsistent across studies. Lempainien et al. reported no remarkable weakness in knee flexion or hip extension movements in manual muscle tests in any of 103 respective cases before undergoing surgery (81). This is contrary to findings by Young et al., who found a reduction of knee flexion strength in 95.1% of their group (65). Other physical impairments that have been hypothesised in narrative reviews include hip abductor weakness, hip adductor weakness, decreased hip external rotator strength, reduced hip extension range of motion, increased anterior pelvic tilt, back extensor weakness and both increased and decreased hamstring flexibility (46, 82). Considering the conflicting information, high-quality cross-sectional research is required to better understand physical impairments that occur in proximal hamstring tendinopathy. This will help influence the design of rehabilitation programs.

1.9.7 Sport-specific considerations

Several biomechanical factors have been proposed to contribute to the development of proximal hamstring tendinopathy in runners (46, 83). Examples include overstriding, excessive anterior pelvic tilt and forward trunk lean, which have all been proposed to increase the load on the hamstring muscle-tendon unit (46, 83). Close inspection of sporting technique (e.g. running gait) is therefore warranted during the assessment.

1.10 Imaging

The diagnosis of tendinopathy is made clinically and does not necessitate imaging (80). Findings suggestive of tendinopathic changes (termed pathology) on MRI include increased signal uptake, tendon thickening, peritendinous and bone marrow oedema (49, 84). Findings on ultrasound include thickening of the tendon, peritendinous fluid, hypoechoic or heterogeneous echotexture, echogenic foci of calcific tendinopathy, cortical irregularity and concordant pain (58). Asymptomatic individuals also present with tendon pathology on MRI (49). As the relationship between tendon structure and symptoms is poor, imaging is believed to have little role in diagnosis (7).

A disparity between imaging findings and symptoms has been reported in proximal hamstring tendinopathy (84). A retrospective study reviewed 506 MRIs of the proximal hamstring tendon complex from 253 asymptomatic individuals with a mean age of 60 (84). Bilateral pathology was present in 130 patients (52%) (84). Another study examined proximal hamstring appearance on MRI in patients with proximal hamstring tendinopathy and healthy controls (mean age not reported) (85). In this study, they found increased internal T1 and T2 signals in more than 90% of hamstring tendons without symptoms of proximal hamstring tendinopathy. A single study found tendinopathic changes in 25% of proximal hamstring tendons (n=16) in those younger than 45 years (86). The findings of this small study suggest that proximal hamstring tendon changes in younger populations (without symptoms) are less prevalent (86). The results that can be gleaned from these studies are that MRI findings are very common in the asymptomatic population, particularly in older adults. Some caution should be taken when considering these results due to the small sample size (86).

These findings reiterate the importance of associating imaging findings with a clinical examination. Further research in this field is required to better understand the changes in the proximal tendon with age.

1.11 Differential diagnosis

1.11.1 Somatic referral

Various structures including the hip joint, lumbar spine, sacroiliac joint and peripheral nerves, may refer to the buttock region (Table 1.1). The area of symptoms caused by somatic referral is typically diffuse (72). In contrast, symptoms of proximal hamstring tendinopathy tend to be

localised (46), particularly when testing against the resistance of the hamstrings on clinical examination.

1.11.2 Sciatic nerve irritation

Entrapment of the sciatic nerve may occur at various locations along the nerve pathway, such as through the posterior buttocks to below the level of the ischial tuberosity (46, 87). Symptoms arising from the sciatic nerve, as well as common aggravating factors (e.g. pain with sitting), are similar to those reported in patients with proximal hamstring tendinopathy (87). Pain provocation tests include the piriformis stretch test, the active piriformis test (88), straight leg raise and slump test (46, 88). Due to the small distance between the sciatic nerve and hamstring at the ischial tuberosity, it has been proposed that swelling of the proximal tendon, which can occur with tendinopathy, may compress the sciatic nerve and, in some cases, tether the hamstring tendon to the sciatic nerve (62). Tethering of the hamstring tendon to the sciatic nerve has been reported in several case series through surgical inspection (47, 58, 65, 87).

Several other peripheral nerves may refer to the lower gluteal region and should be considered as a differential diagnosis. These include the inferior cluneal nerve, pudendal nerve and posterior femoral cutaneous nerve (Figure 1.7).

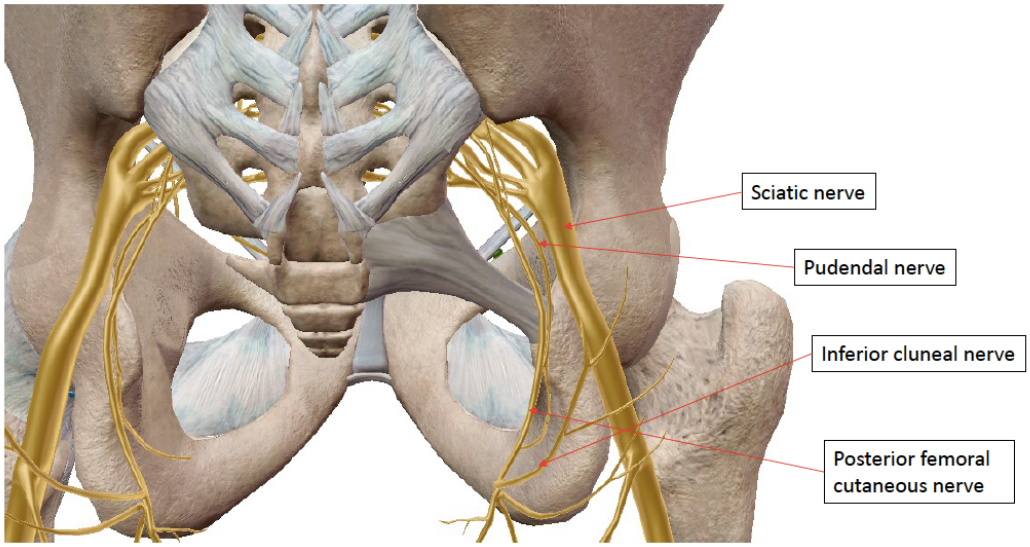


Figure 1.7 Surrounding neural anatomy of the proximal hamstring region
Primal pictures ©

1.11.3 Proximal hamstring ruptures and hamstring strain injuries

The area of the hamstring that is susceptible to injury varies with age. In adolescents, the weakest link in the musculotendinous unit is at the apophyseal attachment (usually 13-16 years) (89). In this age group, common injuries include apophysitis (insidious onset) or avulsion fractures (acute onset) (90). In young adults, the most frequently injured area is the biceps femoris long head, followed by semitendinosus, both of which commonly occur at the musculotendinous junction, and may occur concurrently (91).

Partial tears or ruptures of the hamstring origin can occur from overstretching (e.g. water-skiing) (66). These injuries are less common than injuries to the musculotendinous junction. Injuries to the intra-muscular tendon, which is commonly referred to as the central tendon may also occur (92). These injuries present with a different mechanism of injury to proximal hamstring tendinopathy (acute) and have more variable symptoms (92). The semimembranosus tendon may suffer partial or full thickness tears due to overstretching and are commonly referred to as stretch type hamstring strains (93). This injury is common in dancers (69). The adductor magnus (ischiocondular portion) shares the ischial tuberosity with the hamstring group (94). Tears of the adductor magnus have been reported with the clinical presentation revealing more medial ischial symptoms than those found in proximal hamstring tendinopathy (95).

1.11.4 Spondyloarthropathy

Inflammation of tendinous or ligamentous insertions onto bone is called enthesitis. This is a characteristic finding of spondyloarthropathy (33). Common sites of inflammation at the bone tendon junction are the Achilles tendon and the plantar fascial insertions (96). However, involvement of the ligamentous and tendinous insertions onto the pelvic bones, such as proximal hamstring are also encountered (96). Clinicians should be suspicious of athletes who present with stiffness in the lower back and hip as a secondary complaint to ischial pain (96). Other clinical signs of spondyloarthropathy are summarised in Table 1.1.

1.11.5 Paratendinopathy

Paratendinopathy describes activity-related tendon pain of the tendon sheath (97). Signs and symptoms include pain with movement, tenderness to palpation and swelling around the tendon (97). Paratendinopathy has not been described extensively in the proximal hamstring region (97). Key methods suggested to differentiate paratendinopathy from tendinopathy in the

Achilles region include the palpation of nodules that do not move when the ankle is dorsiflexed, the presence of crepitus with movement and the identification of fluid surrounding the tendon with ultrasound, with the absence of intra-tendon pathology (97). Diagnostic methods to explore paratendinopathy as a differential diagnosis in the proximal hamstring region need to be further explored.

1.11.6 Gluteal tendinopathy

Gluteal tendinopathy is a common cause of lateral hip pain (48). Various terms have been used to describe the condition, such as greater trochanteric pain syndrome and gluteus medius tendinopathy (48). The primary pathology of gluteal tendinopathy is tendinosis of the gluteus minimus and/or medius tendons (98). Bursitis of the trochanteric bursa may occur in conjunction with the condition, but rarely occurs in isolation (99). The condition is commonly aggravated by activities such as sleeping on the affected or unaffected side, walking, ascending stairs and running (48, 100). The condition is most common in post-menopausal women, but does occur in men and younger women (48).

In conjunction with the patients presenting history, gluteal tendinopathy is diagnosed using a battery of pain provocation tests (48, 64). Examples of provocation tests include palpation of the gluteus medius/minimus tendon at the greater trochanter, timed single leg stance (30 seconds), the flexion abduction external rotation test, and the modified external de-rotation test (64, 101). Firm conclusions regarding the diagnostic accuracy of clinical tests are not possible with the use of MRI as a diagnostic reference standard, which is commonly used in the literature, due to the frequency of tendinopathic changes in asymptomatic populations (64, 101, 102).

The effect of various treatments has been assessed in the management of gluteal tendinopathy (103). Common treatments used to manage the condition include injection therapies (corticosteroids, platelet-rich plasma), shockwave therapy, exercise, education and surgery (103). Further high-quality systematic reviews of randomised clinical trials are required to understand which treatment is the most effective.

1.11.7 Other causes of buttock pain

Other causes of buttock pain, which are also insidious in onset, that must be considered as a differential diagnosis for proximal hamstring tendinopathy include ischiofemoral impingement (74, 104), stress fractures of the pubic rami (105), and spondyloarthropathies, such as ankylosing spondylitis (Table 1.1) (33).

1.12 Interventions used to manage proximal hamstring tendinopathy

Treatments for proximal hamstring tendinopathy can be divided into conservative and non-conservative interventions. Conservative management may include exercise therapy, NSAIDs, therapeutic ultrasound and manual therapy. Less conservative options exist on a continuum, and include shockwave therapy, injectables (autologous blood injections, platelet-rich protein, corticosteroid injections) and surgery. Support in the literature for surgical interventions and injection therapies for tendon injuries has weakened over the past decade, with such management options being reserved for recalcitrant cases (106). This is primarily due to the availability of successful non-invasive options, such as exercise (39, 107).

To date the primary outcome measures used in studies on proximal hamstring tendinopathy have been measures of self-reported pain and/or symptoms (108). There are two published patient-reported outcome measures for tendinopathies around the hip and pelvis (109, 110). **Research is required to determine the usefulness of these measures in both research and clinical settings.**

Most of the literature on proximal hamstring tendinopathy reports on the efficacy of invasive interventions, such as surgery and injectables (47, 57, 111). The efficacy of the different interventions is mixed, with shockwave therapy showing the most promising results on physical function and self-reported symptoms (112). **The limited number of studies and small samples sizes used means there is a degree of uncertainty in the interpretation of results**. It is important that available literature is summarised and the quality of evidence is understood.

It is unclear whether the present literature represents current best practice. **Research to understand current best practice will assist in developing future randomised clinical trials that compare the efficacy of different interventions.** Prior to developing expensive

clinical trials, it is critical to have valid, reliable and responsive outcome measures to track responses to treatment. **Research is required to provide researchers and clinicians with robust outcome measures that are valid, reliable and responsive, that reflect the broad impact hip-related tendinopathy has on patients.** This will ensure the impact of tendinopathy of the hip is reflected in research.

1.13 Justification of thesis

The understanding and management of tendinopathy has evolved significantly over the past two decades. However, much of the research on tendon pain in the lower limb has focused on specific sites, namely Achilles and patellar tendinopathy (113, 114). There is a lack of research on proximal hamstring tendinopathy, as evidenced in this chapter. Research is required to provide patients and clinicians with proven management options. Specific information is required before the development of clinical trials. This includes understanding the efficacy of available management options to recognise interventions worth investigating (e.g. current best practice) and to explore the utility of outcome measures that could be used in trials to measure change. This will help ensure the impact of tendinopathies of the hip are captured in research.

1.14 Aims of thesis

Based on the gaps identified in the literature, the aim of this thesis was to: **evaluate the current evidence base for interventions, identify how experts manage the condition and evaluate the use of outcome measures in research of tendinopathy of the hip**. The specific gaps identified will be addressed through the following studies.

Study one Proximal hamstring tendinopathy; a systematic review of interventions.

Aim: to systematically evaluate the literature, and report on the efficacy of invasive and non-invasive interventions on symptoms, physical function and quality of life. Presented as Chapter two.

Study two Proximal hamstring tendinopathy; expert physiotherapists' perspectives on diagnosis, management and prevention.

Aim: to determine what is current best practice in terms of assessment, management and prevention of proximal hamstring tendinopathy. Presented as Chapter three.

Study three Core outcome set development for proximal hamstring tendinopathy (COS-PHT); a study protocol.

Aim: to outline the steps taken to develop a core outcome set for proximal hamstring tendinopathy that can be used in clinical trials. Presented as Chapter five.

Study four Outcome measures in the management of proximal hamstring tendinopathy: a systematic review of their measurement properties.

Aim: to systematically evaluate the measurement properties of outcome measures used to evaluate proximal hamstring tendinopathy. Presented as Chapter six.

Study five Core outcome set for proximal hamstring tendinopathy (COS-PHT); a survey of an international collaboration

Aim: to evaluate the use of outcome measures in studies examining proximal hamstring tendinopathy using the Outcome Measures in Rheumatology (OMERACT) filters of truth and practicality. Presented as Chapter seven

Study six Outcome measures in the management of gluteal tendinopathy; a systematic review of their measurement properties.

Aim: to systematically evaluate the measurement properties of outcome measures used to evaluate gluteal tendinopathy. Presented as Chapter eight.

1.15 Thesis deviations due to Covid-19

An original aim of this thesis was to complete a cross-sectional study to investigate the clinical features of proximal hamstring tendinopathy. Soon after data collection began Covid-19 caused state-wide lockdowns which meant data collection was not possible. The final chapters of this thesis were altered to investigate the use and quality of outcome measures in evaluating proximal hamstring tendinopathy and gluteal tendinopathy.

2 Chapter 2: Proximal hamstring tendinopathy; a systematic review of interventions

2.1 Preface

In the introductory chapter, it was established that various treatments have been used to manage proximal hamstring tendinopathy. It is not known which treatment is the most effective. The premise of this study was to answer the simple question:

"What is the most effective treatment to manage proximal hamstring tendinopathy?"

To answer this question, a systematic search of the literature was performed to capture all types of interventions that have been used to manage proximal hamstring tendinopathy. The primary outcome measures of this systematic review were self-reported symptoms, quality of life and physical function. The secondary outcome measure was adverse events. Additionally, how completely studies reported on the interventions they tested was captured using the TIDiER checklist (Template for Intervention Description and Replication) (115). Capturing information on the completeness of reporting in research is critical for the replication and translation of scientific findings to clinical practice.

The following chapter contains a modified version of the paper

Proximal hamstring tendinopathy; a systematic review of interventions

International Journal of Sports Physical Therapy

Anthony Michael Nasser, Bill Vicenzino, Alison Grimaldi, Jay Anderson, Adam Ivan Semciw

Vol. 16, Issue 2, 2021

URL link to published article: https://doi.org/10.26603/001c.21250

2.2 Abstract

Study design

Systematic review

Background

Proximal hamstring tendinopathy affects athletic and non-athletic populations and is associated with longstanding buttock pain that limits sport participation. Several treatment options have been described in the literature, but it is unclear which is the most effective.

Objective

To assess the efficacy of surgical and non-surgical interventions for treating proximal hamstring tendinopathy.

Methods

Electronic databases were searched through to January 2019. Studies investigating interventions for people with proximal hamstring tendinopathy were eligible for inclusion. Outcomes included symptoms, physical function, quality of life and adverse events. Included studies were screened for risk of bias and overall quality of evidence was rated according to GRADE (Grading of Recommendations Assessment, Development, and Evaluation guidelines).

Results

Twelve trials were included (n=424; males 229). There were two RCTs and ten case series. Interventions included: platelet-rich plasma injection (n=5), surgery (n=4), shockwave therapy (n=1), multi-modal intervention (n=2), autologous whole blood injection (n=1) and corticosteroid injection (n=2). Due to the serious risk of bias within studies and small sample sizes, no conclusions could be made regarding the efficacy of surgery, platelet-rich plasma injection, shockwave therapy, corticosteroid injection or autologous whole blood injection. Surgical studies reported the highest number of adverse events, however this varied considerably between studies, and was inconsistently reported.

Conclusions

As there was no high or moderate-quality evidence for any treatment option, a pragmatic approach would be to initially trial rehabilitation approaches that have proven successful in other tendinopathies. Such programs tend to focus on education, load management and progressive loading of the affected musculo-tendon unit and synergists.

2.3 Introduction

Proximal hamstring tendinopathy presents as persistent buttock pain that occurs with activities such as running, sitting and lunging (47, 112). The condition primarily affects active men and women (81) in sports such as track and field, distance running, soccer and rugby (112). It also afflicts people who do not participate in sport (8, 57). The condition is challenging to manage because of the persistence of symptoms and lack of response to treatment (46, 81).

Interventions for proximal hamstring tendinopathy focus on reducing symptoms and restoring physical function. Common non-surgical interventions include exercise, corticosteroid injection, platelet-rich-plasma injection and shockwave therapy (46, 58). Common surgical procedures include tenotomy, bursal and tendon debridement and removal of adhesions around the sciatic nerve (47, 65).

With many treatment options available, a rational starting point is to systematically review the literature and synthesise information where possible. The aim of this study was to evaluate both surgical and non-surgical interventions used in managing proximal hamstring tendinopathy and understand their impact on symptoms, physical function and quality of life.

2.4 Method

The systematic review protocol was developed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement and preregistered on PROSPERO (ID: CRD42017072678) (116).

2.4.1 Research question

To evaluate surgical and non-surgical interventions used in managing proximal hamstring tendinopathy.

2.4.2 Search strategy

A systematic literature search was conducted from the earliest date available until January 2019 for relevant studies published in MEDLINE, CINAHL, EMBASE, SPORTSDISCUS and PUBMED. A comprehensive search was undertaken, with the assistance of a librarian, using a combination of keywords and medical subject headings (MeSH). The search strategy was formed around two topics; “hamstring” and “tendinopathy”. Synonyms within each concept

were combined with OR Boolean operator; and terms between concepts were combined with AND Boolean operator (Appendix C). Search results for MEDLINE can be found in Appendix D.

After the removal of duplicates, two reviewers (AN and JA) independently scanned titles and abstracts of all papers. Any disagreement was referred to a third reviewer (AS) for consensus. Full-text versions of articles were obtained for all remaining studies. To ensure all relevant articles were included, citation tracking (PubMed/ Google scholar) and reference checking of included studies was performed.

2.4.3 Selection criteria

2.4.3.1 Type of studies

Included: randomised controlled trials (RCTs), prospective (comparative) cohort studies, case-control studies and case series with ten or more participants. Articles were confined to the English language only.

Excluded: biomechanical reports and narrative reviews.

2.4.3.2 Type of participants

Participants of any age diagnosed with proximal hamstring tendinopathy by a health professional were included. Synonyms considered included: hamstring tendinitis, high-hamstring tendinopathy, and hamstring origin tendinopathy. Traumatic injuries such acute proximal hamstring tendon tears, complete hamstring tears or avulsion injuries were excluded.

2.4.3.3 Type of intervention

Surgical and non-surgical interventions were considered in this review. Interventions included, but were not limited to: surgery, shockwave therapy, platelet-rich plasma injections, autologous whole blood injections, corticosteroid injection and multi-modal intervention (NSAIDs, manual therapy, exercise and stretching).

2.4.3.4 Type of outcome

Outcome measures that reported on symptoms, physical function (e.g. return to sport), quality of life (QOL) and ratings of success were included. Short (≤ 12 weeks), medium (> 12 weeks - 1 year) and long (≥ 1 year) time-points were considered. Adverse events were reported as a secondary outcome and were dichotomised as minor or major (117). Minor adverse events

were defined as incidents that had minimal serious or potentially serious effects, such as thigh paraesthesia or a small infection that resolved with anti-microbial drugs. Major adverse events were incidents that had the potential for severe effects, such as deep vein thrombosis or severe infection.

2.4.4 Protocol deviations

We made one amendment to our initial protocol. We originally included all case series. This was amended to exclude case series with <10 participants. This was identified as an oversight in our initial protocol and was revised prior to conducting our search. This amendment resulted in the exclusion of three case series with small sample sizes (118-120).

2.4.5 Data extraction

Pre-specified data was extracted from each study and included eligibility criteria, study design, participant demographics, intervention, outcome parameters, results at all time points and adverse events. Data was extracted by one reviewer (AN) and checked by a second reviewer (AS). Authors were contacted in the case of missing data. If author(s) did not respond, they were contacted again after two weeks. If the author(s) still did not respond raw data was reported.

Means and standard deviations of continuous outcomes for comparative studies (e.g. intervention A vs Intervention B) were converted to standardised mean differences (SMD) with 95% CI using RevMan (version 5.3). When studies reported changes over time, means and standard deviations (SD) were converted to standardised paired differences (SPD) using the 'metafor' package (version 2.1) within the 'R' statistical software package (version 3.5.1). calculations (with 95% CI) require information about pre- and post-test reliability (121, 122). If this was not reported, or could not be located, a conservative estimate was used (r = 0.50). In cases where the SD was not reported, we used the formula provided in the Cochrane Handbook (section 16.1.3.2), using a conservative correlation coefficient of 0.5, to calculate the SD (123). Effect sizes (SMDs and SPDs) of 0.2, 0.5 and 0.8 were considered as small, medium and large respectively (124). When data could not be presented as a SPD, raw data were presented. The proportion of people who returned to pre-injury level of sport was summarized as a percent with 95% confidence interval (CI). Statistical heterogeneity across pooled studies was assessed using the Tau (τ^2) and I^2 statistic. An I^2 value of 25%, 50% or 75% was considered a low, moderate or high level of heterogeneity respectively (125). The

'meta' package, version 4.9-2, of the R statistical software package (version 3.3.1) was used for all statistical analyses (http://www.r-project.org/).

2.4.6 Assessment of risk of bias

Study quality was assessed using separate risk of bias tools, depending on study design (123). The Cochrane Risk of Bias Tool was used to assess RCTs (123). Each of the eight domains were marked as either low risk of bias (+), high risk of bias (-) or unclear risk of bias (?) (123). Studies with the presence of three or more (-) or (?) were considered as having a high risk of bias. The Joanna Briggs Institute Checklist was used to assess the risk of bias in other study designs (126). Items were scored "yes (Y)", "no (N)", "unclear (U)" or "not applicable (N/A)" on the ten-point checklist (126). If an individual study scored ≥ 6 "yes (Y)" scores, it was considered low risk of bias.

2.4.7 Completeness of reporting

Completeness of reporting in intervention and control groups was assessed using relevant items of the Template for Intervention Description and Replication Checklist (TIDieR) (115). Improved reporting provides an opportunity to understand the treatments used and more precisely inform clinicians about the type of interventions, enabling replication of treatments in clinical practice. If items were not relevant for an included study, they were scored not applicable (N/A). For example, completeness of reporting of a control group was scored N/A in studies where there was no control/comparison.

2.4.8 Appraisal of the quality of the body of evidence

Both single group and comparative studies were assessed according to the GRADE guidelines (127). Overall quality was defined as high, moderate, low or very low (127). We followed a staged process whereby RCTs were first rated as high and case series were rated as low-quality evidence (128). Following this step, quality of evidence for individual outcomes were further reviewed, with the potential of being further downgraded by one level for each of the following factors: i) limitations in design (≥25% of the participants from studies with a high risk of bias as determined by the risk of bias tool), ii) inconsistency of results (significant statistical heterogeneity (I^2 >40%) or inconsistent findings across studies (≤75% of the participants report findings in the same direction), iii) indirectness (i.e. generalisation of the findings), iv) imprecision (total number of participants <300 for each outcome and wide confidence intervals) and v) other considerations (e.g. publication bias, flawed design, massive dropout)

(129, 130). Single studies (n<300) were considered inconsistent and imprecise, providing "low-quality evidence", which could be further downgraded to very low-quality evidence if additional items were not satisfied (129, 130).

2.5 Results

2.5.1 Study selection

The PRISMA flow diagram, is shown in Figure 2.1. 1924 studies were identified through database searching. Full texts of 34 studies were assessed for inclusion. Twelve studies met inclusion criteria.

2.5.2 Study characteristics

2.5.2.1 Participants

Participant characteristics, study design and diagnostic criteria are displayed in Table 2.1. A total of 424 (229 males) participants were included. Mean ages in individual studies ranged from 24 to 51 years. One-hundred and ninety-nine (47%) participants were described as professional, competitive or high-level athletes, 75 (18%) were recreational athletes and five participants were non-athletes. Activity levels in 145 (34%) participants were not reported.

2.5.2.2 Diagnostic criteria

Most studies used clinical assessment and imaging to diagnose proximal hamstring tendinopathy (Table 2.1). The clinical assessments varied between studies. Imaging (magnetic resonance imaging or ultrasound) was used to attempt to verify proximal hamstring pathology in 92% of studies.

2.5.2.3 Outcomes

A variety of patient-reported physical function measures and pain measures were used. Mean follow-up times varied from one week to six years. Adverse events were reported in 11 (92%) studies (Table 2.2).

2.5.2.4 Interventions

The most common intervention was platelet-rich plasma injection (111, 131-134). Other interventions included surgery (8, 47, 57, 65), corticosteroid injection (9, 135), autologous

whole blood injection (111), shockwave therapy (112) and multi-modal intervention (NSAIDs, manual therapy, exercise and stretching) (110, 132).

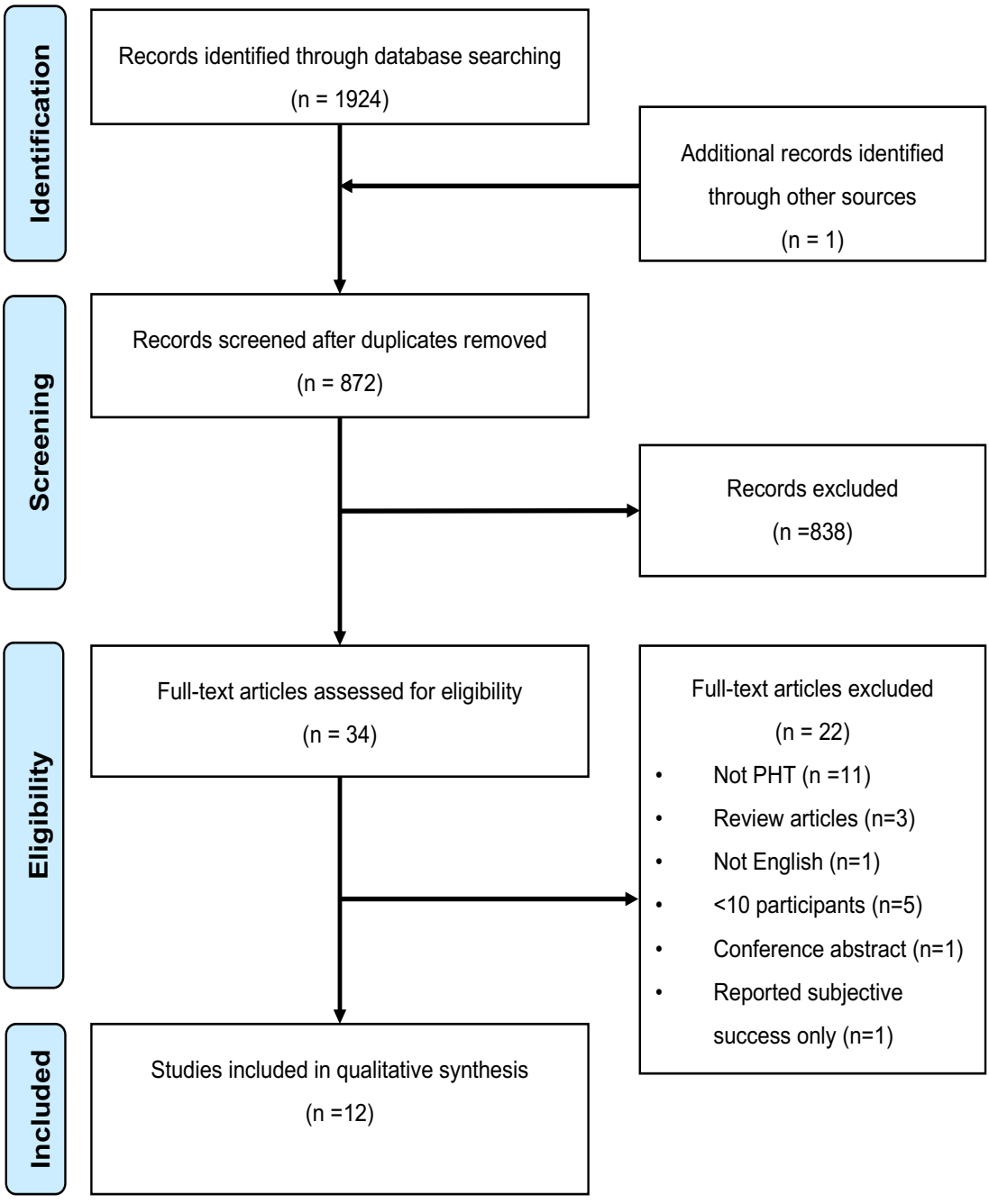


Figure 2.1 PRISMA flow chart illustrating study selection process

Table 2.1 Study design and patient characteristics

Study (Year)	Study design	n (F), mean age (years)	Diagnostic criteria	Tendon involved (%)	Activity level	Average duration of symptoms (range)
Benazzo (2013)	Case series	17 (5F), 27 years	Reporting pain at ischial tuberosity Pain with concentric hamstring contraction: (positive 47%) Tenderness ischial tuberosity (41% positive) Puranen Orava test (88% positive) Leg raising test (23% positive) MRI	Bicep femoris (41%) Semimembranosus (29%) Semitendinosus (6%) Common tendon (23%)	Professional athletes: 9 Competitive athletes: 8	23 months (3-48)
Cacchio (2011)	RCT	40 (13F), SWT 24 years, multi-modal intervention 24 years	At least 2 of the following painful; Puranen Orava test, fast hamstring-stretch test, hamstring strength test MRI	NR	Multi-modal professional athletes: 40	Multi-modal intervention: 21 (13-81) SWT: 19.6 (11-72)
Davenport (2015)	RCT	15 (13F), AWB: 45.4 years, PRP 47 years, 17 cases (2 bilateral)	Clinical diagnosis[a] Positive MRI or US	NR	NR	NR
Fader (2014)	Case series	18 (12F), 43 years	Clinical diagnosis[a] MRI	NR	NR	32.6 months (6-120)
Krauss (2016)	Case series	14 (13F), 47 years	At least 2 of the following; Tenderness to palpation at ischial tuberosity, positive bent knee stretch test, positive supine plank test MRI	NR	NR	4.1 years (5 months to 10 years)
Lempainen (2009)	Case series	90 (32F), 34 years, 103 cases (13 bilateral)	Patient interview and history Pain at ischial tuberosity with hamstring stretch MRI	All semimembranosus	Professional: 5 Competitive: 47 Recreational: 38	NR
Levy (2019)	Case series	29 (22F), 45 years	Clinical diagnosis[a] confirmed with positive MRI	NR	NR	NR
Nicholson (2014)	Case series	18 (10F), 51 years, 22 cases (4 bilateral)	Pain with prolonged sitting Pain with hamstring contraction MRI	NR	Athletes: 18	28 months (2-120)

Puranen (1988)	Case series	59 (14F), Athletes: 25 years, joggers: 39 years, non-athletes: 35 years, 65 cases (6 bilateral)	Pain on active stretching Pain on palpation	NR	Athletes: 50 Joggers: 4 Non-athletes: 5	NR
Wetzel (2012)	Case series	15 (8F), 38 years, 17 cases (2 bilateral)	Clinical diagnosis[a] MRI	NR	PRP group: Collegiate or high-level athletes: 9 NR: 1 Multi-modal intervention group: High level athletes: 2 NR: 3	Multi-modal intervention: 7.8 months PRP: 9.6 months
Young (2008)	Case series	44 (16F), 29 years, 47 cases (3 bilateral)	Pain on palpation of proximal hamstring region Weakness at 30º of resisted knee flexion MRI/US	NR	Professional: 4 Semi-Professional: 7 Recreational: 33	8 cases: < 6 months 15 cases: 6-12 months 10 cases: 12-18 months 14 cases: >18 months
Zissen (2010)	Case series	65 (37F), 37.7 years	Clinical diagnosis[a] MRI/US	NR	NR	8 cases: < 6months 15 cases: 6 months to 1 year 15 cases: > 1 year

AWB: autologous whole blood, F: female, MRI: magnetic resonance imaging, NR: not reported, PRP: platelet-rich plasma, US: ultrasound
[a] Clinical test(s) performed in examination not specified

Table 2.2 Results

Study (Year)	Intervention	Mean follow-up (range)	Symptoms Effect size (CI) unless otherwise stated (-ve indicates reduction in pain)	Physical function Mean effect size (CI) unless otherwise stated	Adverse effects (number %): minor, major
Benazzo (2013)	Surgery: Prone incision from ischial tuberosity to 8-15cm distally. Partial transverse tenotomy performed plus sciatic nerve release.	71.3 months (24-138)	Patient-reported pain rating: VAS: • -6.02 (-8.10, -3.94) SPD[a]	Patient-reported physical function: Tegner score: • -4.08 (-5.53, -2.63) SPD[c] Return to sport (pre-injury level): • Proportion: 17/17 (100%) returned to pre-injury level at average of 4.4 (range 2-9 months)	Minor: • 1 (6%) Major: • 1 (6%)
Cacchio (2011)	Multi-modal intervention: NSAIDs, manual therapy, exercise and ultrasound SWT: 2500 impulses at 0.18 $mJmm^2$/ frequency = 10 shocks in 4 separate sessions at weekly intervals.	1 week, 6 months, 12 months	Patient-reported pain rating: VAS: • Week 1: -1.84 (-2.59, -1.09) SMD[a] • 6 months: -3.23 (-4.28, -2.19) SMD[a] • 12 months: -3.22 (-4.28, -2.16) SMD *-ve favours SWT*	Patient-reported physical function: NPRS: • 1 week: -3.09 (-4.04, -2.15) SMD • 6 months: -2.90 (-3.88, -1.92) SMD • 12 months: -2.42 (-3.33, -1.50) SMD *-ve favours SWT* Return to sport (pre-injury level): Proportion: • SWT: 16/20 (80%) mean time of 9 weeks (range, 6-15 weeks). • Multi-modal intervention: 0/20 (0%)	Minor: • SWT: 0 (0%) • Multi-modal intervention: 0 (0%) Major: • SWT: 0 (0%) • Multi-modal intervention: 0 (0%)
Davenport (2015)	PRP vs AWB: Single U/S guided injection of AWB (5mL) or PRP (3mL).	12 weeks, 6 months	NR	Patient-reported physical function: HOS (ADL): • 6 weeks: -0.56 (-1.58, 0.45) SMD • 12 weeks: 0.03 (-0.96, 1.03) SMD • 6 months: NR *-ve favours PRP* HOS (Sport): • 6 weeks: -0.21 (-1.24, 0.83) SMD • 12 weeks: 0.28 (-0.76, 1.32) SMD • 6 months: 0.17 (-0.86, 1.21) SMD *-ve favours PRP* MHHS: • 6 weeks: -0.28 (-1.28, 0.72) SMD • 12 weeks: 0.05 (-0.95, 1.04) SMD	Minor: • PRP: 1 (10%) • AWB: 0 (0%) Major: • PRP: 0 (0%) • AWB: 0 (0%)

				• 6 months: 0.24 (-0.76, 1.24) SMD *-ve favours PRP* Health related QOL: i-HOT 33: • 6 weeks: -0.80 (-1.86, 0.26) SMD • 12 weeks: 0.01 (-1.00, 1.03) SMD • 6 months: -0.04 (-1.05, 0.97) SMD *-ve favours PRP*	
Fader (2014)	PRP: A single U/S guided injection (2.5-4mL).	46 months	Patient-reported pain rating: VAS (mean pain): • -2.9 MD[a] Patient-reported rating of symptom improvement: • 10 (55.6%) patients had an 80% or greater improvement at 6 months	NR	Minor: • 1 (6%) Major: 0 (0%)
Krauss (2016)	PRP: Single U/S guided injection (4mL) + needle tenotomy (5 passes).	12 weeks	Patient-reported pain rating: VAS (mean pain): • -0.92 (-1.54, -0.29) SPD[a]	Patient-reported physical function: LEFS: • -0.90 (-1.52, -0.28) SPD[c]	NR 4 (29%) had worse physical function at 12 weeks (LEFS)
Lempainen (2009)	Surgery: Prone, transverse or longitudinal incision. Transverse tenotomy performed on semimembranosus + adhesions freed around sciatic nerve as required.	49 months (range, 12-156 months)	NR	Return to sport (pre-injury level): • Proportion: 80/90 (89%) mean 5 months	Minor: • 3 (3%) Major: • 1 (1%)
Levy (2019)	PRP: Single U/S guided injection (6mL)	8 weeks	NR	Patient-reported physical function: • VISA H: -0.44 SPD (-0.82, -0.06) Return to Sport (pre-injury level): • Proportion: 3/29 (10%) at 8 weeks	Minor: • 1 (3%) Major: • 0 (0%)
Nicholson (2014)	CSI: Single injection under fluoroscopic guidance	21 months (VAS) 24.8 months (LEFS)	Patient-reported pain rating: VAS: • -3.28 MD[a]	Patient-reported physical function: LEFS/80: • 60 (48, 72)[e] Level of function (% of full normal activity): • Increased from 28.76% to 68.82%	Minor: • 0 (0%) Major: • 0 (0%)
Puranen (1988)	Surgery	24 months (24-96)	NR	Return to Sport: • Proportion: 52/59 (88%) [f]	Minor: • 4 (6%)

	65% modified Kocher incision with patient lying on unaffected side. 35% straight incision over lower edge of gluteus maximus with patient prone. Taut structures of hamstring at ischial tuberosity divided and tendon ends separated, freeing sciatic nerve.				Major: • 0 (0%)
Wetzel (2012)	Multi-modal intervention (physical therapy + NSAIDs) vs failed multi-modal intervention + single volume PRP injection (6cc) plus additional multi-modal intervention	Multi-modal intervention: 2 months Failed multi-modal intervention (PRP + multi-modal intervention) : 4.5 months	Patient-reported pain rating: VAS: • Multi-modal intervention: -6.2 MD[a] • Failed multi-modal intervention (PRP + multi-modal intervention): -7.5 MD[a]	Patient-reported physical function: NPRS: • Multi-modal intervention: -2.2 MD[c] • Failed multi-modal intervention (PRP + multi-modal intervention): -4 MD[c] Return to sport (pre-injury level): Proportion: • Multi-modal intervention: 2/11 (18%) • Failed multi-modal intervention PRP + multi-modal intervention: 9/9 (100%)	Minor: • 0 (0%) Major: • 0 (0%)
Young (2008)	Surgery Semi-prone, incision in gluteal fold to gluteus maximus dissection continued toward ischial tuberosity. Sciatic nerve freed and prominent bursal tissue removed.	53 months (range, 9-110)	Patient-reported pain rating: VAS: • -1.89 (-2.36, -1.41) SPD[a]	Patient-reported physical function: Weakness score (/10): • -3.69 (-2.76, -4.62) MD[c] *-ve indicates reduction in weakness* Return to sport (pre-injury level): • Proportion: 34/44 (77%)	Minor: • 11 (23%) Major: • 0 (0%)
Zissen (2010)	Single CSI to area of maximum pain under U/S guidance.	48 months (6-96)	Patient-reported pain rating: Number (%) of patient's symptoms resolved: • Complete: 11 (29%), moderate: 8 (21%), mild: 10 (26%), no: 9 (24%)	NR	Minor: • 0 (0%) Major: • 0 (0%)

AWB: autologous whole blood, CI: confidence interval, HOS: Hip Outcome Score, i-HOT: International Hip Outcome Tool, LEFS: Lower Extremity Functional Scale, MHHS: Modified Hip Harris Score, NR: not reported, NRPS: Nirschl Phase Rating Scale , PRP: platelet-rich plasma, RTS: return to sport, SPD: standard paired difference, SMD: standard mean difference, SWT: shockwave therapy, U/S: ultrasound, VAS: Visual Analogue Scale, VISA H: Victorian Institute of Sport Assessment - proximal hamstring tendon

[a] Negative indicates an improvement in pain

[b] Positive indicates improvement in weakness score

[c] Negative indicates an improvement in physical function

[d] Adverse effects not reported on separately from complete tendon ruptures within study

[e] Measure taken post-operatively only

[f] Level of return to sport not reported

2.5.3 Completeness of reporting and risk of bias

Name and description (item 1) and mode of delivery [item 6 (a) and (b)] were reported completely in both intervention and control groups (Figure 2.2). Intervention adherence (TIDieR items 11 and 12) was not reported in any study. No studies provided adequate details of post-surgical protocols. Risk of bias for RCTs and case series are reported in Figure 2.3 and Figure 2.4 respectively.

2.5.4 Evidence synthesis

2.5.4.1 Randomised controlled trials

2.5.4.1.1 Shockwave therapy vs multi-modal intervention

We found very low-level evidence from a single RCT that shockwave therapy was more effective than a multi-modal intervention (one week of rest plus NSAIDs, two weeks of manual therapy and therapeutic ultrasound, and three weeks of exercise - including strength training and stretching) by a large effect in the short (-1.84 SMD; 95% CI -2.59, -1.09), medium (-3.23 SMD; 95% CI -4.28, -2.19) and long-term (-3.22 SMD; 95% CI -4.28, -2.16) on self-reported symptoms (Appendix C) (112). There was also very low-level evidence that shockwave therapy was more effective than the multi-modal intervention by a large effect in the short-term (-3.09 SMD; 95% CI -4.04, -2.15), medium (-2.90 SMD; 95% CI -3.88, -1.92) and long-term (-2.42 SMD; 95% CI-3.33, -1.50) on physical function (Appendix E) (112). Sixteen athletes (80%) returned to pre-injury level of sport, at a mean time of nine weeks (range; 6-15 weeks) post shockwave therapy intervention (Figure 2.5). No participants in the multi-modal intervention group returned to sport at one year (112).

2.5.4.1.2 Platelet-rich-plasma injection vs autologous whole blood injection

There was very-low level evidence from a single RCT of no significant difference between platelet-rich plasma and autologous whole blood injection on physical function in the short [(0.03 SMD; 95% CI -0.96, 1.03) to (0.28 SMD 95% CI -0.76, 1.32)] and medium-term [(0.17 SMD; 95% CI -0.86, 1.21) to (0.24 SMD; 95% CI -0.76, 1.24)]. There was also very low-level evidence of no difference between interventions on quality of life in the short (0.01 SMD; 95% CI-1.00, 1.03) and medium-term (-0.04 SMD; 95% CI -1.05, 0.97) (111). One complication occurred after platelet-rich plasma injection (irritation of the sciatic nerve) (111).

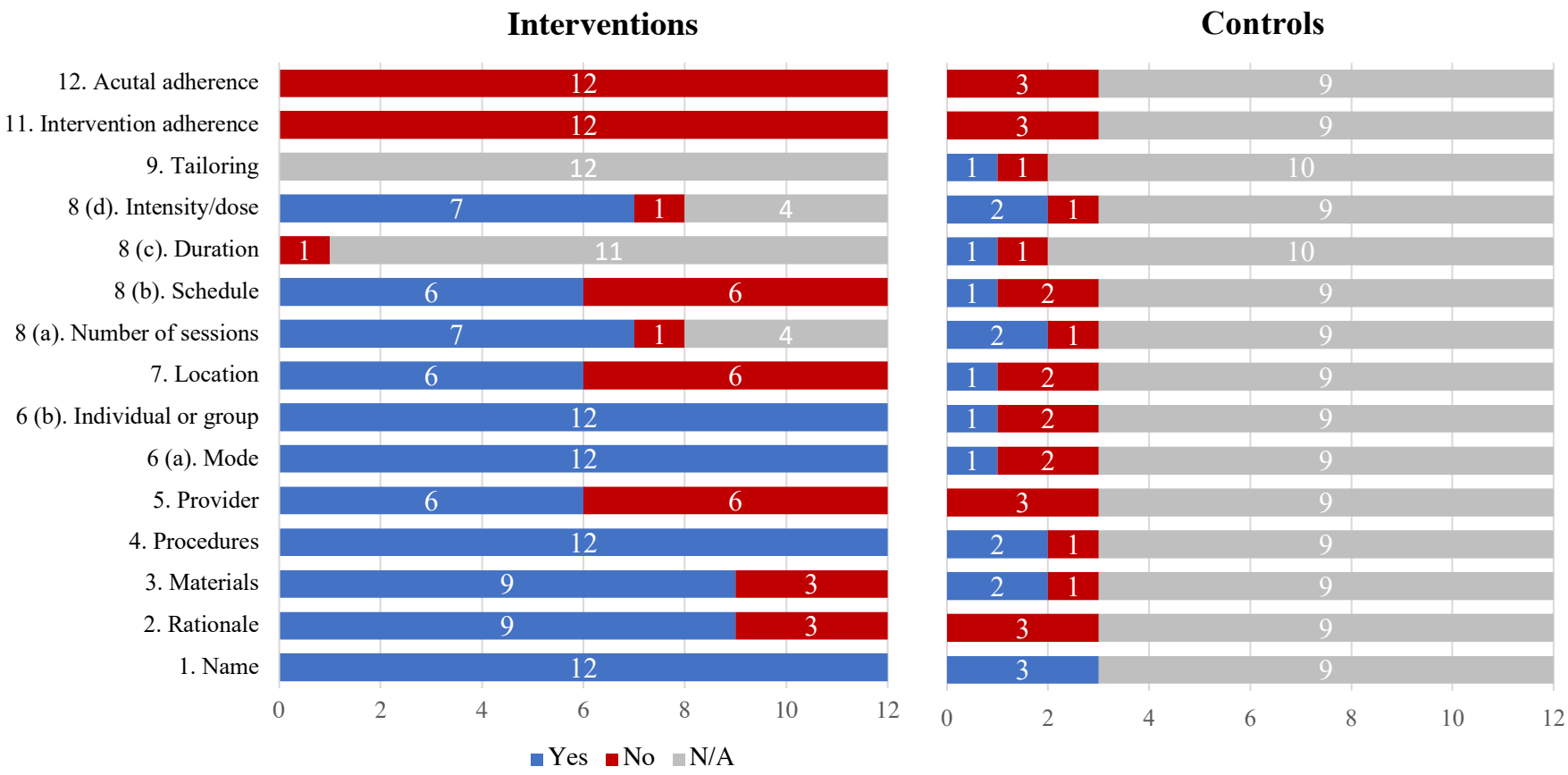


Figure 2.2 Reporting of Intervention and control groups

Percentage of trials achieving each TIDieR item in intervention and control groups of included studies (/12). Each item was scored as 'yes', 'not applicable 'or 'no' for both the intervention and control groups in each included study.

Item 9 was applied to exercise interventions only. Item 10 of the TIDiER scale was not reported on, as it was not able to be assessed in included studies.

In studies that reported on surgery or injection interventions item 11 and 12 were assessed on reporting on adherence to post-surgical protocol.

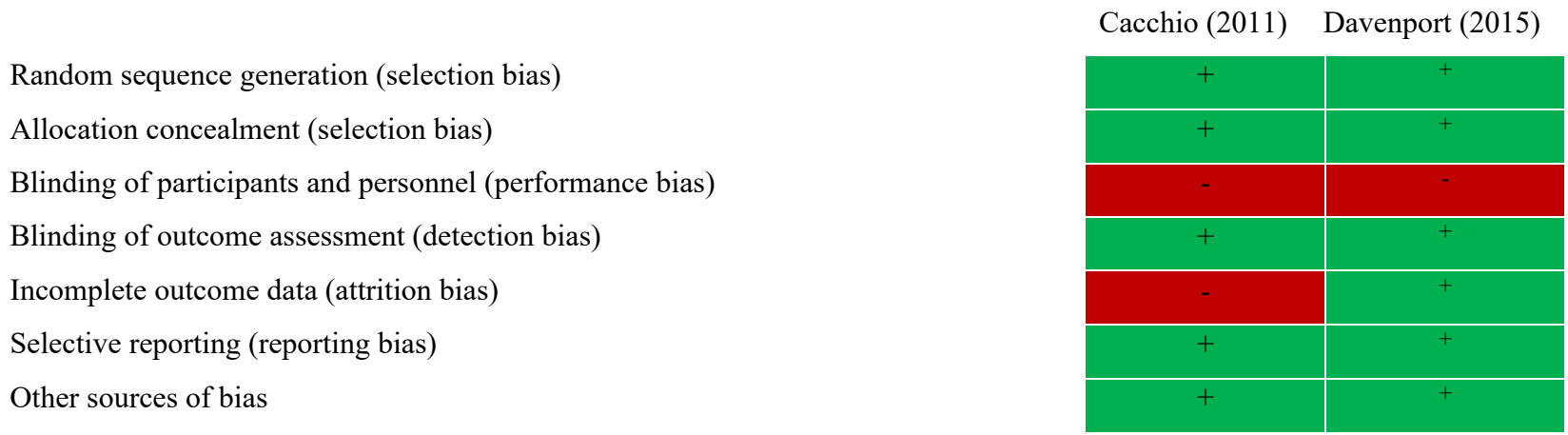


Figure 2.3 Risk of bias in RCTs

Risk of bias for each domain for each RCT. – indicates high risk of bias, + indicates low risk of bias, ? indicates unclear risk of bias

	Benazzo (2013)	Fader (2014)	Krauss (2016)	Lempainen (2009)	Levy (2019)	Nicholson (2014)	Puranen (1988)	Wetzel (2012)	Young (2008)	Zissen (2010)
Clear criteria for inclusion?	N	N	Y	N	U	N	N	N	N	N
Condition measured in a standard, reliable way?	Y	U	Y	U	U	U	U	U	U	U
Valid methods used for identification of condition?	Y	N	Y	N	U	N	N	N	N	N
Consecutive participants?	U	U	U	U	U	U	U	U	U	Y
Complete inclusion of participants?	U	U	U	N	Y	U	U	U	U	N
Reporting participant demographics?	Y	N	N	Y	N	Y	Y	Y	Y	N
Reporting of clinical information?	Y	Y	Y	N	Y	N	Y	Y	Y	Y
Outcomes or follow-up results reported?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Reporting of site(s)/clinic(s) demographics?	N	Y	N	Y	Y	Y	N	Y	N	N
Statistical analysis appropriate?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Figure 2.4 Risk of bias in case series

Items 1-10 were scored as Y = yes, N = no or U = unclear or not applicable = N/A

2.5.4.2 Case series

2.5.4.2.1 Platelet-rich-plasma injection

Three studies, recording the results of platelet-rich-plasma injection alone, documented the change in self-reported symptoms over time (Table 2.2) (131, 132, 134). Outcomes could not be pooled due to heterogeneity in measures and incomplete data reporting. There was very low-level evidence that PRP is associated with a large reduction in symptoms over the short-term (-0.92 SPD; 95% CI -1.54, -0.29) (134). A second study found very low-level evidence that platelet-rich plasma injection was associated with a small, clinically significant, reduction in self-reported symptoms (Table 2.2) (131). Two studies reported changes in physical function following platelet-rich plasma injection(s) over time (Table 2.2) (133, 134).

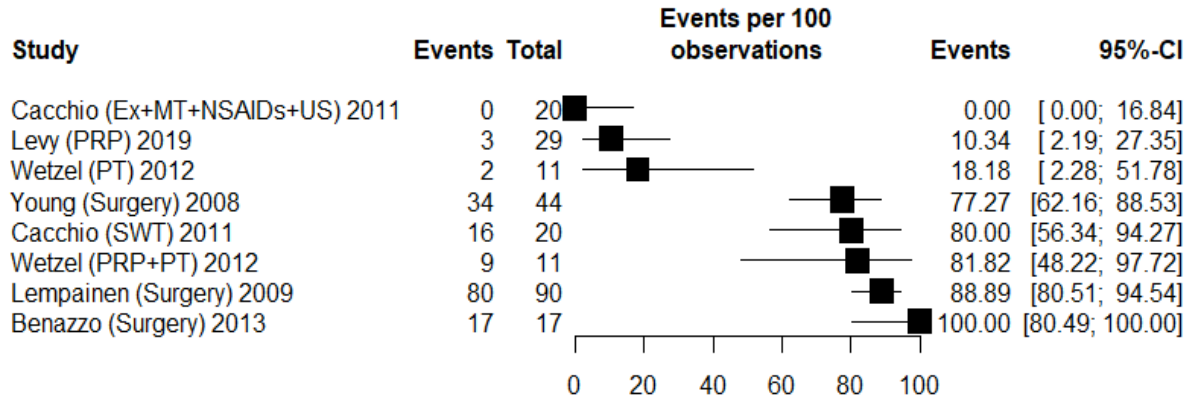


Figure 2.5 Return to pre-injury level of sport

Demonstrates the proportion of participants returning to pre-injury level of sport
Ex: exercise, MT: manual therapy, PT: physical therapy, PRP: platelet-rich plasma injection, SWT: shockwave therapy, US: ultrasound
In Wetzel et al. (2012) - patients who failed initial PT (physical therapy + NSAIDs) went on to receive a single volume PRP injection (6cc) plus additional physical therapy

There was very low-level evidence of a small (-0.44 SPD; 95% CI -0.82, -0.06) (133) to large (-0.90 SPD; 95% CI -1.52, -0.28) (134) improvement in physical function in the short-term (Table 2.2). Ten per cent of participants returned to their pre-injury level of sport after platelet-rich plasma injection at 8 weeks (Figure 2.4) (133). Two (4%) minor adverse events occurred (short-term high levels of pain post-injection) (131, 133).

A single retrospective case series investigated the change in outcomes before and after PRP, in those who had failed a multi-modal intervention (physical therapy + non-steroidal anti-inflammatory drugs) (132). Failure was determined by the persistence of symptoms and an inability to return to baseline activity after the multi-modal intervention had been completed (132). There was very low-level evidence of an improvement in symptoms and physical function over time (Table 2.2) (132). In this study, all athletes who were unable to return to sport with multi-modal intervention alone, returned to sport post PRP plus multi-modal intervention (Figure 2.4) (132).

2.5.4.2.2 Surgery

Two studies reported change in self-reported symptoms after surgery (Table 2.2) (47, 65). There was very low-quality evidence of a large reduction in symptoms in the long-term [(-1.89 SPD; 95% CI -2.36, -1.41) to (-6.02 SPD; 95% CI -8.10, -3.94)] (47, 65). There was very low-quality evidence from a single study of a large improvement in physical function in the long-term (4.08 SPD; 95% CI -5.53, -2.63) (47). Return to pre-injury level of sport following

surgery ranged from 77-100% (Figure 2.5) (8, 47, 57, 65), with two studies reporting on mean return to sport time to pre-injury level (4.4 to 5 months) (47, 57). One study reported that 88% of participants returned to sport, but did not report whether participants returned to their pre-injury level (8).

Adverse events were reported in all studies investigating outcomes of surgery (Table 2.2) (8, 47, 57, 65). One study did not provide sufficient detail to allow events to be dichotomised as minor or major (8). Complications occurred in 21 (10%) cases (Table 2.2). The type and severity of complications varied. Nineteen (9%) patients experienced minor complications. Examples included: notable post-operative soreness (57) and intermittent thigh paraesthesia that resolved spontaneously (47). Two (1%) patients experienced major complications, which included wound abscess requiring surgical drainage (47) and deep vein thrombosis (57).

2.5.4.2.3 Corticosteroid injection

Management of proximal hamstring tendinopathy with corticosteroid injection was described in two studies (9, 135). Both described change in self-reported symptoms over time (Table 2.1) (9, 135). One study provided very low-level evidence of clinically significant improvement in self-reported symptoms in the long-term (Table 2.2) (135). Both studies reported on long-term symptom resolution (9, 135). One study found that 56% of patients did not experience improvement greater than three months (135). The other study found that 56% still reported symptoms at long-term follow-up (9). A single study reported on physical function (post-intervention only), as a percentage of full activity, and found activity level increased by 40% in the long-term (mean 24.8 months) (135). No adverse events were reported (9, 135).

2.6 Discussion

Our primary aim was to evaluate surgical and non-surgical interventions used to manage proximal hamstring tendinopathy. Three main findings emerged from the systematic review: i) there is a lack of rigorous RCTs comparing treatment interventions, ii) patient selection criteria and use of outcome measures are inconsistent and iii) there is inadequate description of treatment interventions used.

Of the twelve studies that met inclusion criteria, only two were RCTs. Both were confounded by small sample sizes. We found no high or moderate-quality evidence for any intervention (symptoms, physical function or QOL). There was very low-level evidence from a single study to suggest that shockwave therapy was more effective than a multi-modal intervention (exercise, manual therapy, NSAIDs, ultrasound) in professional athletes, at both medium and long-term time points (112). Whilst there were limitations in the completeness of reporting on adverse events, we found surgery resulted in the highest level of minor and major adverse events.

2.6.1 Surgery

In managing tendinopathy, surgical interventions are typically reserved for recalcitrant cases that have not yet been resolved with other interventions. Whilst there were improvements over time following surgery (symptoms and physical function) in case series studies without comparator groups, we do not know if these are real treatment effects or whether results are caused by other factors, such as natural history or postoperative rehabilitation. An insight into the likely treatment effects of surgery may be gleaned from a recent systematic review of upper and lower limb tendinopathy, which found surgery was not superior to sham surgery or physiotherapy on pain, function, range of motion and success ratings in the longer term (136). Recommendations given by the authors were that surgery should not be seriously entertained until 12 months of an evidence-based loading program has been credibly trialled (136). We propose that these recommendations are applied to proximal hamstring tendinopathy, until such time as there are adequately designed comparator studies.

2.6.2 Injection therapies

Injection therapies may be an attractive option because they are less invasive than surgery (137). Whilst multiple types of injections were reported (platelet-rich plasma, corticosteroid and autologous whole blood), the overall quality of evidence for all injections in proximal hamstring tendinopathy was found to be low or very low (Appendix E). Consequently, at this stage it is not possible to recommend any type of injection over another or no injection.

Corticosteroid injections are a commonly prescribed treatment for tendinopathy. Whilst it is important to consider the limitation in overall quality of evidence for corticosteroid injection in our review, the findings in single group case series indicate a positive change in symptoms

over time in the short-term (Table 2.2). Systematic reviews of high-quality RCTs in tendinopathies, which compared corticosteroid injection to other interventions, found a similar trend of a positive short-term effect on symptoms, that are nullified in the medium and long-term (138, 139). Whilst this early improvement is desirable, it is worth noting that corticosteroid injection has been associated with delayed healing compared to wait and see (lateral epicondylalgia) (140, 141) and increased collagen disorganization and necrosis in vitro (142).

The popularity and cost of regenerative therapies warrants continued research to improve the evidence base (143). This systematic review found no high or moderate-quality evidence for platelet-rich plasma injection; therefore its utility remains uncertain in this condition. Its effect on symptoms in other tendinopathies was recently summarised in a systematic review (144). The authors reported platelet-rich plasma injections were more efficacious than alternative injections on pain severity in tendinopathies (0.47 CI 95% 0.22 to 0.72) (144). However, it is worth noting that there were several limitations, including the type of injections used as comparisons (e.g. corticosteroid injection or saline) and the tendon involved.

2.6.3 Shockwave therapy and multi-modal interventions

Shockwave therapy has been proposed for a host of upper and lower limb tendinopathies (112, 145-149) and other musculoskeletal conditions (150, 151). Our review found preliminary support that shockwave therapy was superior to a multi-modal intervention (exercise, manual therapy, NSAIDs, ultrasound) for improving symptoms and physical function. However, as there was only a single study on shockwave therapy in proximal hamstring tendinopathy, with a small sample size, the results may not be representative of the wider population (112). Furthermore, there are several factors that may have made the comparison group (multi-modal intervention) less effective. In this study, exercise was included for 3-weeks of the 6-week program. In the literature, successful rehabilitation programs for other lower limb tendinopathies consistently report significantly longer timeframes (107, 152, 153). Avoiding excessive compressive load on the tendon at the enthesis has also been proposed to be a consideration in early to mid-stage exercise selection (32, 46). In the multi-modal intervention (exercise, manual therapy, NSAIDs, ultrasound), exercises selected likely placed the insertion of the proximal hamstring tendon under high levels of compression, early in rehabilitation (lunge, hamstring stretch, exercise bike) (112). For further information regarding a

management program consistent with these parameters, readers are directed to a narrative review (46).

2.6.4 Limitations and directions for future research

A limitation of our review was the lack of high-quality trials with consistent patient outcome measures, inclusion criteria and time points. Our findings highlight the need for consensus on patient selection criteria, outcome measures and frequency of follow-up, to allow the pooling of data. Another limitation was the high number of pre-post study designs. We have calculated SPDs in an attempt to provide a standardised measure across these study designs. It is important that these SPDs are not misconstrued as treatment effects, because there were no randomised comparisons that remove confounders such as regression to the mean, natural recovery and testing. Consequently, we discussed these studies separately from comparative designs (154). Interventions such as load management, heavy slow, strength training, platelet-rich plasma vs placebo and shockwave vs sham shockwave should be avenues for future research. Future research should prospectively report post-surgical protocols and adherence to interventions.

2.7 Conclusion

There was very low-level evidence that shockwave therapy was more effective than a multi-modal intervention. There was very low-level evidence of no difference between autologous whole blood injection and platelet-rich plasma injection. The results of this systematic review highlight the need for high-quality studies and the standardisation of selection criteria, outcomes and reporting across studies. This will assist in determining the most effective option for the management of proximal hamstring tendinopathy.

2.8 Summary of findings

This systematic review aimed to evaluate the efficacy of different interventions in managing proximal hamstring tendinopathy. Twelve studies were included – two were RCTs. Both had small sample sizes (111, 112). There was low-level evidence that shockwave therapy was more effective than a multi-modal intervention in the medium and long-term (physical function and self-reported symptoms) (112). No intervention was supported by high or moderate-quality evidence.

The representation of studies outlined in this systematic review on invasive options, such as surgery or injection therapies was high, considering rehabilitation is commonly advocated as a first-line treatment for tendinopathy (7). As the single RCT involving rehabilitation included only three weeks of exercise, further research is required to assess the efficacy of exercise-based management performed over a longer duration (112). Specific exercise parameters of rehabilitation delivered including total exercise dose may be a factor in treatment outcomes. Further research is required to understand current best practice for managing proximal hamstring tendinopathy, such as exercise parameters, which can be later tested in RCTs.

Other than the lack of high-quality RCTs comparing interventions, several other findings emerged from the systematic review (155). Firstly, **patient inclusion criteria were inconsistent between studies.** Having clear and consistent inclusion criteria that are clearly described is critical to ensure all patients recruited in the study and compared between studies have similar characteristics. This is particularly important for tendinopathy where the diagnosis is based on clinical findings, with no diagnostic gold standard (2). A recent Consensus process in 2018, International Scientific Tendinopathy Symposium (ICON), reported that the following data should be reported for cases and controls in tendinopathy: sex, age, standing height, body mass, history of tendinopathy, whether imaging was used to confirm pathology, loading tests, pain location, symptom duration and severity, level of disability, comorbidities, physical activity level, recruitment source and strategies, and medication use history (80). Whilst most studies included were published before this consensus statement was published, this level of description was not included in any of the studies in the systematic review (155).

Separate to the lack of clarification of subjects included in studies, we identified a lack of completeness of reporting in studies, using the TIDiER checklist (115). Poor reporting of interventions reduces the transparency of research and can hinder the translation of evidence to clinical practice (156). **Outcome measures used in studies were inconsistent**. There were eleven unique outcome measures used in the 12 included studies. Adequate reporting of data and consistency in the selection of measures, within similar health domains, are essential for the comparison of outcomes. **Further research is needed to evaluate the quality of outcome measures used to assess proximal hamstring tendinopathy.** This will ensure measures used to capture the effect of proximal hamstring tendinopathy are meaningful to patients and responsive enough to reflect either improvement or worsening of the condition.

3 Chapter 3: Proximal hamstring tendinopathy; expert physiotherapists' perspectives on diagnosis, management and prevention

3.1 Preface

The findings from the systematic review in Chapter two (study one) highlighted several gaps in the literature. Of note was the void of literature on exercise-based interventions (155). Considering the scarcity of evidence to inform decisions from the literature, Chapter three (study two) aimed to answer the question:

"How do expert physiotherapists diagnose, manage and prevent proximal hamstring tendinopathy?"

To achieve this aim, a qualitative study was designed that involved interviewing expert physiotherapists and exploring their thoughts on diagnosis, treatment and prevention.

The following chapter contains a modified version of the paper

Proximal hamstring tendinopathy; expert physiotherapists' perspectives on diagnosis, management and prevention

Nasser AM, Pizzari T, Grimaldi A, Vicenzino B, Semciw AI

Physical Therapy in Sport

Vol 48, p-67-75, 2021

URL link to published article: https://doi.org/10.1016/j.ptsp.2020.12.008

3.2 Abstract

Objectives

To explore and summarise expert physiotherapists' perceptions on their assessment, management and prevention of proximal hamstring tendinopathy.

Methods

We conducted semi-structured interviews with expert physiotherapists until data saturation was met (n=13). Interviews were transcribed verbatim, and data were analysed systematically and organised into categories and sub-categories according to study aims.

Results

Experts report using a clinical reasoning-based approach, incorporating information from the patient interview and results of clinical load-based provocation tests, in the physical examination to diagnose proximal hamstring tendinopathy. Experts manage the condition through education and progressive loading targeting the hamstring unit and kinetic chain, avoiding provocative activities in positions of compression in early-mid stage rehab and a gradated and controlled return to sport. Passive therapies including injection therapies and surgery were believed to have limited utility. Prevention of recurrence primarily involved continuation of hamstring and kinetic chain strengthening programs and management of physical workload.

Conclusion

Experts rely on a combination of information from the patient interview and a battery of pain provocation tests to diagnose proximal hamstring tendinopathy. Education and graded exercise of the hamstring group and synergists, minimising early exposure to hip flexion, were the foundation of management of the condition.

3.3 Introduction

Proximal hamstring tendinopathy affects athletic and non-athletic populations and is associated with longstanding ischial pain (46, 58, 81). The research on prevalence is limited, however proximal hamstring tendinopathy has been consistently identified in sports involving running such as Australian Rules football, tennis and track and field, as well as sedentary populations (46, 47, 112). Pain is frequently aggravated by activities such as hill running and sitting (46, 58, 112).

A recent systematic review identified multiple potential interventions, including exercise, corticosteroid injection, platelet-rich plasma injection, shockwave therapy and surgery (155). This review reported a lack of unbiased estimates of strong treatment effects to guide treatment selection (155). With no high-quality evidence, clinicians are left with a lack of direction to guide management (46).

In the absence of strong empirical evidence, qualitative research provides insight into the expert's clinical reasoning process and assists in understanding the decisions and complexities faced in current practice (157, 158). This expert opinion can be used to assist in the development of management protocols, which can be used clinically and tested in randomised clinical trials. Our aim was to explore and then summarise expert physiotherapists' perceptions on their assessment, management and prevention of proximal hamstring tendinopathy.

3.4 Methods

3.4.1 Study design

We conducted semi-structured interviews consisting of questions devised a priori to gauge the opinion of expert physiotherapists on clinical aspects of proximal hamstring tendinopathy (Appendix F). Our data were analysed continuously during collection to establish when saturation was reached – indicating that no new information or themes were observed. Ethical approval was obtained by the University of Queensland Human Ethics Committee (ID 2018001158).

3.4.2 Participants

We selected participants using purposeful sampling, with authors ensuring the sample of experts were from a range of geographical locations and had experience across different

sporting populations. Participant inclusion criteria were decided a priori. Expert physiotherapists were selected by the investigators as individuals who had published in the topic area and/or had extensive clinical experience in treating patients with proximal hamstring tendinopathy. Expert physiotherapists (participants) were also required to be: i) registered physiotherapists with experience treating people who have proximal hamstring tendinopathy ii) hold a Master's degree or Doctor of Philosophy and iii) have a minimum of 10 years' experience.

3.4.3 Data collection

Experts were approached via email by a single investigator (AN). Interviews were performed in English online using Zoom © software or via telephone. A male interviewer (AN) performed, recorded and transcribed all interviews. The interviewer was a PhD candidate, and an Australian registered physiotherapist of 10 years, with a Master's degree in sports physiotherapy. Expert physiotherapists were emailed the list of interview questions prior to the interview. All interviews were recorded and transcribed verbatim and returned to participants for comment and/or correction. Two of the experts interviewed were authors on the paper, but were not involved in coding or analysing results, and had no conflicts of interest. Interviews were conducted until the same constructs repeated themselves
and no additional new themes emerged. No repeat interviews were performed.

3.4.4 Data analysis

Qualitative content analysis was used to analyse data. Audio files were transcribed and read multiple times to gain a sense of recurrent themes. Two independent researchers (AN & TP) analysed data systematically for meaning and then condensed, coded and organised these into categories and sub-categories according to the study aims. The computer software NVivo 12 (QSR International Pty Ltd version 12.5.0) was used to assist with the organisation of data. Categories and sub-categories were then compared for similarities and differences.

3.5 Results

After the 13 interviews there were 14 main themes, which were organised into 40 categories and 119 subcategories.

3.5.1 Expert demographics

Fourteen experts were contacted and agreed to participate. One expert subsequently declined due to illness. The average clinical experience of the 13 experts (8 males) was 25 years (range 11-42). Experts worked in Australia (9), Ireland (1), New Zealand (1), Qatar (1) and Scotland (1) with experience across multiple sports at an elite level. The mean interview time was 44 minutes (range 31-78). All experts had experience working with the general population and professional level sport. Sports included netball, Australian Football, soccer, ballet, race walking, touch football, weightlifting, rugby and running.

3.5.2 Diagnosis

All respondents used a combination of findings from the patient interview and clinical examination to come to the diagnosis of proximal hamstring tendinopathy (Figure 3.1).

3.5.3 Patient interview findings

All experts reported that the onset of proximal hamstring tendinopathy was insidious and associated with an increase in mechanical load through the proximal hamstring tendon (Table 3.1). All experts reported that patients described pain at the proximal hamstring insertion at the ischial tuberosity and most experts agreed that the pain did not shift or spread. *"It has to be around the proximal hamstring tendon"* (Expert 5). A shift or spread in pain was often expressed to indicate a differential diagnosis or co-morbidity. Four experts (31%) reported that the symptoms primarily occurred at the hamstring insertion, although pain did at times spread down the hamstring, but not past the knee.

Typically, patients described a spike in energy-storage-release (fast tensile) loads, particularly in combination with hip flexion (e.g. increased volume of running up hills). Provocative activities included activities such as lunging, running up hills and hamstring stretching, or activities that placed compressive loads through the proximal hamstring tendon unit, such as sitting.

3.5.4 Clinical tests

3.5.4.1 Loading tests

All experts performed a battery of clinical tests to form a diagnosis (Table 3.1). Most respondents used progressive load-based response tests, where loads placed on the hamstring

tendon were progressed from low to high through a sequence of successive tests. Tests were considered positive if localised pain at the ischial tuberosity was reproduced and increased with tasks that placed greater loads on the hamstring. Experts commonly used loads involving forward trunk inclination in standing with extended knee(s) (e.g. single leg Romanian deadlift, arabesque, trunk on leg flexion). Other, isolated pain-provocation tests commonly used were the single leg bridge test in 90 degrees of hip flexion, isometric knee flexion at 90 degrees of hip flexion and knee extension, bent knee stretch test and modified bent knee stretch test.

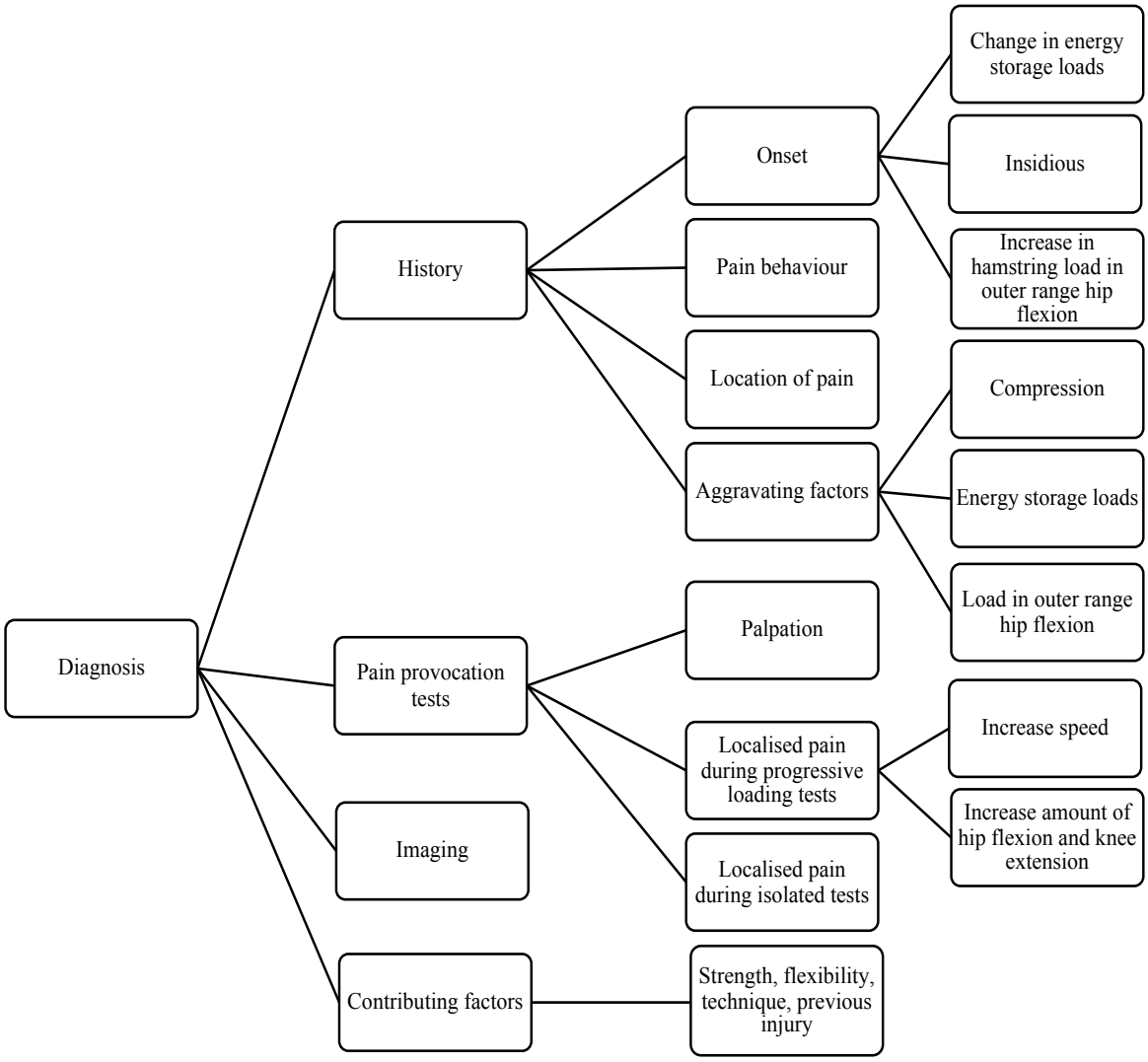


Figure 3.1 The diagnostic process used by experts in diagnosis and identification of contributing factors

Table 3.1 Key points on diagnoses of proximal hamstring tendinopathy

Patient interview	
Key points:	Related quotes in response to questions regarding the patient interview
• Insidious onset	• *"The mechanism is the big thing. Particularly in rugby, you know the ones you want to scan straight off, whereas the history for these types of things are going to be the load related factor. Most of the time being a lingering thing that has got a bit worse. Insidious like."* **Expert 6**
• Spike in training load	• *"History leading up to it. So, in the younger running group, or the active group, there is often some sort of history of change in load. So, it might have been coming up to an event and they've increased mileage or very commonly in our runners it is doing hill running, and often back-to-back hill running, and often without adequate rest between."* **Expert 12**
• No shift in direct area of pain at ischial tuberosity with loading	• *But the thing I'm looking for there is a really well localised area. That is the area that they get any time they stretch it or do their provocative exercise. And critically, the area doesn't change."* **Expert 13**
• Aggravated by direct compression (e.g. sitting) • Aggravated by activities that involve hip/trunk flexion and contraction of the hamstring unit	• *"... I'm looking for aggravating factors that I would associate with being high tendon load. Often compression is an aggravating factor and particularly those combined things for tendons."* **Expert 2** • *"These are worse I think in more hip flexion and worse with direct pressure. So, that is why if you are sitting with your hip in flexion, you're getting that compressive load on the underside of the tendon."* **Expert 7**
Clinical examination	
Key points:	Related quotes in response to questions regarding the clinical examination:
• A battery of load-based tests should be used to diagnose PHT • Screen for contributing factors (e.g. impairments) o Requires individual assessment o Relationship to previous injuries	• *"...we'd start with slower movements and increase speed and then in terms of the proximal hamstring, we would increase speed with compression. So, going into a body on leg movement, an arabesque and we're looking for localised pain with those low load tests where as you increase the speed and therefore the load for the tendon the pain remains localised".* **Expert 2** • *"So, looking for localised pain that stays localised to load, and a load dependent increase in pain. So, reproducing that (the athlete's pain) with a double leg body on leg flexion, then a single leg body on leg flexion, double leg fast body on leg flexion, then single leg fast body on leg flexion"* **Expert 11**
• Findings on palpation and imaging should be interpreted with caution	• *"Palpation is a funny one. Sometimes I think patients have hamstring tendinopathy, but are not necessarily sore on palpation. But I definitely use it in the bag of things you use to come to clinical diagnosis."* **Expert 3** • *"Probably one of the other big things is palpation, so palpation at the tendon insertion, and that reproduces their pain and localises the area of their pain."* **Expert 4**
• Hamstring, triceps surae and gluteus maximus strength often reduced • Hamstring muscle atrophy	• *"I think observation is really important. Trying to pick up if there is any hamstring wasting or not. If there is no hamstring wasting um, I tend to prick up my ears a bit more. Because people who have had long standing ones, apart from the middle-aged women, people will tend to have inhibition of hamstring if they've been long standing."* **Expert 13**
Imaging	
Key points:	Related quotes in response to questions regarding the utility of imaging:

• Imaging is not required to diagnose PHT, but is used to aid differential diagnosis	• *"I don't think you have to have imaging for the diagnosis. I'm quite comfortable making a clinical diagnosis because I've seen a lot of these."* **Expert 1** • *"I don't. I don't, so my thinking would be if they are not progressing in 12 weeks, then I would probably send them for imaging at that point."* **Expert 5**
• Signs of PHT on imaging are common in the asymptomatic population	• *"On occasions yes, I would use imaging. With the difficult ones probably, a combination of MR and ultrasound... It does give you some indication if there are some perineural issues. And it does, you know, define the pathology site. But the thing is that, like all tendinopathies we can have changes, this doesn't necessarily describe their pain."* **Expert 13** • *"I don't send for imaging, if they have imaging, I think where it is most helpful is when you have negative imaging."* **Expert 2**
Common differential diagnosis	
Key points:	Related quotes in response to questions regarding differential diagnosis:
• Screen for other musculoskeletal structures that may refer to the region • Sciatic nerve pathology is a common differential diagnosis	• *"Peri-neural irritation of the sciatic nerve. You often get a different response to load (e.g. you won't get a load dependent increase in pain, so sometimes it might be affected by faster activities in an unloaded situation."* **Expert 11** • *"As you know the big differential diagnosis is sciatic neuritis and um, sometimes the only difference with a whole history and examination has been these subtle changes in the pain site - when you really break it down."* **Expert 13**
PHT = proximal hamstring tendinopathy	

3.5.4.2 Palpation

Some experts believed it was important to be able to reproduce the patient's symptoms on palpation, whereas other respondents thought palpation was either of no use or had limited diagnostic value (Table 3.1). Several experts believed that failing to palpate the entire footprint of the hamstring insertion, in particular the attachment of the semimembranosus on the lateral aspect of the ischium, was why clinicians have difficulty reproducing the patient's symptoms on palpation.

3.5.4.3 Contributing factors

As part of the physical examination, all experts screened patients for contributing factors they believed relevant to the development of the condition. Examples included load capacity tests of hip and knee movements (e.g. hip extension strength and knee flexion strength), calf endurance capacity tests, and range of motion of the hip, knee, ankle and first metatarsophalangeal joint.

There was a consistent theme that analysing performance "on-field" was vital. *"There are times where we really need to get a better idea of what has actually created this within their playing environment and we can do things, often to address this"* (Expert 13). Commonly observed features of gait in runners with proximal hamstring tendinopathy were overstriding, low cadence, sitting low (crouching type gait) and excessive anterior pelvic tilt.

3.5.4.4 Imaging

Experts rarely used imaging to diagnose proximal hamstring tendinopathy, preferring to use information gained in the patient interview and physical examination. Experts usually referred for imaging when they believed a different condition was masquerading as a tendinopathy, or when the condition was unresponsive to management. The absence of tendinopathic changes on imaging (MRI or ultrasound) was suggested to be useful in helping rule the condition out.

3.5.4.5 Common differential diagnosis

Experts highlighted that a primary differential diagnosis was pain originating from the sciatic nerve or nerve sheath (Table 3.1). This pathology was reported to occur either concurrently with proximal hamstring tendinopathy or as a separate entity. A theme was that a more widespread distribution of symptoms through the buttock and down the posterior thigh was a

common feature of sciatic nerve involvement, with the pain location slightly more lateral from the ischial tuberosity compared to proximal hamstring tendinopathy. Three experts highlighted that symptoms (sciatic nerve) could also be localised to the ischial region without peripheral spread. Tests used to diagnose sciatic nerve involvement included the slump test, performed in both lumbar flexion and extension, and the straight leg raise.

Several experts mentioned that undiagnosed systemic inflammatory conditions were at times present in patients referred for treatment of proximal hamstring tendinopathy. Examples of systemic drivers here included ankylosing spondylosis or psoriatic arthritis.

3.5.5 Management

The primary management options utilised were education and exercise (Table 3.2). Passive interventions were included by some experts, but only as an adjunct to education and exercise.

3.5.5.1 Education

Patient education covered a variety of different elements (Table 3.2). The delivery was adapted to individual goals and specific limitations. *"...I think empowering people and giving them like really positive sort of self-efficacy and not only reassuring them that they'll get better, but giving them the tools to get there"* (Expert 2). Tendinopathy specific pain education with the key message being that pain does not always mean harm and pain 24 hours post activity could be used to judge how well the tendon had tolerated an activity. All respondents agreed that the condition required significant rehabilitation time, *"often a good 3-6 months in most people"* (Expert 9), others mentioned 6-12 months.

3.5.5.2 Exercise

3.5.5.2.1 Targeted hamstring rehabilitation

All respondents prescribed targeted exercises to load the hamstring musculotendon unit from early to late-stage management (Figure 3.2). Exercises were progressed from low-load exercises in positions of minimal hip flexion (e.g. isometric long lever bridge or supine plank) to high-load exercises, depending on factors, such as individual pain response to the exercise. Exercises were advanced as early as tolerated, which was determined primarily by pain response to load (e.g. 24-hours post). Experts reported athletes could often tolerate heavy load, performed slowly, early on (e.g. single leg prone hamstring curl, long lever hip bridge). Key

Table 3.2 Key points on management of proximal hamstring tendinopathy

Education	
Key points:	Related quotes in response to a question regarding messages in education:
• Discuss the persistent nature of tendinopathy (not a quick fix) • Provide education on low importance of tendon changes on imaging	• *"We'll have to have a chat about what tendinopathy is - telling them that we don't really understand where the pain is coming from, but these are our best guesses, this is how it normally behaves and most importantly this is what we think we can do about it…So the way I often wheel that discussion around is to start off talking about is I'd say well if we had a weak muscle what do you think we'd do about it…well muscles are different, you can strengthen them by doing exercise, you know these muscles attach to tendons, and their turn over is slower, but how do you think your body can make them stronger, and then we can have a chat about different types of exercise."* **Expert 7**
• Deliver self-monitoring strategies (e.g. 24 hour-rule) • Educate that loading/strengthening is the key treatment	• *"One of the things I talk about is load tolerance, so you've got to develop load tolerance, independently, and the only way to do that is loading…And then you'd give them some sort of framework. So ok, letting them know they can load it if it's 5/10 that ok, keep going as long as you recover well after 24 hours, if it is more than that, or if the recovery is longer then you've got to reduce what you're doing. So, giving them a framework, and that comes with the belief that pain is not equal to damage."* **Expert 5**
• Minimise compressive loads and energy storage and release loads • Employ methods to reduce sitting pain • Modify high tendon loads and compressive loads	• *"At the gym - finding alternatives to those exercises or positions that are creating high amounts of compression, or combinations of compression and tension of the hamstring tendons. Sitting - if that is a big problem it is something that we need to address… Also, we advise them to try to share between sitting and standing, if they can get a sit/stand desk set up. And then teaching them just basic functional load sharing things, such as during bending and squatting, making sure they are using their knees when they are bending."* **Expert 12**
Early management	
Key points:	Related quotes in response to questions regarding early management:
• Avoid exercises involving moderate hip flexion motion range	• *"Starting with a slow or static movement that is not provocative for a tendon out of compression. So, examples would be a long leg bridge whether in pelvic hip neutral a prone hamstring curl for either isometric or isotonic. In terms of the isotonic what I'm looking for as well is time under tension, so we would look at a full three seconds concentric/eccentric."* **Expert 2**
• Provide isolated loads to the hamstring muscular-tendon unit • Athletes can often tolerate heavy loads early in management if applied in an isometric or as slow isotonic fashion	• *"I like starting them on hamstring curls machine, or a cable machine. Something where we can weight them up pretty heavy. Probably set them up on a plan of somewhere around the range of 3 times per week of doing some kind of sets towards 4-6 reps. So, depends where the athlete is in terms of and pain it might take them a little more time to get them there. You might take them through 10 reps. 8 reps, 6 reps then even 4 reps. So really heavy and really slow. 4 sets of 6 would be ideal and keep that as maintenance."* **Expert 6**
• Start movement technique re-training targeting i) control of lateral pelvic posture ii) control of anterior pelvic tilt posture iii) excessive trunk flexion iv) excessive hip flexion	• *"So do a lot of work on controlling around the hip, in particular avoiding excessive anterior pelvic tilt, and single leg stance around the hip, with the thought patterns being that if you are struggling to maintain your pelvic position in single leg stance then you might be at risk of falling into anterior pelvic tilt rather than controlling that lumbopelvic position, which will increase the work load on the hamstring and hamstring*

	tendon while increasing compression." **Expert 10**
• Use manual therapy as an adjunct to target physical impairments (e.g. reduce hip extension, hamstring muscle flexibility)	• *"Ah, no not really, unless, no. If it was sciatic nerve then obviously, freeing up the path of the nerve, so the glutes and all of that. Hamstring tendon, no. I don't think there is anything that will really help besides load."* **Expert 11** • *"Yes, in the early stages I use a fair bit of soft tissue massage through the hamstrings. You can do other techniques like dry needling depending on patient preference. This would be just in the muscle belly, not the tendon. Restoring adequate hip mobility is also important. I think performing techniques to increase and restore adequate hip extension is important, particularly if they are hyper-lordotic."* **Expert 8** • "*With shockwave therapy we know that it can be effective, but there is uncertainty whether it is different to placebo. So, I tell patients that this is an intervention that does help some people's pain, but the effectiveness has not been conclusively proven. Would you like to try it? …. Some people get a good response. Aside from the loading really, there is not much else I'd do as well as the shockwave therapy"* **Expert 5**
Mid-to-late-stage management	
Key points:	Related quotes in response to progressing past early management:
• Hip flexion range of motion starts in very limited motion range, and very slow, and is gradually progressed as tolerated • Continue isolated hamstring loading using slow heavy loading • Progress to faster movements that require energy storage and release loads through the tendon once a strong strength base is built	• *In response to a question regarding progressing into more hip flexion ROM - "split squats where you can go into progressive amounts of hip flexion, and different amounts of compression. I really like stairs, so I can get people going up and down stairs and then as you get them going up 2 and 3 stairs at a time and really driving. So really combining that tension and compression.*" **Expert 2** • *"So that would be things like increasing range of hip thrusters, they could maybe do some sideways sumo steps with band around their legs and carrying weights as well. So, barbells. Getting into more and more flexion. I'll progress from sideways to doing forward lunges to maybe only 40 odd degrees and then gradually increase the range from 40 at the hip and knee to 50/60/70/80/90. All progressed by how much discomfort they are getting at the time and how much pain they are getting 24 hours later…. then low sled push, where I start really erect and then add more and more and more hip flexion. So, it is only concentric, but we are starting to get really good loads through the whole kinetic chain. And that starts to give confidence to them into range. There are some that really struggle with introducing hip flexion, so I might tend to go soft on that, and even get to stage 3 where I'm starting to do faster stuff, but more upright. Then just add the hip flexion as I can get it, otherwise you can be waiting forever."* **Expert 13**
• Biomechanics - movement technique re-training	• *"The mechanics would be things like over-striding, anterior pelvic tilt and landing with a very straight knee in running, which goes with over-stride. That basically would be the main ones and you might also get lots of trunk flexion, more so in sports. They'd be the main ones."* **Expert 5**
Return to sport	
Key points:	Related quotes in response to assessing for readiness for return to sport:
• Assess hamstring strength and symptom response to loading tasks	• *Return to running – "I like to look at hand held dynamometry. I like to see that they can do 80-90% of what the other leg is doing. We do 90/45 and 0. I also look at glute bridges off a box. Knee 20 degrees of flexion. And then performing a single leg. Getting to 20 with good lumbopelvic control throughout."* **Expert 8** • *"I would also like them to be pain free on an isometric muscle tests, and I like the supine plank exercise as a*

	bit of a general screen for what their load tolerance and load capacity is like, so I like them to be able to lift one leg off, lift the other leg off, without pain and feeling like they've got full control of that. I've usually had them doing a walking program on the flat, so I like them to be pain free walking for at least 30 minutes." **Expert 12**
• Most athletes return to sport with symptoms • Individuals in team sports can often continue to compete while recovering	• *"If they are an elite sportsman trying to keep them in the sport, but making adjustments in other aspects of their life, to keep them playing – that seems to work better than if I have to pull them out. I think the hardest ones seem to be the ones where you really think that they're current sporting loads are compromising them, and that is where I've found it more difficult, where you actually have to stop them playing their sport for a period of time, that is a little unpredictable because it varies between people, so they are the ones that I've found more difficult, so I think getting patient by-in, really good education from the start and then having a really clear progressive program of functional steps, but also strengthening and loading steps so they're not booming or busting through the course of management."* **Expert 3** • *"Return to load, return to their sport with a manageable pain condition. So, I don't think it is about zero pain. I think it is about managing it. So, if you have someone who can run, they can continue to run if they are managing that pain the next day.* **Expert 11**
• Strength and conditioning exercises that required larger ranges of hip flexion, such as deadlift, leg press are only reintroduced if necessary once full return to sport is reached and condition is stable	• *I like to see equal glute strength to hamstring strength. So, I think you might know – we use load cells to look at strength over the hip and over the knee. So, I like to see pretty equal strength there. So, no pain on any of the drop catches – including internal rotation. No pain on forward lunges or arabesque. And then getting them to do, if they are a sprinter/footballer, that they've done either their starts on their change of direction, inside ball sort of work and come up pain free.* **Expert 13**
Monitoring	
Key points:	Related quotes in response to monitoring:
• VISA-H is rarely used • Self-monitoring using NRS of load-based tests (e.g. arabesque, single leg bridge) • Change in sitting pain 24/hours post sport/activity • Isometric strength – knee flexion dynamometry • TAMPA scale, psychological readiness for sport	• *"So, yeah, we should use the VISA-H but we don't. We use our clinical experience and outcome measures that are more patient based."* **Expert 11** • *"So, we are looking at the level of discomfort after games. This is the main focus. Then the duration of those symptoms. So, if we play on a Saturday and then generally we get home by the Monday and they've still got relatively low levels of symptoms, say 3-4/10 on single leg bridge, ongoing pain with sitting and that generally feeling of tightness when they are walking, we start to worry that that is starting to linger in terms of symptoms."* **Expert 10**
Preventing recurrence of PHT	
Key points:	Related quotes in response to preventing recurrence:
• Running technique	• *"…if they have a history of other conditions, so they if they have a history of ankle sprains, or say if they'd had an ACL reconstruction then you'd need to work out what part of the kinetic chain is vulnerable to losing strength and making sure they remain strong."* **Expert 11**
• Strength training for the kinetic chain (gluteus maximus and triceps surae)	• *"Running, so how they run. Changing their technique so they land under their centre of mass. That also might be an important consideration."* **Expert 4**

• Hamstring strength	• *"I would say people have to have sufficient strength and kinetic chain strength, and then have tendon loading that's really consistent and that's your best way of preventing it."* **Expert 2**
• Education regarding load management	• *"Education about changes to load, about… some of these people are trying to prepare for a race and they get through the race, then they have time off and then they don't keep up any management strategies."* **Expert 3** • *If they are not appropriately spreading those higher load days, I think they are much more likely to get recurrence…And then of course just activity – so what type of activity they are doing. There are certain sports or activities that will just be higher load or have higher compressive load and higher combinations of compressive and tensile load for the individual"* **Expert 12** • *"I think improving them with specific retraining to the tasks they want to do with progressive overload you'll get them to the point where they can take that load. I think that is how you prevent reoccurrence."* **Expert 6**

AN: Anthony Nasser (interviewer), EP: Expert physiotherapist, NRS: numerical rating scale, PHT: proximal hamstring tendinopathy, VISA-H: Victorian Institute of Sport – proximal hamstring tendon,

characteristics of exercise selection, shared by experts, included that initial exercises were in near-neutral hip flexion and were performed unilaterally as early as possible.

3.5.5.2.2 Kinetic chain rehabilitation

Increasing the capacity of the entire kinetic chain to improve load distribution was a recurring theme. Other areas targeted in rehabilitation were based on individual assessment, and were often specific to goals and deficits. Deficits targeted were often related to past injuries (e.g. previous ankle sprain and triceps surae wasting).

3.5.5.2.3 Aberrant biomechanics

Movement retraining was also recommended if aspects of the performance of a task were seen to place increased load on the hamstring unit. While not exhaustive, this typically targeted i) control of frontal plane femoro-pelvic position ii) control of sagittal plane pelvic position (i.e. anterior/posterior tilt) iii) excessive forward trunk inclination iv) excessive hip flexion.

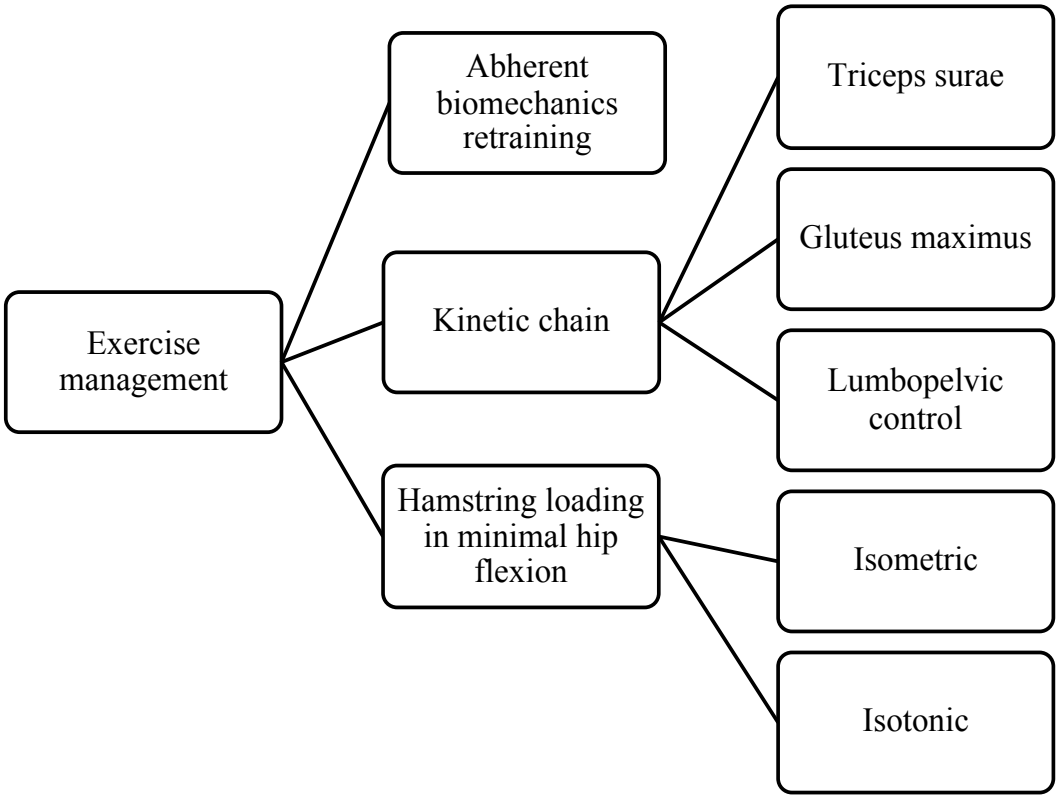


Figure 3.2 Considerations in exercise selection during early-stage management

3.5.5.3 Mid to late-stage exercise management

Rehabilitation was progressed by i) increasing load ii) increasing speed of contraction iii) hamstring exercises in increasing ranges of hip flexion and iv) increasing complexity (e.g. dynamic sports-specific drills). Experts progressed athletes into more hip flexion, in controlled environments, in a graded manner. Examples of exercises included step-ups, split squats, stairs and slow sled push.

Athletes were then progressed to faster movements that required energy storage and release loads through the tendon. Once athletes had successfully transitioned and were tolerating their sporting requirements in match play, strength and conditioning exercises that required larger ranges of hip flexion, such as deeper deadlifts, were reintroduced if deemed necessary.

3.5.5.4 Passive interventions

Passive interventions, such as manual therapy and injection therapies were not considered integral by any expert physiotherapist. Most experts used massage therapy, as an adjunct, in the early stages, as they felt it would assist in settling the tendon down when it was in a reactive state. Manual therapy was also used as an adjunct to target associated physical impairments (e.g. soft tissue massage to address hamstring muscle flexibility). One expert mentioned they sometimes utilise shockwave therapy alongside a loading program. Other expert physiotherapists didn't use any passive management strategies. More invasive management, including injection therapies and surgery were not recommended. No expert physiotherapist referred patients on for platelet-rich plasma injections or corticosteroid injections.

3.5.5.5 Return to sport

Athletes were often able to continue to compete in sport while recovering through adjustment of training loads and incorporation of targeted strength and conditioning interventions. *"It is pretty rare that we stop someone all together"* (Expert 9). A major theme was that due to the nature of competitive sport, return was often rushed and rehabilitation incomplete. A notable exception to those views were for track and field athletes and amateur runners who had deconditioned significantly and did not have a history of strength and conditioning training – these often-required time away from sport to recover.

3.5.5.6 Monitoring

Most experts were aware of the Victorian Institute of Sport Assessment - Proximal Hamstring Tendon (VISA-H), a proximal hamstring tendinopathy-specific questionnaire, but rarely used it. They found the questionnaire not very responsive to change. Question-based monitoring of specific patient issues and provocative loading tests (Table 3.2) were used to monitor rehabilitation.

3.5.5.7 Prognosis

Several experts stated that sitting pain, a common feature of proximal hamstring tendinopathy, often took over a year to resolve and lagged behind return to function. Patients with concurrent pathology, comorbidities and athletes in mid-season were reported to be more difficult to manage with delayed recovery.

3.5.5.8 Preventing recurrence

The rationale for ongoing management was reiterated due to the high potential for recurrence. In particular, the importance of strength in the hamstring and the kinetic chain was echoed across respondents, as well as addressing areas that were vulnerable to atrophy, such as deficits associated with past injuries. *"I would say people have to have sufficient strength and kinetic chain strength, and then have tendon loading that's really consistent and that's your best way of prevention"* (Expert 2). *"So, getting a program that targets lumbar spine, lower glute, hamstring strength, adductors, so all of those exercises you just need to make sure that they continue them forever more, basically"* (Expert 9). Re-testing key objective measures, such as strength with hand-held dynamometry or in gym-based exercises, following breaks from sport, was seen as important since a spike in workload upon sport resumption could cause a recurrence.

3.6 Discussion

We aimed to explore and summarise expert physiotherapists' clinical reasoning around the assessment, management and prevention of proximal hamstring tendinopathy. The diagnosis was typically made by combining information gained in the history and confirmed with multiple pain provocation tests. Proximal hamstring tendinopathy was primarily managed through patient education and progressive load-based rehabilitation, targeting the hamstring unit and kinetic chain. This involved avoiding activities in positions of end-range hip flexion

early in rehabilitation and a graded return to sport. Prevention measures involved maintenance of hamstring and kinetic chain strengthening programs and control of physical workload.

3.6.1 Diagnosis

Whilst there is consensus that tendinopathy is a clinical diagnosis (159), there is no consensus regarding which diagnostic tests should be used when diagnosing proximal hamstring tendinopathy. This is supported by a systematic review demonstrating inconsistencies in participant selection criteria (155). Three evidence-based pain provocation tests – Puranen-Orava, bent-knee stretch, modified bent-knee stretch, have moderate to high sensitivity (0.76 to 0.89) and specificity (0.82 to 0.91) in detecting MRI-defined tendinopathy in symptomatic participants compared to healthy controls (61). These tests were only used by a small percentage (23%) of physiotherapists interviewed. Most used a number of other provocation tests, that have yet to be validated – isometric knee flexion in supine with 90 degrees of hip flexion and the arabesque.

Expert physiotherapists suggested clinicians must be wary when interpreting imaging findings due to the disparity between tendon changes on imaging and symptoms. The evidence suggests that there is a high prevalence of MRI-defined structural changes in the proximal tendon in the asymptomatic population – e.g. 65% (84) and 90% (mean age not reported) (85) of those imaged. This theme aligned with a recent International Consensus, reporting that imaging is not necessary when diagnosing tendinopathy (159).

The relationship between proximal hamstring tendinopathy and the sciatic nerve, highlighted in this study, has long been acknowledged. Surgeons frequently perform debridement of the nerve in conjunction with treatment of the affected proximal tendon (47, 57, 65). The anatomical proximity has been considered a reason for the relationship between these two conditions (62), with theories including swelling of the hamstring tendon causing direct compression on the sciatic nerve, and the sciatic nerve becoming impinged during activities due to adhesions forming between the two structures (47, 58, 87). Due to the uncertainty expressed around the diagnosis of sciatic nerve pathology, future research into diagnosis is required.

3.6.2 Management

Experts consistently reported that rehabilitation should include education, load management and progressive exercise. There is little research on management. A multi-modal intervention that included 3-weeks of exercise (exercise, manual therapy, NSAIDs, ultrasound) was less effective in improving symptoms and physical function than shockwave therapy (112). This rehabilitation program differed significantly from what expert physiotherapists told us. Major deviations were the consideration of compressive load on the enthesis when selecting exercises and the need for 3-6 months of rehabilitation. Whilst there is a lack of research on load-based interventions in proximal hamstring tendinopathy (155), such interventions have demonstrated promising results in other tendinopathies of the lower limb (38, 107, 153). Future rehabilitation research should consider examining the effectiveness of load management and exercise – attending to specific parameters – that experts currently use in practice.

3.6.3 Return to sport

Experts suggested athletes can often continue competing with proximal hamstring tendinopathy. Supporting evidence, in Achilles tendinopathy, suggests that pain within acceptable levels (maximum pain score of 5/10 on a 0-10 scale) during activity does not adversely affect outcomes (160). Proximal hamstring tendinopathy may affect performance to a point where a period of recovery and progressively rebuilding load tolerance may be indicated. Gradual rebuild with progression related to increasing load, speed of contraction, range of motion and complexity was endorsed in our results.

3.6.4 Preventing recurrence

There is little evidence on secondary prevention in tendinopathy. The use of corticosteroid injection (CSI), which was recommended against in our study, has been shown to delay recovery and lead to higher recurrence rates than control (138). Whilst the long-term effect of CSI use in proximal hamstring tendinopathy is unknown, two retrospective studies showed >50% of patients did not have long-term symptom resolution (9, 135). In our study, expert physiotherapists advised careful load management in activity and sport, along with long-term continuation of hamstring and kinetic chain strengthening.

3.6.5 Limitations

As we only included expert physiotherapists, our results may be biased towards non-invasive interventions. Most of our experts were from Australia (69%) and while there were no major differences in opinions between experts, this may affect external validity.

3.7 Conclusion

Expert physiotherapists diagnose proximal hamstring tendinopathy using a combination of findings from the patient interview and pain provocation tests – implying no single test is adequate. There was consensus that progressively loading the tendon, to check for pain response, was useful in diagnosis, whereas views on the value of palpation differed. Education to improve patient understanding of pain and tendon load, to allow self-monitoring and progression, combined with a progressive rehabilitation program, were cornerstones of management and the prevention of recurrence. Passive management strategies were perceived to be of little benefit.

3.8 Summary of findings

A qualitative design was used to understand the current best physiotherapy practice in the assessment, management and prevention of proximal hamstring tendinopathy (161). Proximal hamstring tendinopathy was diagnosed through findings in the patient interview and physical examination (161). Education to encourage self-monitoring and progression, combined with progressive rehabilitation, were foundations of both management and prevention (161). Passive interventions were thought to have minimal benefit.

The diagnosis of tendinopathy was achieved through clinical assessment (2). Whilst key basic elements of diagnosis achieved consensus (e.g. combination of history and clinical tests) many particulars (e.g. specific test used) were inconsistent (161). Evidence accumulated in Chapter two demonstrated inconsistencies in patient eligibility in studies (inclusion criteria) (155). A similar issue of a lack of consistency in eligibility criteria was reported in a recent editorial on Achilles tendinopathy (162). The emphasis of the editorial can be gleaned from the title – "Diagnosing Achilles tendinopathy is like delicious spaghetti carbonara: it is all about key ingredients, but not all chefs use the same recipe" (162). The International Tendinopathy Consensus (ICON) mandatory reporting guidelines have set a standard for studies on tendinopathy (80). Clearer descriptions of inclusion criteria and patient characteristics aid the interpretation of results. This is particularly important in the case of proximal hamstring tendinopathy where there is yet to be consensus on diagnostic or inclusion criteria.

The importance of considering compressive loads in the development of rehabilitation programs and activities of daily living was a recurring theme in management. Unloading of activities that place high amounts of compressive load has been promoted in the rehabilitation of insertional tendinopathy (32). No comparisons of rehabilitation programs with and without patients reducing compression management in proximal hamstring tendinopathy exist. Results from an RCT in 2011, showed that a multi-modal program that involved patients exercising in positions of compression (e.g. deadlift, lunge, hamstring stretch) had a poor effect on improving physical function and symptoms in the short and long-term compared to shockwave therapy (112).

The VISA-H, the only condition-specific outcome measure for proximal hamstring tendinopathy, was not considered valuable or practical by experts (161). The lack of take up

of the VISA-H as a monitoring tool, suggests that the validation of practical questionnaires, such as single-item questions or questionnaires with limited items, should be explored. It is also critical that the measurement properties of VISA-H (e.g. content validity and responsiveness) are scrutinised. Experts in the qualitative study instead used simple question-based monitoring of specific patient issues (e.g. patient-specific functional scale) or load monitoring tests (e.g. NRS pain with an arabesque each morning) (161). These were used as they were thought to be more meaningful to patients, and of more use in monitoring symptoms due to their responsiveness to symptoms.

4 Chapter 4: Coming to consensus – progressing scientific research on tendinopathy

4.1 Preface

The aim of Chapter four is to summarise findings from the International Scientific Tendinopathy Consensus (ICON) statements published in 2019 and 2020 which provide context for the remaining chapters of the thesis.

4.2 Coming to a consensus

A group of researchers with expertise in tendinopathy conducted a consensus process between 2018 and 2019 that aimed to address inconsistencies in research and increase research transparency and quality that is halting progress in this field (163). Factors impeding progress included inconsistencies in clinical terminology, an absence of any agreed-upon tendon health-related domains and incomplete/poor reporting of patient characteristics/inclusion criteria in studies (2, 80, 164).

4.2.1 Clinical terminology

The evolution in the terminology used to describe tendon pain was outlined in Chapter One. Inconsistency in terminology for painful tendon disorders is a problem in many regions throughout the body. The ICON consensus aimed to address this by providing clear and uniform terminology for clinicians to use (2). Consistent terminology is critical for patients as they may feel they are receiving multiple different diagnoses - for example, if they are diagnosed with proximal hamstring tendinopathy and tendinitis by two different professionals (2). This may reduce confidence in the health professional's advice and reduce compliance to management strategies, such as an exercise-based loading program (2).

The consensus group agreed tendinopathy is the preferred term for persistent tendon pain and that imaging was unnecessary for the diagnosis of tendinopathy (2). Lateral elbow tendinopathy, medial elbow tendinopathy, patella tendinopathy, peroneal (fibularis) tendinopathy and Achilles tendinopathy were agreed upon as terms for their respective tendon disorders (2). Nomenclature at other tendon sites, such as proximal hamstring, tibialis posterior, gluteal tendon, or rotator cuff was not addressed.

4.2.2 Health-related domains

Core health-related tendinopathy domains were agreed upon as part of the ICON consensus process. This was to ensure the impact of tendinopathy on patients was reflected in research findings. Domains were initially considered via an online survey and agreed upon during a meeting at the 2018 International Scientific Tendinopathy Symposium (ISTS, Groningen, the Netherlands) (164).

The domains that were decided upon were: patient overall rating, participation, pain on activity/loading, disability, function, physical function capacity, quality of life, psychology, and pain over a specified timeframe (164). Domains that were not considered to be necessarily reported were physical activity, structure, medication use, adverse events/effects, economic impact(costs), pain elicited with clinician-applied stress/examination, clinical examination findings, palpation, range of motion, dropout or discontinue treatment, sensory modality-specific pain and pain without further specification (164).

4.2.3 Outcome measures

Now that there is a consensus on which domains should be measured, outcome measures need to be agreed upon for individual tendinopathies that measure each of the aforementioned core health domains (e.g. gluteal tendinopathy, proximal hamstring tendinopathy and Achilles tendinopathy). These outcome measures would then become candidates to be a part of a core outcome set. Core outcome sets are the minimum sets of measures that should be tested when conducting research. A core outcome set can be developed through the methods proposed in the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN), Core Outcome Measures in Effectiveness Trials (COMET) and Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT) initiatives (165-167). Core outcome sets are important to ensure homogeneity in the selection and reporting of results in research, which facilitates the pooling of results in meta-analyses (168).

5 Chapter 5: Core outcome set development for proximal hamstring tendinopathy (COS-PHT): a study protocol

5.1 Preface

Chapter five builds on findings from the ICON consensus outlined in Chapter four, which highlighted the need for the development of core outcome sets to evaluate the impact of tendinopathy. This chapter describes the protocol used to develop a Core Outcome Set for Proximal Hamstring Tendinopathy (COS-PHT). The completion of a core outcome set would recognise a collection of measures that could be used in research of proximal hamstring tendinopathy. This will ensure the impact of proximal hamstring tendinopathy is captured in research.

The following chapter contains a modified version of the paper

Core outcome set development for proximal hamstring tendinopathy (COS-PHT): a study protocol

Nasser AM, Semciw AI, Grimaldi A, Rio E, Pizzari T, Vicenzino B,

Physical Therapy Reviews

P1-7, 2022

URL link to published article: https://doi.org/10.1080/10833196.2022.2077066

5.2 Abstract

Background

Proximal hamstring tendinopathy is a cause of buttock pain that is common in running-based sports. A recent systematic review found outcome measures used in research studies on this condition were inconsistent, which impedes the synthesis of evidence. To understand the impact tendinopathy has on patients an international collaboration (International Scientific Tendinopathy Symposium Consensus Group) agreed on nine core health domains that should be measured in research on tendinopathy. The aim of this study is to develop a core outcome set for proximal hamstring tendinopathy that covers all nine tendinopathy domains. This study describes the steps taken to develop this core outcome set.

Methods

This mixed method study will follow two key phases. Phase one: a systematic review will produce a comprehensive list of outcome measures that have been used in studies evaluating proximal hamstring tendinopathy. Outcome measures extracted will be matched to the tendinopathy outcome domains by a steering committee of clinicians and researchers. Phase two: Following a Delphi process, outcome measures will be considered by experienced clinicians, researchers and patients. Within this study outcome measures will be screened and selected using the Outcome Measures in Rheumatology filters of truth, feasibility and discrimination, with a threshold of 70% agreement set for consensus.

Discussion

This project aims to provide and disseminate a core outcome set that can be recommended for use in future studies related to proximal hamstring tendinopathy. This will minimise research waste, allow pooling of studies in meta-analysis and assist in directing future research.

Trial registration

Protocol registered with the Core Outcome Measures in Effectiveness Trials in May 2021 (http://www.comet-initiative.org).

Key Points

- There is no agreed set of outcomes that are consistently used when researching proximal hamstring tendinopathy

- This protocol outlines the methods to derive a set of outcome measures to use when researching proximal hamstring tendinopathy
- Researchers, clinicians and patients will participate in a process of selecting outcome measures against the nine core domains for tendinopathy
- Where domains do not have an outcome measure the participants will provide recommendations for future research.
- Consistent use of outcome measures will allow pooling of data in meta-analysis and minimise research waste

5.3 Background and objectives

Proximal hamstring tendinopathy is a common cause of buttock pain that occurs in running-based sports, such as distance running and Australian football (161). The condition also afflicts people who do not participate in sport, in particular, post-menopausal women (46, 161). Symptoms are often persistent and slow to respond to treatment (155, 161).

A recent systematic review of the efficacy of interventions used in managing proximal hamstring tendinopathy identified a range of different outcome measures used in individual studies (155). This made it difficult to generate meaningful conclusions of treatment efficacy and thus clinical recommendations. A remedy for this would be to agree to a set of outcomes that will be reported in all trials of proximal hamstring tendinopathy. Meaningful conclusions will lead to relevant clinical recommendations.

To understand the impact tendinopathy has on patients – the International Scientific Tendinopathy Symposium Consensus (ICON) collaboration – agreed on core health-related domains that should be measured in research in tendinopathy (164). These include: patient rating of condition, participation in life activities, pain on activity/loading, function, psychological factors, physical function capacity, disability, quality of life and pain over a specified time (164). It was recommended that these domains serve as a basis for the development of a core set of outcome measures that validly and feasibly measure the impact of the tendinopathy on a patient (164). The objective of this study is to develop and publish a core outcome set for use in future clinical trials on any intervention in adults with proximal hamstring tendinopathy (COS-PHT).

5.4 Methods

This mixed method study follows the Outcome Measures in Rheumatology guidelines (167). The study will follow two key phases. In phase one, a systematic review will identify a comprehensive list of outcome measures that have been used in previous studies evaluating proximal hamstring tendinopathy. Outcome measures extracted will be matched to the established core tendinopathy by a steering committee (164). In phase two, outcome measures matched to core health domains will be considered by experienced clinicians, researchers and patients. The methodology for this study is summarised in Figure 5.1.

This protocol follows guidance from the Core Outcome Set Standardised Protocol checklist (164)and handbook (166). The project was prospectively registered with the Core Outcome Measures in Effectiveness Trials Initiative: https://cometinitiative.org/Studies/Details/1876 in May 2021. Definitions used throughout the consensus process can be seen in Table 5.1.

5.4.1 Steering committee

The steering committee will consist of six members and includes clinical investigators (researchers) and health care professionals with expertise in proximal hamstring tendinopathy. The steering committee's role is to identify all outcome measures used in research on proximal hamstring tendinopathy, preliminarily map the outcomes against core health domains as defined by the ICON consensus process (164), facilitate the development and completion of surveys, patient interviews, consensus meeting and analyse and disseminate the results.

5.4.2 Recruitment of participants

5.4.2.1 Working group

Professionals in the working group will consist of experienced clinicians and clinical investigators with expertise in proximal hamstring tendinopathy. Authors who have published on proximal hamstring tendinopathy in the last ten years will be invited to join the working group via email from a single author (AN). Organisations that are likely to see a high volume of patients with proximal hamstring tendinopathy will also be approached via email (e.g. sport associations). Clinical expertise will be recognised through either the completion of relevant post-graduate training (e.g. Master's degree) or a minimum of ten years of clinical experience. Professionals will be recruited with the aim of achieving diversity of profession (e.g. physiotherapists, orthopaedic surgeons etc.), experience across different sports, geographical location and sex. This will assist in capturing diverse viewpoints.

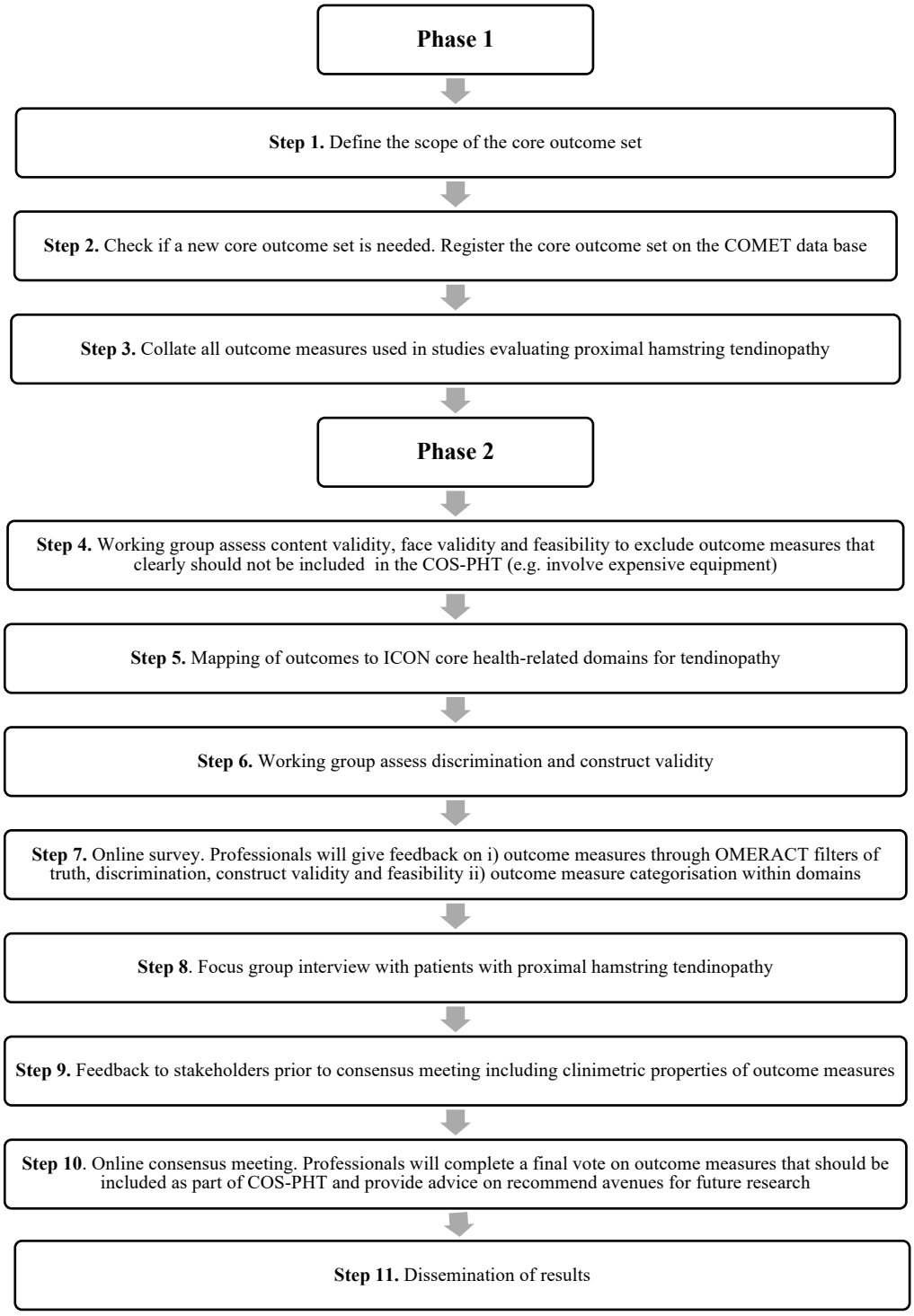


Figure 5.1 Steps to develop the core outcome set for proximal hamstring tendinopathy COMET = core outcome measures in effective trials, COS-PHT = core out set - proximal hamstring tendinopathy, International Scientific Tendinopathy Symposium Consensus Group (ICON), Outcome Measures in Rheumatology (OMERACT)

Table 5.1 Truth, Discrimination and Feasibility

Truth	Does the outcome measure used to assess proximal hamstring tendinopathy measure what it intends to measure? Is the result of the outcome measure unbiased and relevant? Truth captures the face, content, construct and criterion validity.
Discrimination	Does the outcome measure distinguish between situations that are of interest? Discrimination captures reliability and sensitivity to change.
Feasibility	Can the measure be applied easily to patients with proximal hamstring tendinopathy, given constraints of time, money, and interpretability? Feasibility addresses the practicalities of the measure

This table is adapted from Tugwell et al. 2007 (doi: 10.1186/1745-6215-8-38).

We aim to recruit a minimum of 15 professionals. Participation in all stages of the project is voluntary, with members of the working group able to withdraw from the study at any time. The number of experts included in Delphi studies varies between 15 and 60 (169). There will be no maximum number of professionals. Fluency in written and spoken English will be required to participate in surveys and focus groups.

5.4.2.2 Patients

Patients older than 18 years of age will be recruited through clinical practices who have been diagnosed with proximal hamstring tendinopathy by a health care professional in Australia. Participants must have had the condition for a minimum of three months. Participation will be voluntary. Patients will need to be fluent in written and spoken English to participate in a focus group. A target number of five patients will be recruited.

5.4.3 Phase one

5.4.3.1 Identification of available outcome measures for a COS-PHT

A search of the literature will be conducted to identify published clinical trials or systematic reviews that report outcome measures used to evaluate proximal hamstring tendinopathy. Outcomes that had been measured in previous trials will be summarised to inform a set of standard outcomes for future research studies. Participants in included studies must be diagnosed by a healthcare professional with proximal hamstring tendinopathy, with or without radiological confirmation, and have undergone any type of intervention (e.g. rehabilitation, injection therapy, surgery). All outcome measures will be included. Only studies in the English language will be considered for inclusion. Studies on acute partial tears, or full-thickness ruptures will be excluded. Outcome measures will be extracted by two independent

reviewers into a pre-formulated excel spreadsheet. An independent third reviewer will provide consensus in the case of discrepancies.

5.4.4 Phase two

5.4.4.1 Initial assessment of truth and feasibility

Members of the steering committee will conduct an initial assessment of content validity, face validity and feasibility of each outcome measure extracted in phase one. The purpose of this phase is to exclude outcome measures that clearly should not be included in a core outcome set (e.g. specialised equipment not widely available for all clinical trials are not feasible such as expensive biomechanical motion analysis equipment). For each filter, a traffic-light rating will be applied. A green light means the instrument will be included at this stage, an amber light means it is included at this stage with caution, and a red light means it should be excluded. All six members of the committee will vote on all outcome measures, and differences in opinion will be discussed until consensus is reached using the online program Qualtrics^{R} (Provo, Utah, USA).

5.4.4.2 Mapping of outcomes to core health-related domains for tendinopathy

Outcome measures will be mapped to one, or more, of the nine core health domains independently by the six members of the working group using the online program Qualtrics^{R} (Provo, Utah, USA) (164). Any differences in opinion will be discussed until a consensus is reached. If the consensus is that an outcome measure fits into more than one domain, the outcome measure(s) will appear in both core health domains when presented to the professionals in the first stage of the Delphi process. If there is consensus amongst the steering committee that an outcome measure(s) does not fit into any domain, the outcome measure will be presented to the professionals who will have an opportunity to recommend mapping an outcome to a core health domain. This will also ensure transparency of outcome measures excluded from the process on these grounds.

5.4.4.3 Online survey of professionals

A survey using the software Qualtrics^{R} (Provo, Utah, USA) will be constructed and piloted for usability (including reducing complexity) and technical functionality. The survey will be developed and reported in accordance with the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) (170). An invitation to complete the survey will be sent prior to the

release of the survey and will include information on the objectives, study design, expectations, and consensus definitions. A link to access the survey online will be sent to the professional's email addresses.

Demographic information of respondents will be collected in the survey (e.g. age, professional discipline, sex). Participants who contribute will be authors of a future manuscript if their contributions to the final publication of the core outcome set meet the requirements of the International Committee of Medical Journal Editors. To prevent multiple entries from the same individual, the data of participants will be checked by the steering committee (unique name and email address).

We anticipate the online survey will take approximately 30-45 minutes to complete, with the survey to be completed within a timeframe of four weeks. A reminder to complete the survey will be sent to participants who have not completed the survey one-two weeks after the survey has been released. The survey will include a copy of each outcome measure and a summary of how often the measure has been used in the literature, including citations, to assist professionals to make an informed decision.

In the survey, respondents will be asked to consider the truth and feasibility (Table 5.2) of each outcome measure within each core health domain(s). All outcome measures for each core health domain will be presented in random order within each domain – respondents will be made aware of this. For each question, respondents will be given three possible responses (agree, disagree or unsure). Following each domain, we will ask respondents for any comments about the outcome measures we present and others they may suggest (e.g. ones that were missed in the literature review or ones in development and not yet published). If we are not able to find any outcome measures for a domain, we will provide the example outcome measures that were provided in the original ICON domains publication (164).

The participants' response to each question will be calculated as a percentage. Consensus definitions and the criteria for outcome measure reduction, progression or addition for the survey are detailed in Table 5.2. In brief, we have decided in advance that at least 70% of participants will need to agree or disagree for an outcome measure to be included or excluded, respectively – if there are no less than 15% of respondents voting for the opposite. To ensure transparency around missing data, the number of professionals addressing each question will

be reported, as will any dropouts at different stages of this consensus process. A second Delphi round will be utilised if required to assess the truth and feasibility of any new outcome measures identified in the survey.

Table 5.2 Consensus definitions, criteria for item reduction, progression or addition

Criteria for inclusion	Consensus for Inclusion in the COS-PHT: ≥70% of participants selected 'agree' responses and <15% selected 'disagree'.
Criteria for exclusion	Consensus for Exclusion from the COS-PHT: ≥70% of participants selected 'disagree' responses and <15% selected 'agree' responses
Undecided	Undecided: any outcome instruments that have reached neither the criteria for inclusion nor exclusion
Criteria for item reduction	Any items for which consensus for exclusion was reached will not be taken forward for further consideration.
Criteria for item progression	Any items for which consensus for exclusion was not reached, will be taken forward for Quality assessment.
Addition of Items	From the suggestions within text boxes within the survey following each domain, any currently available outcome measures that match these domains will be added for consideration in subsequent steps. Outcomes measures not currently available, or in development, will be added to the discussion at the consensus meeting, if the study working group agrees the suggestion represents a potentially truthful and feasible measurement.

5.4.4.4 Focus group interview

To ensure the core outcome set reflects the impact of proximal hamstring on patients, a focus group will be conducted with patients who have proximal hamstring tendinopathy. Two members of the steering committee will conduct the focus group – with at least one who has experience with focus groups. We aim to recruit a minimum of seven participants, based on the COSMIN recommendations (171). The purpose of the focus group is to gain patient insight on whether an outcome measure is feasible, accurately captures relevant aspects of the condition, is comprehensive, and comprehensible. If patients do not have a say in the development, the COS may overlook important outcomes, and ultimately research on proximal hamstring tendinopathy may fail to provide useful information about treatments (167). To ensure patients are not overly burdened the focus of the meeting will be on outcome measures that have achieved consensus (e.g. ≥70% of participant's select 'agree' responses). If time permits and patients are agreeable, the other outcome measures will be considered. The focus group will be performed online in English using Zoom © software. The meeting will be recorded and transcribed to allow researchers to review and analyse data. Qualitative content analysis will be used to analyse data. The responses will be read multiple times to gain a sense

of themes by two independent reviewers, who will analyse data for meaning and then organise data into categories and sub-categories. Any discrepancies in interpretation will be discussed and if they are unable to be resolved a third reviewer will provide consensus. Demographic information of the patients will also be collected (e.g. age, sex, duration of symptoms). Patients will be sent the final report from the focus group to ensure that their views are appropriately represented.

5.4.4.5 Assessment of measurement properties

We will conduct a search for studies or systematic reviews that evaluate the psychometric properties of outcome measures used in proximal hamstring tendinopathy. The measurement properties that will be reviewed include: content validity, structural validity, construct validity, reliability (Cronbach's α), repeatability (test-retest reliability), responsiveness and interpretability. Measurement properties will be rated as either sufficient (+), insufficient (-) or indeterminate (?) and a rating of the overall quality of evidence as high, moderate, low or very low using the Consensus-based Standards for the Selection of health status Measurement Instruments (COSMIN) guidelines (172). Ratings of each outcome measure will be provided to participants prior to the final consensus meeting.

5.4.4.6 Feedback to stakeholders prior to the consensus meeting

Prior to the final consensus meeting results of the survey and patient interviews will be prepared and provided to professionals in the working group. This report will be provided a minimum of two weeks prior to the final online consensus meeting. Responses to all open-ended questions in the survey will be summarised and included in this feedback. This will allow consensus meeting participants to consider ideas of all members of the group to assist them to make an informed decision.

5.4.4.7 Online consensus meeting

The consensus meeting will be attended by professionals and held online in English using Zoom © software. There will be an overview of the process to date and of the results from the survey and patient interviews. Attendees will then discuss and vote on any of the items that did not reach the criteria for either inclusion or exclusion in the Delphi round 1 survey. Voting will be completed using an online polling function on ZOOM ©. The criteria for inclusion in

a COS-PHT are summarised in Table 5.3. Professionals will also discuss gaps in evidence for any core health domains where a suitable outcome measure has not been identified.

Table 5.3 Consensus meeting criteria

Consensus definitions	Consensus for inclusion: ≥70% of participants selected agree responses Consensus for exclusion: ≥70% of participants selected disagree responses. Undecided: any outcome instruments that have reached neither the criteria for inclusion or exclusion

Dropouts from the initial survey to the consensus meeting may result in missing data. The number of attendees at the consensus meeting will be reported, as well as the number of attendees who vote on each item.

5.4.4.8 Publication of COS-PHT consensus and recommendations

If the consensus process identifies at least one suitable outcome measure in each of the core health domains, a COS-PHT will be published and disseminated for use in studies of people with proximal hamstring tendinopathy (164). If we are unable to recommend any outcome measures, we will publish recommendations for further development and validation of new or existing instruments. These recommended outcome measures will serve as priorities for research.

5.5 Discussion

This project ultimately aims to disseminate a core outcome set that can be recommended for use in future studies on proximal hamstring tendinopathy. This will improve research quality, allow pooling of studies in meta-analyses, and assist in directing future research. This study seeks to involve a range of stakeholders to encourage broad acceptance and implementation of the outcome measures. The involvement of patients will assist the core outcome set to reflect the impact of proximal hamstring tendinopathy.

Inconsistencies in outcomes used within studies on the same condition causes problems for users of health care research (173). For example, a review of trials on oncology found that more than 25,000 outcomes appeared two times or less (168). The development of a core outcome set addresses the problem of lack of standardisation of outcome measures across

studies, which hampers evidence synthesis (173). In addition to this, outcome measure quality varies considerably (e.g. reliability), and it is often not clear whether the most reliable and valid measure is selected within a study (174). The development of a core outcome set aims to provide researchers with a set of robust outcome measures which may improve research quality. The process provides an opportunity to ensure the quality of an outcome measure is adequate prior to recommendation (166).

For an outcome measure to be recommended it must have sufficient measurement properties. The COSMIN guidelines set rigorous methods for outcome measure evaluation and recommendation based on measurement properties (172). In this study, there is a possibility that there will be insufficient evidence on the measurement properties of an outcome measure to represent one or more of the core health domains (164). This will likely affect the working groups confidence in making recommendations.

The success of a core outcome set is reliant on the uptake in research and clinical practice. A recent systematic review explored the application of core outcome sets in randomised controlled trials and found uptake varied considerably across different areas of health research (173). For example, full core outcome set uptake was 0% in randomised controlled trials on gout, compared to 82% in trials on rheumatoid arthritis (173). Due to such variability, it is critical that steps are undertaken to consider barriers to uptake when developing a core outcome set. Barriers to uptake suggested are limited patient and key stakeholder involvement, as their involvement improves the relevance of the core outcome set (13). This reiterates the importance of including a patient focus group in the consensus process. Other barriers include lack of awareness, difficulties with implementation, and lack of resources (16). Another barrier suggested was the practicality of the outcome set, with a survey for paediatric acute and chronic pain finding that six domains were considered too many (175), as it is possible that the perceived burden on patients to complete an exhaustive list of outcome measures may cause reduce implementation by researchers (173). Although no relationship has been found between the number of domains and uptake (173), as the ICON included nine core health domains (164), the practicality of completing outcome measures will be an important consideration to minimise the burden of completing the core outcome set.

This core outcome development will be conducted in the English language only – a limitation. For practicality reasons we decided to limit the search to one language as there are other matters that need to be considered when translating across languages in a valid and reliable manner. Whilst efforts will be made to invite a broad and representative selection of contributors in the selection of the study working group and patients, the authors recognise that certain groups may be under-represented, as there is the possibility that certain sub-groups may hold more influence during parts of the Delphi meeting, such as the patient focus group or consensus meeting. To minimise this risk, the steering group will maintain oversight of the process and use an anonymous voting system in the consensus process.

5.6 Conclusion

Variability exists in the outcome measures used when researching proximal hamstring tendinopathy. Outcome measures used in clinical practice and research must be reliable, valid and responsive to change and cover the range of core health domains acknowledged in the ICON consensus process. This protocol describes the approach to initiate the development of a Core Outcome Set – Proximal Hamstring Tendinopathy (COS-PHT) with the aim to publish a patient-centered selection of outcome measures.

5.7 Summary of findings

Due to the issues caused by the variability of outcome measures when researching proximal hamstring tendinopathy, a problem highlighted in Chapter two, it is critical outcome measures selected are consistent between studies. This protocol described the approach to initiate the development of a Core Outcome Set – Proximal Hamstring Tendinopathy (COS-PHT) with the aim to publish a patient-centred selection of measures.

A rigorous methodology was described that involved the recruitment of international experts from a variety of health professions. Thresholds for consensus were set to guide the selection process. The result aims to provide a patient-focused, user-friendly outcome set with the goal of broad implementation in future research.

6 Chapter 6: Outcome measures in the management of proximal hamstring tendinopathy; a systematic review of their measurement properties

6.1 Preface

Chapter five (study three) outlined the methodology to develop a core outcome set for proximal hamstring tendinopathy (COS-PHT) (176). Part of this process involves collating outcome measures used in research on proximal hamstring tendinopathy and then classifying them into the nine established tendinopathy core health domains: patient overall rating, participation, pain on activity/loading, disability, function, physical function capacity, quality of life, psychology, and pain over a specified timeframe (164). Following this step, information on the quality of outcome measures is required to report back to the core outcome set development working group. To achieve this a systematic review of the literature was performed, following COSMIN guidelines, to identify the measurement properties of outcome measures used in research on proximal hamstring tendinopathy.

The following chapter contains the following publication in its entirety

Outcome measures in the management of proximal hamstring tendinopathy; a systematic review of their measurement properties

Mr Anthony Nasser, Prof Bill Vicenzino, Dr Alison Grimaldi, Dr Ebonie Rio, Mr Aidan Rich, Dr Tania Pizzari, Dr Adam Semciw

Submitted to Physical Therapy in Sports (Under-review)

6.2 Abstract

Objective

To summarise outcome measures in proximal hamstring tendinopathy research, map measures to domains, and evaluate their measurement properties.

Methods

MEDLINE, CINAHL, EMBASE, SPORTSDISCUS and PUBMED were searched to identify outcome measures used in proximal hamstring tendinopathy (February 2022). Measures were mapped to the International Tendinopathy Scientific Consensus core health domains. A second search (MEDLINE, CINAHL, EMBASE and PubMed) identified (February 2022) studies evaluating measurement properties of outcome measures captured in the initial search. Included studies evaluated measurement properties of outcome measures in participants with proximal hamstring tendinopathy (diagnosed by a healthcare professional). Consensus-based-Standards for the Selection of Health Instruments methodology were followed—including risk of bias assessment and synthesis of findings.

Results

Four studies (n=302) evaluated the measurement properties of the Victorian Institute of Sport Assessment – Proximal Hamstring Tendinopathy (VISA-H). For the VISA-H there was moderate-quality evidence of sufficient construct validity, low-quality evidence of sufficient responsiveness, reliability and measurement error, very low-quality evidence of sufficient relevance and comprehensibility and very low-quality evidence of insufficient comprehensiveness.

Conclusion

No outcome measure demonstrated sufficient measurement properties to be recommended for use. The VISA-H is currently better than other measures of disability - based on few studies and should be used with caution.

6.3 Introduction

Proximal hamstring tendinopathy presents as pain on the hamstring insertion at the ischium (161). The condition is associated with long-standing pain and disability and is consistently identified in athletes participating in running-based activities (81, 161). The condition is also common in active post-menopausal women (46, 161). Research on the prevalence, risk factors, diagnosis and management of proximal hamstring tendinopathy is limited (155, 161). Suitable outcome measures are required to evaluate interventions and identify change over time.

Consistency in the measurement of health constructs and outcome measures in studies on treatment efficacy is critical for the comparison of interventions in systematic reviews, from which clinical guidelines are often formed (e.g. pooling of studies in meta-analysis) (174). Outcome measures that have not been validated for use in patients with proximal hamstring tendinopathy, may impact results, reducing our confidence to inform practice. To capture the broad impact of tendinopathy, outcome measures should reflect the core health domains (i.e. health-related aspects that may be impacted by the condition) established by the International Scientific Tendinopathy Symposium Consensus (ICON) (164) in 2019 (patient overall rating, participation, pain on activity/loading, disability, function, physical function capacity, quality of life, psychology, and pain over a specified timeframe) (164).

This systematic review is part of a project to produce a core outcome set for use in clinical trials of patients with proximal hamstring tendinopathy. The aim of this study was to examine the types of outcome measures used in studies on proximal hamstring tendinopathy, categorise measures into the core health domains and evaluate their clinimetric properties following the COSMIN (Consensus-based Standards for the Selection of Health Measurement Instruments) guidelines (165).

6.4 Methods

6.4.1 Study design and registration

This study consisted of three parts: (part one) a systematic review to identify all outcome measures used in research on proximal hamstring tendinopathy, (part two) mapping of outcome measures to ICON core health domains and (part three) completion of a second systematic review to evaluate the measurement properties of the identified outcome measures of proximal hamstring tendinopathy. We listed the proposal to develop a core outcome set for proximal

hamstring tendinopathy on Core Outcome Measures in Effectiveness Trials in 2021(https://comet-initiative.org/Studies/Details/1876) and published the protocol (176). The systematic review protocols were developed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (177) statement and preregistered on PROSPERO (ID: CRD42021237507).

Part One

6.4.2 Search strategy

A search was conducted to find outcome measures used in studies of treatment efficacy in proximal hamstring tendinopathy. A search was updated from a systematic review published in 2021 on proximal hamstring tendinopathy (155). MEDLINE, CINAHL, EMBASE, SPORTSDISCUS and PUBMED were searched in February 2022, with the assistance of a librarian, using a combination of keywords and medical subject headings (MeSH). The search strategy was formed around the concepts “hamstring” and “tendinopathy”.

6.4.3 Study selection

Titles and abstracts of articles identified from all databases were uploaded to an online systematic review screening tool (COVIDENCE). After the removal of duplicates, two independent reviewers (AN and BEW) screened titles and abstracts against the inclusion criteria. A third independent reviewer (AS) was available in the case of discrepancies. Full texts of the remaining articles were examined for inclusion by two independent reviewers (AN and BEW). We included studies published on any intervention that reported on patients with proximal hamstring tendinopathy (e.g. randomised clinical trials, cohort studies) and case series (with ten or more participants). Participants of any age diagnosed with proximal hamstring tendinopathy by a healthcare professional were included. Traumatic injuries such as acute proximal hamstring tendon tears, complete hamstring tears or avulsion injuries were excluded.

6.4.4 Data extraction

Outcome measures used in eligible studies were extracted into a pre-formulated Microsoft Excel ® spreadsheet (version 16.16.27) by two independent reviewers (AN, BEW). A third independent reviewer (AS) was available in the case of discrepancies.

Part Two

6.4.5 Suggested domain mapping

Outcome measures extracted from part one were independently mapped by five of the authors (AN, AS, ER, AG, BV) to the core health domains of tendinopathy using the online survey software Qualtrics. Any disagreements were discussed by the authors until a consensus was reached (164).

Part Three

6.4.6 Search strategy

Based on the results of parts one and two, a search strategy was developed to identify studies that reported on the measurement properties of outcome measures in patients with proximal hamstring tendinopathy. MEDLINE, CINAHL, EMBASE and PUBMED were searched (February 2022) using search filters recommended by COSMIN (Appendix G).

6.4.7 Study selection

Titles and abstracts of articles identified from all databases were uploaded to an online systematic review screening tool (COVIDENCE). After the removal of duplicates, two independent reviewers (AN and AR) screened titles and abstracts against the inclusion criteria. Any disagreement was resolved by an independent third reviewer (AS). Full texts of the remaining articles were examined for inclusion by two independent reviewers (AN and AR). Cases of disagreement were referred to an independent third reviewer (AS) for consensus.

To be included studies must have i) included patients with the diagnosis of proximal hamstring tendinopathy made by a health professional and ii) included an assessment of one or more measurement properties of outcome measures identified in the systematic review in part one that could be mapped to one or more of the nine core health domains of tendinopathy (part two). Studies reporting on measurement properties of outcome measures that could not be mapped to one or more of the nine core health domains were excluded. Measurement properties included: content validity, reliability, construct validity, structural validity, internal consistency, measurement error, cross-cultural validity, responsiveness, interpretability and feasibility. Animal studies, narrative reviews, case reports, protocol papers, systematic reviews and studies with less than ten participants were excluded.

6.4.9 Data extraction

Two reviewers (AN and AS) independently extracted information on the characteristics of the study populations and all data on measurement properties into preformulated Microsoft Excel ® spreadsheet (Version 16.16.27). Any discrepancies in extraction were resolved by a third independent reviewer. Data were extracted on study characteristics (e.g. author, year, country of administration, outcome measure evaluated) and all information pertaining to the following measurement properties: content validity, reliability, responsiveness, cross-cultural validity, construct validity, internal consistency, measurement error, interpretability and feasibility.

6.4.10 Methodological quality of included studies

The methodological quality for each measurement property in each study was assessed by two independent reviewers (AN and AR) using the COSMIN Risk of Bias Checklist (178). A third independent reviewer (AS) was invited in the case of disagreement. Each criterion on the checklist for each measurement property received a rating of either: very good, adequate, doubtful or inadequate (172). As per the guidelines, the worst score counts approach was used (e.g. if one of the criteria was rated inadequate the overall rating for that measurement property was rated inadequate regardless of what other items for that measurement property were rated) (172, 178).

6.4.11 Data analysis

Extracted results on content validity and the content of the outcome measure itself were rated by two independent reviewers (AN and AR) against the Ten Established Criteria for Good Content Validity (179). Each of the criteria was rated: positive (+), negative (-), or indeterminate (?). The content of the outcome measure itself was also scored by two independent reviewers (AN and AR), with a third independent reviewer (AS) providing consensus in the case of disagreements (reviewer ratings) (179).

All other measurement properties were rated by two independent reviewers against the updated criteria for good measurement properties (e.g. measurement error, responsiveness) (180). A third independent reviewer was invited in the case of disagreement (172). Each criterion could be scored: positive (+), negative (-), or indeterminate (?) (178). The smallest detectable change was calculated using the formula provided by Terwee (SDC = 1.96 x $\sqrt{2}$ x SEM) (180). When possible, results from different studies on a single measurement property were pooled using

the software R (A language and environment for statistical computing. Version 4.5.0. Vienna, Austria; 2019; http://www.r-project.org/). An intraclass correlation coefficient (ICC) was used for test-retest reliability analysis. Values were transformed to z-values using Fisher's method and then back-transformed for presentation in forest plots (181). Heterogeneity in effect sizes was estimated using (I^2), with values interpreted as follows: 0% meant no inconsistency, 25% low inconsistency, 75% high inconsistency and 100% total inconsistency (181). Results were then summarised to obtain an overall score and rated as sufficient (+), insufficient (−), inconsistent (±), or indeterminate (?) (172). The overall rating was scored indeterminate if there was not enough information available. The hypothesis was developed *a priori* using Pearson's correlation coefficient (r) that strong correlations ($r \geq 0.50$) would be found with instruments measuring similar constructs, correlations with instruments measuring related, but dissimilar constructs would be lower ($r = 0.3\text{-}0.5$) and correlations with instruments measuring unrelated constructs would be low ($r<0.3$).

6.4.12 Quality of the body of evidence

The overall level of evidence was summarised for each measurement property using a modified Grading of Recommendations, Assessment, Development and Evaluations approach (172). This rating reflects the confidence in the available evidence (e.g. ''high'', ''moderate'', ''low'', or ''very low''). All measurement properties started at 'high-level evidence' were rated down by up to three levels (e.g. to very low) for risk of bias and up to two levels for inconsistency, imprecision or indirectness (172, 178).

The COSMIN guidelines were followed to direct overall recommendations on the use of outcome measures. An outcome measure was recommended for use if it had any level of evidence showing sufficient content validity and at least low-quality evidence for sufficient internal consistency (165). An outcome measure was not recommended for use if there was high-quality evidence for an insufficient measurement property the outcome measure (165). If an outcome measure did not fit into either statement the outcome measure with the best evidence for content validity was provisionally recommended (165).

6.4.13 Formulating recommendations

If an outcome had sufficient content validity (any level of evidence) AND at least low-quality evidence for sufficient internal consistency it could be recommended for use (165). If there

was high-quality evidence for an insufficient measurement property the outcome measure could not be recommended for use (165). If an outcome measure did not fit into either of the above statements the outcome with the best evidence for content validity should be provisionally recommended for use (165).

6.5 Results

Part one

6.5.1 PROMs used in proximal hamstring tendinopathy

A total of 2517 records were identified, of which 1150 were duplicates (Figure 6.1a). Following title and abstract screening, a total of 76 articles were retrieved, which resulted in 14 studies meeting inclusion criteria. Several studies were excluded due to study design (182, 183), sample size (184), or population (e.g. acute partial or full-thickness hamstring tear) (185, 186).

6.5.2 Outcome measures

Twenty-seven distinct outcome measures were used in studies on proximal hamstring tendinopathy (Appendix H). The outcome measure most reported was adverse events (12 studies), followed by return to sport (pre-injury level – six studies). Pain (visual analogue scale) was also commonly reported with most studies reporting this outcome at average follow-up time (e.g. 71.3 months) (47), rather than a pre-specified time point (e.g. 12 weeks) as recommended in the ICON consensus (164).

Part two

6.5.3 Suggested domain mapping

Participation and disability were the most reported core health domains (Appendix I). Psychology, physical function capacity and pain on activity/loading were not reported, and quality of life was reported in a single study. Various outcomes measures were unable to be mapped to any core health domains. Examples included adverse events, scores of patient satisfaction and pain at a mean follow-up time, with five studies not meeting requirements as they did not specify a reference period over which the patient was rating pain.

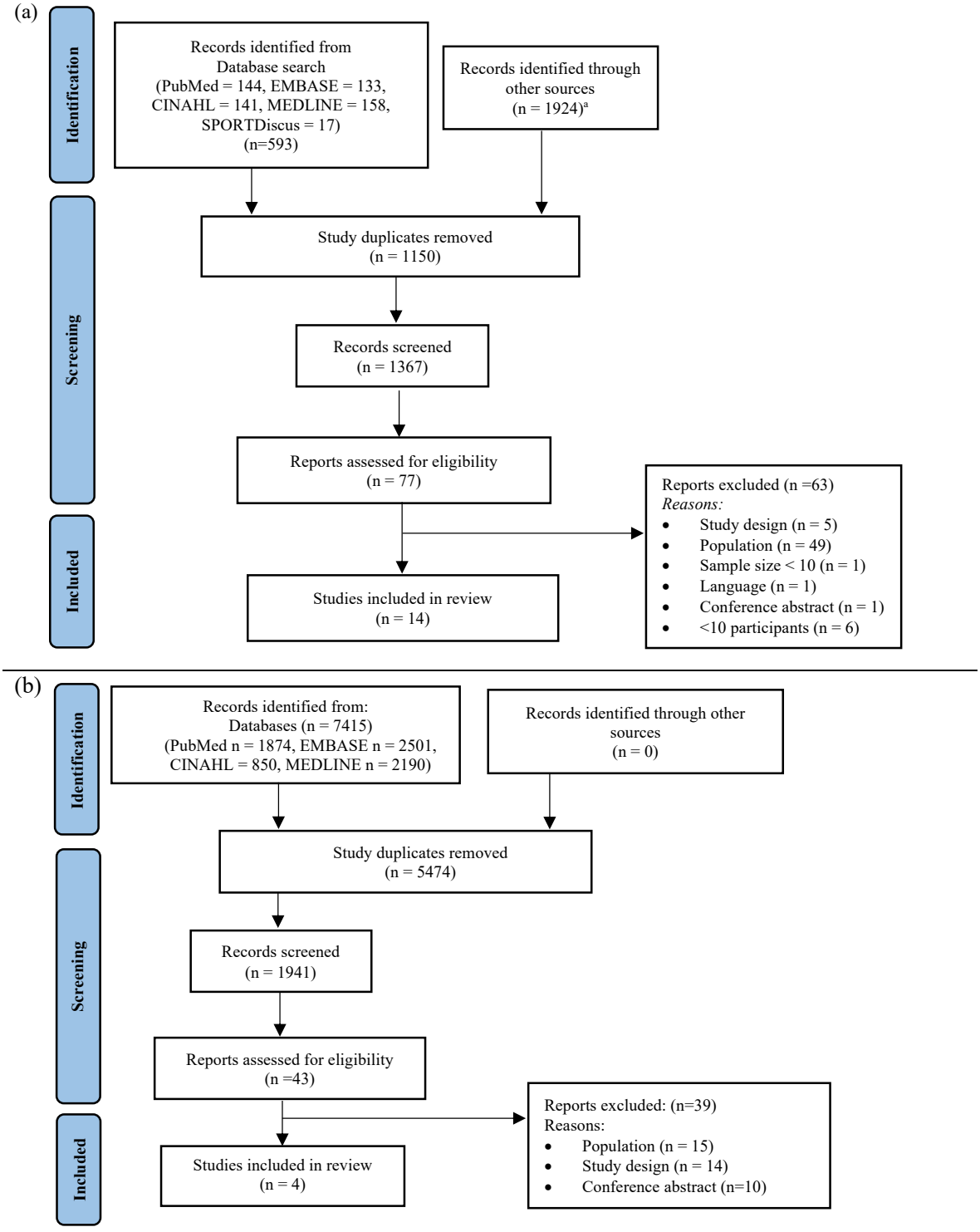


Figure 6.1 PRISMA flow diagram
Figure (a) shows the process taken to identify articles examining outcome measures in patients with proximal hamstring tendinopathy. Figure (b) shows the process taken to identify all articles evaluating clinimetric properties of outcome measures in patients with proximal hamstring tendinopathy

[a] Proximal hamstring tendinopathy; a systematic review of interventions (Nasser et al. 2021)

Part three

6.5.4 Study selection clinimetric properties

A total of 7415 records were retrieved of which 5474 were duplicates (Figure 6.1b). Following title and abstract screening, 1941 articles remained; four met eligibility criteria (60, 110, 187, 188). Several studies were excluded due to the population (189-192) (acute hamstring strain injury rather than proximal hamstring tendinopathy).

6.5.5 Study characteristics

Four studies reported on the measurement properties of the Victorian Institute of Sport Assessment – Hamstring (VISA-H) (60, 110, 187, 188). Measurement properties were not evaluated on any other measure in patients with proximal hamstring tendinopathy. Studies on the VISA-H included 302 participants, of which 137 had proximal hamstring tendinopathy (60, 110, 187, 188). Sixty participants took part in the original validation of the VISA-H (Table 6.1) (110). The mean age of participants with proximal hamstring tendinopathy varied from 21 to 40 years. The mean age of healthy controls ranged from 23 to 39 years.

6.5.6 Content validity

The VISA-H was developed in 2014 for patients with proximal hamstring tendinopathy (110) and has been translated into French (VISA-H.F) (187), Spanish (VISA-H.Sp) (188) and Brazilian Portuguese (VISA-H.Br) (60). The original development study did not involve a pilot test or cognitive interview to test the comprehensibility or comprehensiveness of the questionnaire (110). Consequently, the development study was rated inadequate(110). Three additional studies of doubtful quality reported on the comprehensibility of the VISA-H (60, 187, 188). The construct 'severity of symptoms' of proximal hamstring tendinopathy was considered when assessing the relevance and comprehensiveness of the VISA-H, rather than multiple concepts of pain, function and sporting activity (considering the VISA-H is calculated as a single score) (110). The ratings for comprehensiveness were based on reviewer ratings only, as no studies reported on these measurement properties. Overall, there was very low-quality evidence of sufficient comprehensibility and relevance, and insufficient comprehensiveness (Table 6.2).

Table 6.1 Characteristics of included studies

Source	Country of administration, Language version	Population	PROM	N (men%), PHT, control/other	Mean age (years), range	Inclusion criteria	Clinimetric properties assessed										
							1	2	3	4	5	6	7	8	9	10	11
Cacchio (2013)	Italy, English	PHT, asymptomatic	VISA-H	60 (70%), PHT non-surgical management: 20 (70%), PHT surgical management: 10 (80%), healthy: 30 (67%)	Non-surgical management: 23.7 years, range 18–25 Surgical management: 21.4 years, range 18–23 Controls: 23.1 years, range 18–26	> 18 years + clinical diagnosis of PHT + MRI. Clinical diagnosis: pain in the lower gluteal region, tenderness in the ischial tuberosity area and positive in at least two of the following three pain provocation tests: Puranen-Orava test, the bent-knee stretch test and the modified bent-knee stretch test	✓		✓	✓		✓	✓	✓	✓	✓	✓
Locquet (2019)	Belgium, French	PHT, asymptomatic	VISA-H.F	51 (64.7% male), PHT: 16, healthy: 35[1]	Total population: 32.4 ±12.0[2]	NR		✓		✓		✓		✓		✓	✓
De-la-Cruz-Torres (2021)	Spain, Spanish	PHT, asymptomatic	VISA-H.Sp	101 (58%), 51 PHT (63%), 50 (54%) healthy	PHT 39.8±9.3 Healthy controls: 38.8 ± 9.2	Clinical diagnosis of PHT with tendon changes verified by ultrasound, >18 years Clinical diagnosis: history of pain in the lower gluteal region for at least three months, tenderness in the ischial tuberosity area, positive in at least two of the Puranen-Orava test, bent-knee stretch test, and modified bent-knee stretch test		✓	✓	✓		✓	✓	✓	✓	✓	✓
Lima (2022)	Brazil, Portuguese	PHT, asymptomatic	VISA-H.Br	90 (47.8%), PHT: 40 (57.4%), 50 healthy (42.6%)	PHT 36.3±8.92 Healthy controls: 25.6±6.90	Clinical diagnosis: reporting presence of pain in the lower gluteal region and tenderness in area of the ischial tuberosity, and at least two positive results among the following four tests: Puranen- Orava, bent-knee stretch, modified bent-knee stretch and palpation		✓		✓		✓	✓	✓		✓	✓

NR = not reported, PROM = patient-reported outcome measure, PHT = proximal hamstring tendinopathy, VISA-H = Victorian Institute of Sport Assessment - Proximal hamstring tendinopathy, VISA-H.F = Victorian Institute of Sport Assessment - Proximal hamstring tendinopathy - French, VISA-H.Sp = Victorian Institute of Sport Assessment - Proximal hamstring tendinopathy – Spanish, VISA-H.Br = Victorian Institute of Sport Assessment - Proximal hamstring tendinopathy – Brazil

[1]20 practiced sports 'at risk' of developing PHT, 15 practiced a sport not at risk of developing PHT

[2] Mean age of groups not reported separately

1 = PROM development, 2 = Content validity, 3 = Structural validity, 4 = Internal consistency, 5 = Cross-cultural validity\measurement invariance, 6 = Reliability, 7 = Measurement error, 8 = Hypotheses testing for construct validity, 9 = Responsiveness, 10 = Interpretability, 11 = Feasibility

Table 6.2 Summary of findings

VISA-H	ROB	Inconsistency	Imprecision	Indirectness	Quality of Evidence (high, moderate, low, very low)	Overall Rating +/-/?
Content validity – relevance	Extremely serious	No	N/A	No	Very low	+
Content validity – comprehensiveness	Extremely serious	No	N/A	No	Very low	-
Content validity – comprehensibility	Extremely serious	No	N/A	No	Very low	+
Structural validity	No	Very serious	No	No	-	?
Internal consistency	Very serious	No	No	Serious[2]	-	?
Cross-cultural validity	NIA	NIA	NIA	NIA	-	?
Reliability	Serious	No	No	Serious[2]	Low	+
Measurement error	Serious	No	No	Serious[2]	Low	+
Construct validity (other outcome measures)	No	No	No	Serious[2]	Moderate	+
Construct validity (known group)	No	No	No	Serious[2]	Moderate	+
Responsiveness (comparison with other PROMs)	Serious	No	No	Serious[2]	Low	+
Responsiveness (before and after intervention)	No	No	Serious[1]	Serious[2]	Low	+

Overall rating:

- + = sufficient, - = indeterminate, ? = indeterminate

Quality level:

- High –very confident that the true measurement property lies close to that of the estimate of the measurement property
- Moderate –moderately confident in the measurement property estimate: the true measurement property is likely to be close to the estimate of the measurement property, but there is a possibility that it is substantially different
- Low – confidence in the measurement property estimate is limited: the true measurement property may be substantially different from the estimate of the measurement property
- Very low - very little confidence in the measurement property estimate: the true measurement property is likely to be substantially different from the estimate of the measurement property
- NIA = no information available, VISA-H = Victorian Institute of Sport – proximal hamstring tendinopathy

All studies started at high-quality and were down grounded for:

- Risk of bias: -1 = serious, -2 very serious, -3 = extremely serious
- Inconsistency: -1 serious, -2 very serious

- Imprecision: -1 total n = 50-100, -2 total n <50
- Indirectness: -1 serious, -2 very serious

[1] Imprecision downgraded due to sample size

[2] Indirectness downgraded due to pooling of multiple language versions of the VISA-H

6.5.7 Structural validity

Two studies of adequate quality assessed the structural validity of the VISA-H (110, 188) (Table 6.3). The VISA-H is calculated as a total score out of 100, and therefore is assumed to measure a single construct. One study performed principal component analysis and found two factors accounted for 73% of the total variance of the eight items (items one to six accounted for 34% and items seven and eight accounted for 39% of the variance). A second study using exploratory factor analysis found evidence a single structure (uni-dimensionality) accounted for 72.1% of the variance. Overall, structural validity was rated inconsistent (Table 6.2).

6.5.8 Internal consistency

As per the COSMIN guidelines, there must be evidence of sufficient structural validity before the evaluation of internal consistency (165). The results for individual studies (60, 110, 187, 188) on internal consistency are presented in Table 6.3, however, no further evaluation was completed as the results of structural validity were inconsistent.

6.5.9 Test-retest reliability

Test-retest reliability was reported in four studies of doubtful quality(60, 110, 187, 188) (Table 6.3). The ICC ranged from 0.88(95% CI 0.81-0.93) to 0.99 (95% CI 0.99-1.00). When results were pooled (Figure 6.2) the ICC was 0.95 (95% CI 0.84-0.98). Reliability of the VISA-H was rated down one level for risk of bias and one level for indirectness (multiple language versions) and overall provided moderate-quality evidence of sufficient test-retest reliability (Table 6.2).

6.5.10 Measurement error

Measurement error was reported in three studies of doubtful quality (60, 110, 188) (Table 6.3). The smallest detectable change (SDC) was reported in three studies and varied from 4.02 to 5.96 in patients with proximal hamstring tendinopathy (60, 110, 188). A single study reported on the SDC in patients awaiting surgery which was 4.32 (110). Two studies reported the minimal important change (MIC) of the VISA-H in non-surgical patients with proximal hamstring tendinopathy. One study used an anchor-based method with a seven-point global rating of change score (1 or 2 to classify a worsened patient, 3–5 to classify a stable patient, and 6 or 7 to classify an improved patient) and found the MIC (minimal important change) was 22 points (110) (using scores of 6 or 7 to classify an improved patient). A second study reported

Table 6.3 Structural validity, internal consistency, reliability and measurement error of the VISA-H

PROM, study (year)	Structural validity			Internal consistency			Reliability (test-retest)			Measurement error		
	n	ROB	Result	n	ROB	Result	n	ROB	Result	n	ROB	Result
VISA-H, Cacchio (2014)	60	A	Two factors accounted for 34.1% (items 1-6) accounted for 39.3% (items 7 and 8) of the variance	60	I	0.84 (95% CI 0.77 to 0.89)	55	D	Asymptomatic (n= 30): 0.92 (95% CI 0.80 to 0.97) Non-surgical (n=16): 0.92 (95% CI 0.80 to 0.97) Surgical (n=9): 0.90 (95% CI 0.63 to 0.97)	55	D	Asymptomatic (n= 30): SEM = 0.25, SDC = 0.69 Non-surgical (n=16): SEM =1.35, SDC = 3.74 Surgical: (n=9): SEM = 1.56, SDC = 4.32
VISA-H.F, Locquet (2019)	NT	-	-	16	I	0.85	16	D	0.92 (95% CI 0.80 to 0.97)	NT	-	-
VISA-H.Sp, De-la-Cruz-Torres (2021)	51	A	One factor accounted for 72.1% of the variance	101	D	0.88	51	D	0.99 (95% CI 0.99 to 1.00)	51	D	SEM: 1.45 SDC: 4.02
VISA-H.Br, Lima (2022)	NT	-	-	90	I	0.96	40	D	0.88 (95% CI 0.81-0.93)	90	D	SEM: 2.15 SDC: 5.96
Pooled result	±			?			+			+		

+ = sufficient, - = insufficient, ± = inconsistent, ? = indeterminate
CI = confidence interval, ROB = risk of bias (V = very good, A = adequate, D = doubtful, I = insufficient), n = sample size, NT = clinimetric property not tested in study, PROM = patient-reported outcome measure, PHT = proximal hamstring tendinopathy, SEM = Standard error of measurement, SDC = smallest detectable change, VISA-H = Victorian Institute of Sport – proximal hamstring tendinopathy, VISA-H.F = Victorian Institute of Sport Assessment - Proximal hamstring tendinopathy - French, VISA-H.Sp = Victorian Institute of Sport Assessment - Proximal hamstring tendinopathy – Spanish, VISA-H.Br = Victorian Institute of Sport Assessment - Proximal hamstring tendinopathy – Brazil

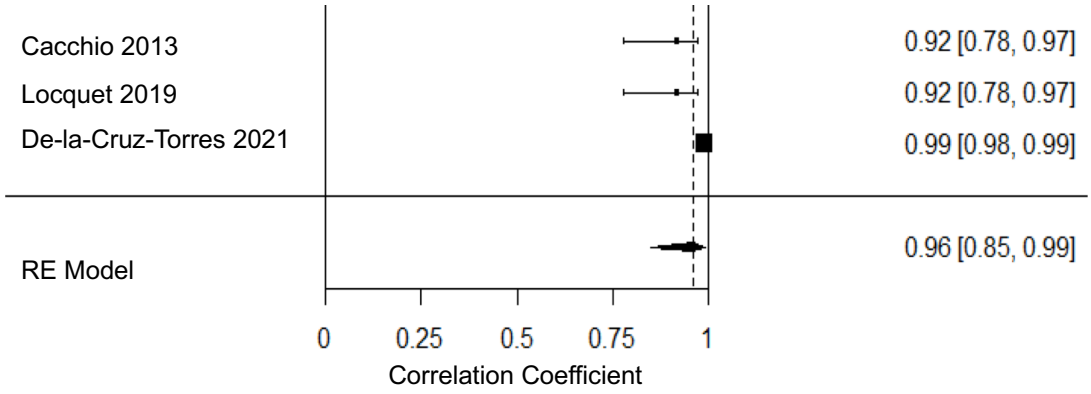


Figure 6.2 Meta-analysis of studies reporting on test-retest reliability of the Victorian Institute of Sport Assessment – Proximal Hamstring Tendinopathy using the intraclass correlation coefficient

RE = random effects model

the clinically significant improvement using a 30% change in VISA-H score from a baseline of 15 points (188). The measurement error was rated down one level for risk of bias and indirectness (multiple language versions) and overall provided low-quality evidence of sufficient measurement error (Table 6.2).

6.5.11 Hypothesis testing for construct validity

Comparison with other outcome measure instruments was explored in three adequate quality (60, 187, 188) and one study of inadequate quality (110) (Table 6.4). The following outcome measures were used as comparators: Functional Assessment Scale for Hamstring Injury, Short Form-36, Lower Limb Functional Index, Lower Extremity Functional Scale, Nirschl Phase Rating Scale, Generic tendon grading system - Curwin and Stanish (60, 110, 187, 188). When the results were summarized 24 results were found to be in line with the hypothesis and 5 results were not in line with the hypothesis (Appendix J). Overall, there was moderate-quality evidence of sufficient construct validity (Table 6.4). Known group validity was assessed in one study of very good quality (110), two studies of adequate quality (60, 188) and one study of doubtful quality (187). When results were summarised, eight results (100%) were in line with the hypothesis. Overall, there was moderate-quality evidence for sufficient hypothesis testing for construct validity (known group) (Table 6.2).

Table 6.4 Construct validity and responsiveness of the VISA-H

PROM, study (year)	Hypotheses testing for construct validity (other outcome measures)			Hypotheses testing for construct validity (known group)			Responsiveness (comparison with outcome measures)			Responsiveness (pre and post intervention)		
	n	ROB	Result	n	ROB	Result	n	ROB	Result	n	ROB	Result
VISA-H, Cacchio (2014)	25	I	8 in line with hypothesis (+8)	55	V	3 results in line with hypo (3+)	55	I	2 results in line with hypo (2+)	16	V	Non-surgical patients (n=16) AUC = 0.90 MIC = 22 ES = 2.2 SRM = 1.6 surgical patients (n=9): ES = 3.3 SRM = 2.2
VISA-H.F, Locquet (2019)	51	A	5 in line with hypothesis (+5) 3 not in line with hypothesis (-3)	51	D	3 results In line with hypo (3+)	NT	-	-	NT	-	-
VISA-H.Sp, De-la-Cruz-Torres (2021)	51	A	10 in line with hypothesis (+10) 2 not in line with hypothesis (-2)	101	A	2 results in line with hypo (2+)	101	A	2 results in line with hypo (2+)	51	D	Discharge: ES = 2.75 SRM =3.1 3-month follow-up: ES = 1.51 SRM = 0.59
VISA-H.Br, Lima (2022)	90	A	1 in line with hypothesis (+1)	90	A	2 results in line with hypo (2+)	NT	-	-	NT	-	-
Pooled result	+ 24 (24+) in line with hypothesis 5 (-5) not in line with hypothesis			+			+			+		

+ = sufficient, - = insufficient, ± = inconsistent, ? = indeterminate
AUC = area under curve, ES = effect size, hypo = hypothesis testing, CI = confidence interval, ROB = risk of bias (V = very good, A = adequate, D = doubtful, I = insufficient), n = sample size, NT = clinimetric property not tested in study, PROM = patient-reported outcome measure, SRM = standard response mean, VISA-H = Victorian Institute of Sport – proximal hamstring tendinopathy, VISA-H.F = Victorian Institute of Sport Assessment - Proximal hamstring tendinopathy - French, VISA-H.Sp = Victorian Institute of Sport Assessment - Proximal hamstring tendinopathy – Spanish, VISA-H.Br = Victorian Institute of Sport Assessment - Proximal hamstring tendinopathy – Brazil

6.5.12 Cross-cultural validity

None of the translations of the VISA-H underwent cross-cultural validation, which requires the assessment of measurement variance across different populations. As such the cross-cultural validity of the VISA-H was rated indeterminate.

6.5.13 Responsiveness

Two studies reported on the responsiveness of the VISA-H through comparison of VISA-H change scores with other outcome measures (110, 188) (Table 6.4). One study of insufficient quality performed hypothesis testing with the Nirschl Phase Rating Scale and Generic tendon grading system - Curwin and Stanish in non-surgical patients and patients awaiting surgery for proximal hamstring tendinopathy (110). Another study of adequate quality performed hypothesis testing with runners and triathletes with proximal hamstring tendinopathy, with the Functional Assessment Scale for Hamstring Injury and Lower Limb Functional Index (188). When results were summarised three results (100%) supported the hypothesis. One study of very good quality (110) and one study of doubtful quality (188) reported on the responsiveness of the VISA-H through comparison of the VISA-H before and after an intervention. One study included either surgical or a non-surgical intervention that included relative rest, avoiding activities and/or exercises that would increase the severity of symptoms and shockwave therapy (110). An anchor-based method using a seven-point Global Rating of Change (GROC) score was used (1 or 2 to classify a patient whose condition worsened, 3–5 to classify a stable patient, and 6 or 7 to classify an improved patient). Effect sizes were 2.2 and 3.3 for non-surgical and surgical patients respectively. The area under the curve (AUC) was reported for non-surgical patients only and was 0.90 (95% CI 0.70 to 0.98) (110). In the second study of doubtful quality, all participants underwent non-surgical management which included relative rest, restriction of the activities or exercises that increase the severity of symptoms, exercise, manual therapy, education, and electrotherapeutical modalities (188). The effect size was 2.75 and 1.51 at discharge and 3 months follow-up respectively. The standard response mean was 3.1 and 0.59 at discharge and 3 months follow-up respectively. Overall, there was low-quality evidence of sufficient responsiveness of the VISA-H for hypothesis testing for comparison with other outcome measure instruments and low-quality evidence of sufficient responsiveness before and after intervention (Table 6.2).

6.5.14 Feasibility and interpretability

Aspects of interpretability and feasibility were reported in four studies on the VISA-H (Appendix K and L) (60, 110, 187, 188). The time to complete the VISA-H was reported in two studies, which was 2 minutes 15 seconds (VISA-H.Br) (60) and less than 5 minutes (VISA-H.Sp) (188). Floor and ceiling effects were reported in four studies (188). No patient with proximal hamstring tendinopathy achieved the highest or lowest possible score on the questionnaire (60, 110, 187, 188). Mean scores on the VISA-H in healthy controls were reported in three studies and ranged from 97.7 (± 4.77) to 100 (95–100) points (100 points indicate no disability). Mean baseline scores in patients with proximal hamstring tendinopathy undergoing non-surgical management ranged from 50.2 (±14.5) to 59 (±15.81) (60, 110, 187, 188). Mean baseline scores of patients before undergoing surgery were reported in one study which was 45.8 (±12.2) (110).

6.6 Discussion

This study aimed to summarise outcome measures used in research on proximal hamstring tendinopathy, map outcome measures to the tendinopathy ICON core health domains and evaluate their measurement properties. A total of twenty-seven outcome measures were used in studies on proximal hamstring tendinopathy. The most common outcome measures used were adverse events and return to sport (pre-injury level). Only a single outcome measure was validated for use in patients with proximal hamstring tendinopathy – the VISA-H. Four studies reported on the measurement properties of the VISA-H (60, 110, 187, 188), however, the outcome measure did not demonstrate sufficient measurement properties to warrant recommendation.

Many outcome measures used to evaluate proximal hamstring were originally designed to evaluate other conditions, such as intra-articular hip pathology (e.g. hip osteoarthritis or labral pathology). Outcome measures developed for intra-articular hip conditions, such as the International Hip Outcome Tool and Hip Outcome Score, may miss key health-related aspects of proximal hamstring tendinopathy and focus on other features that may not relevant (e.g. items related to joint stiffness). The VISA-H was developed in 2014 and was the first condition-specific measure to evaluate proximal hamstring tendinopathy (110). A key advance in the development and evaluation of outcome measures is the positioning of patients as experts, in determining the content validity (165). The content validity of the VISA-H was

established informally through interviews with patients and focus groups of clinical experts (110). COSMIN guidelines instruct that evaluation of content validity should include judgment on the relevance of items included (to ensure they represent the key concepts of the condition), that items cover all key concepts (comprehensiveness) and that questions are appropriately phrased (comprehensibility) (165). Future research, placing more central involvement on patients, is required to better establish the content validity of outcome measures used to evaluate proximal hamstring tendinopathy.

It is important that we have outcome measures to evaluate tendinopathy that are relevant for all populations that may be impacted by the condition. The VISA-H was designed for athletic populations - multiple items on the scale ask questions suited to physically active populations with a strong focus on running (110). For example, items three and four enquire about pain during running and sprinting (110). Whilst the condition commonly occurs in runners who participate in sport (46), it also occurs in less active populations who do not engage in running or sprinting (161). The sporting focus is also present in items seven and eight, which evaluate pain during physical activity, with the heavy weighting of this section likely to have a considerable impact on the total score (193, 194). For example, a sedentary patient who does not participate in sports/physical activity may score zero on items three, four and seven and eight, meaning they could score a maximum of 40/100 points. A separate scale or subscale is required for less active individuals or those who do not participate in running.

An outcome measure with the absence of subscales should only measure a single construct (uni-dimensionality) if calculated as a single score (165). The VISA-H is calculated as a single total score out of 100 points. The original development study of the VISA-H reported the outcome measure had a two-factor structure that was designed to assess three domains (pain, function and sporting activity) (110). If this is true, the VISA-H should contain separate subscales. Exploratory factor analysis suggested that the VISA-H.Sp (Spanish version) measures a single construct (severity of symptoms), which would validate the use of a single total score (188). Further research with larger sample sizes using confirmatory factor analysis is required to assess the structure of the VISA-H. Until this time, caution should be taken when interpreting the VISA-H as a total score. Researchers using the VISA-H should consider reporting data for each item of the VISA-H separately (as a supplementary file) so these data can be re-analysed in the future if a lack of uni-dimensionality is confirmed.

To better capture outcomes for proximal hamstring tendinopathy and improve the synthesis of studies, it is critical that we have valid, reliable and responsive outcome measures that capture the nine ICON core health domains (164). The VISA-H was categorised within the disability domain, however, items within the questionnaire may capture other core health domains. This includes the pain on loading (e.g. item two enquires about pain during or immediately after hamstring stretching and item five after a lunge) and the participation domain (item eight asks about participation in sport). A qualitative study that interviewed expert clinicians on the condition proposed several other outcome measures that could be used to measure pain on loading (161). This included the numerical rating scale of pain with an arabesque or single-leg bridge (161). These items of the VISA-H may be of use in the interim for assessing the pain on loading and participation domains, but require research to evaluate their measurement properties.

There are several limitations of this systematic review. In our search to identify outcome measures, we did not include case series with less than ten participants or protocol papers, meaning we may have missed some outcome measures studied in proximal hamstring tendinopathy. The results for measurement error should be interpreted with caution as the minimal important change used to calculate measurement error is not a fixed value and may vary according to the method of calculation, context and population (195). Finally, whilst COSMIN is the accepted standard for completing systematic reviews evaluating measurement properties, limitations of the COSMIN have been cited including a lack of studies establishing the reliability of these guidelines (172).

6.7 Conclusion

The only measure with clinimetric properties in proximal hamstring tendinopathy was the VISA-H – which itself requires further evaluation and should be used with caution. Measures need to be developed or existing measures validated in proximal hamstring tendinopathy to allow clinicians and researchers to better understand the impact the condition has on patients.

6.8 Summary of findings

This study aimed to examine the types of measurement properties used in research on proximal hamstring tendinopathy, categorise measures into the core tendinopathy domains and evaluate their measurement properties. There were 27 distinct outcome measures identified. The most common outcome measure used in studies on proximal hamstring tendinopathy were adverse events and return to sport (pre-injury level). The health domain most reported in studies were the participation and disability domains. Outcome measures matching to the psychology, physical function capacity and pain on activity/loading domains have not been reported in the literature on proximal hamstring tendinopathy. Four studies reported on the measurement properties of the VISA-H. The measurement properties of no other outcome measure have been evaluated in patients with proximal hamstring tendinopathy. We found no outcome measure demonstrated sufficient measurement properties to be recommended for use. Until further evidence is provided the VISA-H is the best available outcome measure to evaluate the disability caused by proximal hamstring tendinopathy.

7 Chapter 7: Core outcome set for proximal hamstring tendinopathy (COS-PHT); a survey of an international collaboration

7.1 Preface

In the preceding chapter outcome measures that had been used in research on proximal hamstring tendinopathy were mapped to the nine established core health domains for tendinopathy.

It is not known whether these outcome measures mapped to core health domains assess important aspects of proximal hamstring tendinopathy (assessed by the OMERACT definition of truth) or if they are practical for use. To determine the truth and feasibility of outcome measures we recruited and surveyed thirty experts in proximal hamstring tendinopathy.

7.2 Abstract

Objective

Various measures have been used to assess treatments for proximal hamstring tendinopathy. This study aims to survey an international group of healthcare practitioners to determine which measures should be considered for use in future trials for each of the nine core tendinopathy health domains.

Methods

We surveyed health care practitioners to evaluate outcome measures that have been used in research on published trials on proximal hamstring tendinopathy. Outcome measures were evaluated using the Outcome Measures in Rheumatology filters of truth, feasibility and discrimination. A threshold of 70% agreement was set a priori for consensus.

Results

Results were obtained from 30 health care practitioners who worked in eight different countries. Of the 21 outcome measures presented, four outcome measures met consensus for inclusion for both truth and feasibility: two outcome measures in participation in life activities domain (return to sport pre-activity level, return to previous level of activity, and one outcome measure in each of the patient rating of condition (global rating of change) and disability (VISA-H) domains. Two outcome measures met the criteria for exclusion: return to sport (level not defined) and the Modified Harris Hip Scale. No outcome measures used previously in research for proximal hamstring tendinopathy met consensus for pain on activity/loading, function, psychological factors, physical function, quality of life, or pain over a specified period of time. A key theme was that measures used to assess intra-articular pathology of the hip did not meet thresholds for 'truth' and therefore should not be used in research on proximal hamstring tendinopathy.

Conclusion

We gained consensus on the truth and feasibility of a measure for the patient rating of condition, participation in life activities and disability domains. Patient opinions of these measures must be considered to further understand their utility.

7.3 Introduction

Proximal hamstring tendinopathy is a cause of long-standing buttock pain (58, 161). The tendinopathy is common in sports that involve running such as distance running and Australian Rules football but also impacts active, post-menopausal women (46, 161). Pain is often provoked by activities that involve lunging, running and sitting (155, 161).

To understand the effect of interventions on proximal hamstring tendinopathy, outcome measures used in studies must be reliable, valid and responsive (165). A recent systematic review on proximal hamstring tendinopathy recognised that a variety of different outcome measures have been used in research (155). This made it challenging to deduce meaningful conclusions on the efficacy of interventions and generate recommendations for clinical practice. Consistent selection of outcomes in clinical trials on proximal hamstring tendinopathy is important for the assessment of interventions between studies. This allows for the synthesis of evidence through meta-analysis. The selection of outcome measures that have poor validity, reliability or responsiveness in populations with proximal hamstring tendinopathy may influence results.

A solution would be to agree on a group of outcomes for proximal hamstring tendinopathy that are used in all future trials. This study is a part of a project to develop a core outcome set for clinical trials reporting on proximal hamstring tendinopathy (COS-PHT). The initial phase of establishing a COS-PHT was to conduct a systematic search for outcomes used in research and to then map the outcome measures to the core health domains for tendinopathy (164). This study now uses this data to survey health professionals on the truth and feasibility of outcome measures used in studies on proximal hamstring tendinopathy, within each of the International Tendinopathy Symposium Consensus (ICON) nine health domains (164).

7.4 Methods

7.4.1 Study registration

This study is part of the development of a Core Outcome Set for proximal hamstring tendinopathy. The study was pre-registered in May 2021 with COMET (Core Outcome Measures in Effectiveness Trials) (http://www.comet-initiative.org). The detailed methods of this study are reported in Chapter five.

7.4.2 Participants

Participants were recruited into a study working group. These professionals consisted of researchers and health care professionals with significant experience in the area of proximal hamstring tendinopathy. Researchers who had published on proximal hamstring tendinopathy in the previous ten years (2011-2021) were invited to join the study working group. Expertise was defined through the completion of relevant post-graduate training or clinical experience (minimum of ten years of experience). Professionals were recruited to achieve diversity of profession (e.g. Sports Physicians, Physiotherapists), sex and country of work. Participation in this study was voluntary. If authors did not respond to the initial email, the invitation to participate was re-sent after 10 days.

Professionals were invited to take part in the project prior to the release of the survey. The invitation included basic information on the study, including objectives and expectations. If the professional agreed to participate a link to access the survey was sent to the professional's email address. Emails were found online from freely available information or through published work. If professionals did not reply to the second email, they were excluded. Professionals were invited to be authors of the resulting publication if their contributions met the International Committee of Medical Journal Editors requirements.

7.4.3 Outcome measures

The preceding study (Chapter six) involved retrieving outcome measures from the literature and mapping them to one, or more, of the nine ICON domains. There were 27 outcome measures identified. The most common outcome measure used in research on proximal hamstring tendinopathy was adverse events and return to sport (pre-injury level). No outcome measures were mapped to the psychology, physical function capacity and pain on activity/loading domains in any study.

7.4.4 Survey

The survey was developed using the online software Qualtrics^{R} (Provo, Utah, USA) and reported in accordance with the Checklist for Reporting Results of Internet e-Surveys (CHERRIES) (170). Ethics were granted by the La Trobe University Human Ethics Committee (HEC21210). Respondents were given four weeks to complete the survey. A reminder email

was sent to professionals who had not completed the survey two weeks after the survey was released.

Respondents were required to answer questions on the truth and feasibility of outcome measures that had been used in research on proximal hamstring tendinopathy. The assessment of truth asks respondents to consider if the outcome measure used measures what it intends to measure and if the result of the outcome measure is unbiased and relevant (167). Feasibility assesses the ease with which an outcome measure can be administered given restraints of time, money, and interpretability (167).

In the survey, outcome measures were listed in random order in the relevant ICON core health domain (164). Respondents were made aware of the random order. For items related to a measure, respondents were given the following options: agree, disagree or unsure. Following questioning on truth and feasibility within a domain, an open-ended question was used to capture information to support their decision. In cases where no outcome measures could be matched to a domain all examples of outcome measures listed in the ICON core health domains article were provided for respondents to consider. Respondents were also invited to propose outcome measures for consideration (164).

The survey included 25 questions (Appendix M). Links to a copy of each outcome measure were provided, as well as information on the frequency of its use. Demographic details were also collected in the survey (sex, age, profession and country of work).

7.4.5 Analysis

The percentage of agreement required for consensus inclusion and exclusion was decided a priori and can be seen in Table 7.1. The survey responses were exported from Qualtrics to Microsoft Excel (Version 16.16.27). Responses were reported as frequencies and percentages for the total number of responses obtained for each question. The number of professionals that started the survey and the completion rate of the survey were also reported.

Table 7.1 Consensus definitions

Inclusion	≥70% of participants selected ‘agree’ responses + <15% selected 'disagree'.
Exclusion	≥70% of participants selected 'disagree' responses + <15% selected ‘agree’ responses
Undecided	any outcome measures that reach neither criteria

7.5 Results

7.5.1 Participant demographics

The demographics of respondents in the working group are summarised in Table 7.2. Forty participants (28 males and 12 females) responded to the invitation, of ten which declined. All 30 participants who consented to being involved in the study completed the survey (all questions). There were 24 men and six women with a mean of 48 years of age (ranging from 31 to 67 years) from eight countries. The professionals that completed the survey self-identified as physiotherapists, clinician-researchers, Sports Medicine Physicians and researchers (Table 7.2). Professionals worked in eight countries, with Australia and America being the most common. All respondents answered all questions in the survey (no missing data).

Table 7.2 Participant demographics

Participant characteristics	Healthcare professionals (n=30)
Sex: Male	23
Age: mean (range)	48 years (31-67)
Profession (participants could choose more than one)	
Physiotherapist	21
Clinician researcher	16
Sports Medicine Physician	9
Orthopaedic surgeon	5
Researcher	1
Countries where professionals work:	
Australia	16
America	7
United Arab Emirates	2
Finland, Italy, Norway, Netherlands, Spain	1

7.5.2 Patient rating of the condition

One of the three outcome measures presented (global rating of change) achieved consensus for inclusion regarding truth (93.3%) and feasibility (96.7%). Two outcome measures (rating of symptom improvement, and overall rating of outcome) were undecided concerning truth but met consensus for feasibility (Table 7.3). Respondents expressed that the overall rating was a "simple and practical" measure to gather the patient rating of the condition domain. Two participants recommended the use of a 7 or 11-item Global rating of change scale and one preferred the use of a 15-item scale. Other measures identified as possible alternatives were the Patient Acceptable Symptom State, Single Assessment Numeric Evaluation, Numerical Rating Scale of improvement, Visual Analogue Scale, percentage of improvement, ability to sit comfortably (yes/no) and ability to return to running (yes/no).

7.5.3 Participation in life activities

Two of six outcome measures achieved consensus for inclusion regarding both truth and feasibility (return to pre-injury level of sport and return to previous level of activity). One outcome measure met consensus for feasibility alone (Tegner Score), and one outcome measure (return to sport – level not defined) achieved consensus for exclusion (truth). In open-ended responses, a key theme was that respondents thought that several measures such as the International Hip Outcome Tool were tailored towards hip joint pathology (e.g. intra-articular pathology) as opposed to tendinopathy. Examples included items within outcome measures that focused on hip mobility, which "are not a big issue for proximal hamstring tendinopathy" (Responder 11).

7.5.4 Pain on activity or loading

No outcome measures previously used in research on proximal hamstring tendinopathy were mapped to the pain on activity or loading domain. Measures that were proposed by respondents involved activities where the hamstring musculotendon unit was loaded in positions of hip flexion, such as the arabesque/single leg deadlift, lunging with the affected leg placed forwards. One responder also mentioned that the Victorian Institute of Sport Assessment – Proximal Hamstring (VISA-H) includes activities that could be used to fulfil this domain.

Table 7.3 Results of Truth and Feasibility

Filter	Truth (n = 30)			Feasibility (n = 30)		
Domain	Y	N	?	Y	N	?
Patient Rating of Condition						
Global rating of change scale	93.3%	3.3%	3.3%	96.7%	0%	3.3%
Rating of symptom improvement (e.g. yes, no)	46.7%	36.7%	16.7%	93.3%	0%	6.7%
Overall rating of outcome (e.g. poor, fair, good or excellent)	53.3%	23.3%	23.3%	90.0%	3.3%	6.7%
Participation in life Activities						
Return to Sport (level not defined)	6.7%	70.0%	23.3%	60.0%	33.3%	6.7%
Return to Sport (pre-injury level)	80.0%	13.3%	6.7%	93.3%	6.7%	0.00%
Return to previous level of activity	70.0%	16.7%	13.3%	93.3%	6.7%	0.00%
Tegner Score	43.3%	33.3%	23.3%	70.0%	13.3%	16.7%
International Hip Outcome Tool (iHOT-33 Sports and recreational activities subscale)	56.7%	20.0%	23.3%	66.7%	20.0%	13.3% %
International Hip Outcome Tool (iHOT-33 Job related concerns sub scale)	26.7%	53.3%	20.0%	56.7%	30.0%	13.3%
Function Domain						
Lower Extremity Functional Scale (LEFS)	50.0%	33.3%	16.7%	66.7%	10.0%	23.3%
Level of Function (% of full activity)	33.3%	33.3%	33.3%	86.7%	6.7%	6.7%
Disability						
Victorian Institute of Sport Questionnaire - Proximal Hamstring (VISA-H)	82.8%	3.5%	13.8%	93.1%	0%	6.9%
Nirschl Phase Rating Scale (NPRS)	27.6%	37.9%	34.5%	62.1%	17.2%	20.7%
Hip Outcome Score - Activities of Daily Living (HOS-ADL)	34.5%	51.7%	13.8%	79.3%	13.8%	6.9%
Hip Outcome Score (HOS-Sport)	37.9%	37.9%	24.1%	82.8%	13.8%	3.5%
Modified Harris Hip Scale (MHHS)	6.9%	79.3%	13.8%	62.1%	27.6%	10.3%
Lower Extremity Functional Scale (LEFS)	37.9%	41.4%	20.7%	76.0%	17.2%	6.9%
International Hip Outcome Tool (iHOT-33 Symptoms and functional limitations)	24.1%	48.3	27.6%	62.0%	27.6%	10.3%
Quality of Life						
International Hip Outcome Tool (iHOT-33 social emotional and lifestyle concerns subscale)	44.8 %	27.6%	27.6%	65.5%	24.1%	10.3%
Pain Over a Specified Timeframe						
Mean pain at a specified time point (e.g. 12 weeks)	48.3%	37.9%	13.8%	75.9%	13.8%	10.3%

Data are % responses, with green representing ≥70% agree, red ≥70% disagree, and amber neither green or red
There were no outcome measures that were mapped to pain on activity loading, psychological factors, physical function capacity domain in studies on proximal hamstring tendinopathy.
Y = yes, N = no, ? = unclear

7.5.5 Function domain

None of the two outcome measures met consensus for both truth and feasibility for the function domain. One outcome (level of function - % of full activity) met consensus for inclusion (feasibility only). A key theme was that respondents thought that several measures such as the lower extremity functional scale (LEFS) included outcome measures that were tailored towards hip joint issues (e.g. intra-articular pathology).

Patient-reported outcome measures suggested by respondents for this domain included the VISA-H, Lower Limb Functional Index, relevant items of the International Hip Outcome Tool -12, patient-specific functional scale, visual analogue scale, return to prior level of activity (and if no % of improvement), running up a hill, gym-based exercise including lower back raise, leg press and leg extension.

7.5.6 Psychology

No outcome measures that had been previously used in research on proximal hamstring tendinopathy were mapped to the psychology domain. Measures suggested by respondents included the pain catastrophizing scale, Tampa Scale for kinesiophobia, pain self-efficacy questionnaire, Hospital Anxiety and Depression Scale, Hip-Return to Sport after Injury scale, short-form Orebro Musculoskeletal Screening Questionnaire, Patient Health Questionnaire, pain coping inventory, EuroQol-5D, Assessment of Quality-of-Life instrument, short-form 36 and Optimal Screening for Prediction of Referral and Outcome tool. Most respondents suggested that the psychology domain was an important domain to capture in patients with proximal hamstring tendinopathy, however, one responder thought that collection of this domain routinely (as a core measure) was unnecessary.

7.5.7 Physical function capacity

No outcome measures that had previously been used in research on proximal hamstring tendinopathy were mapped to the physical function capacity domain. Several respondents expressed isometric dynamometry of the knee flexors should be considered (isometric dynamometry measured in 90 degrees hip flexion and knee extension was most commonly proposed). Other options were single-leg bridges off a box (number performed/compared to contralateral side), timed stair walk, number of hops, single-leg squats, time sitting on a hard surface, repetitions of hip extension exercise, single-leg deadlift, active knee extension test and repeated sprint ability test.

7.5.8 Disability

Of the seven items that measured the disability domain, one item met consensus for inclusion for both truth and practicality, the VISA-H. Three measures met the threshold for inclusion for feasibility only: the Hip Outcome Score – Activities of Daily Living, the Hip Outcome Score – Sport, and the Lower Extremity Functional Scale. One measure met the threshold for

exclusion based on truth: the Modified Harris Hip Scale. A theme from respondents was that the VISA-H was important to include as it was the only condition-specific patient-reported outcome measure. However, respondents expressed that it may not be relevant to non-sporting populations and item eight within the VISA-H was impractical, as it was confusing to respondents. One responder commented "The Nirschl Scale is a simple measure that could apply particularly for runners but is not specific to proximal hamstring tendinopathy and clinimetrics would need to be done to establish its usefulness" (Responder 22). Another key theme was that many of the included outcome measures were tailored towards hip joint issues (e.g. intra-articular pathology or groin pain) – "the others are all hip joint related scores with many items not applicable or missing for proximal hamstring tendinopathy and the wording hip/groin not appropriate" (Responder 15). There was one additional outcome measure suggested for consideration which was the Copenhagen Hip and Groin Outcome Score-sport.

7.5.9 Quality of life

One outcome measure used in research to measure the quality-of-life domain (International Hip Outcome Tool 33 – social, emotional and lifestyle concerns subscale), however, it did not reach a consensus for inclusion for either truth or feasibility. As with other domains, there were suggestions that there was too much emphasis on the hip joint, as opposed to the proximal hamstring tendon, throughout the measure. Other suggestions included the EuroQoL 5D 5L, short form-36 and the musculoskeletal Health Questionnaire. A recurring theme was that the outcome measure chosen to represent this domain should include items related to the impact of the condition on sitting (e.g. "sitting meeting friends" (Responder 19) and 'sitting at work' (Responder 25)).

7.5.10 Pain over a specified time

There was one outcome measure included for pain over a specified timeframe (mean pain at a specified timeframe). This outcome met consensus for inclusion for practicality but was indeterminate for truth. A key theme was that the question used to report pain at a specified time was unclear (mean pain at a specified time point (e.g. 12 weeks)) – with the suggestion from several respondents that the outcome measure must include further descriptions (e.g. worst pain in the last week).

There was also a suggestion that because the pain comes and goes with tendinopathy asking for pain at a given time point might not reflect their condition status (e.g. pain level now). To this point, one responder suggested that the timeframe of "…a one-week period may be enough for most athletes (runners, field sport athletes) as it captures their regular training (interval sessions, matches, training etc.)" (Responder 18). There were suggestions that when considering this domain patients should not be asked to recall pain over a period longer than one week.

Respondents suggested several other options including: worst pain over the last week, mean pain in the last 24 hours, pain constancy (e.g. % of time pain has been present over the last week), morning pain (when getting out of bed), average estimate throughout the day (including extremes), greatest pain during a meaningful task, maximal pain over the past 24 hours, mean pain in the last 24 hours, the mean level of pain over the past week.

7.6 Discussion

This study was part of an international, multi-disciplinary collaboration to establish a core outcome set for proximal hamstring tendinopathy. This study aimed to investigate the truth and feasibility of outcome measures that have been previously used in the literature on proximal hamstring tendinopathy. We surveyed 30 respondents from eight countries. Four candidate measures met inclusion criteria for both truth and feasibility filters. Two measures in the participation in life activities domain (return to sport pre-activity level and return to previous level of activity), and a single measure from the patient rating of condition domain (global rating of change) and disability domain met consensus for inclusion for both truth and feasibility (VISA-H). No outcome measures met consensus for inclusion in the other six domains. Two outcome measures met consensus for exclusion: return to sport (level not defined) and the Modified Harris Hip Scale, which lie in the participation in life activities and disability domain respectively.

The most important measurement property is content validity, which within the COSMIN criteria is captured by the relevance, comprehensiveness and comprehensibility of an outcome measure (165). Outcome measures originally developed to evaluate hip joint pathology have been frequently used in research on proximal hamstring tendinopathy (155). Examples include the Hip Outcome Score and the Modified Harris Hip Score, which were based on the Harris

Hip Score (196, 197). In this study, many respondents reported that such measures included items that were not relevant for patients with proximal hamstring tendinopathy. Examples included items related to hip 'clicking' and hip joint mobility impairments. This included activities such as getting in and out of a car or stretching the leg into abduction. If an item(s) of an outcome measure does not commonly impact patients with proximal hamstring tendinopathy it will not be responsive to change when used. Respondents to a questionnaire (e.g. patients) may also feel dissatisfaction or become frustrated if questions do not relate to the specific physical impairments of the condition. As such, questionnaires should be designed and reviewed by the population they are being used in to ensure that the majority of items presented are relevant.

Before the development of the VISA-H in 2013, no specific patient-reported outcome measure had been developed and validated for patients with proximal hamstring tendinopathy. The framework for outcome measure development has advanced since the development of the VISA-P, from which the VISA-H was based. This has led to the need to re-evaluate outcome measures to ensure they have measurement properties that are satisfactory for use (193). Specific concerns about the VISA questionnaires include that many of the questions are not relevant for the varying activity levels of patients who present with the condition, the lack of patient consultation in many and the lack of evidence for uni-dimensionality (structural validity) and the lack of practicality of item-eight (194, 198, 199).

Pain over a specific time-point is one of the nine core health domains encouraged to guide reporting in trials in tendinopathy (164). It is often evaluated using an 11-item visual analogue scale (VAS) or numerical rating scale (NRS). Reporting on pain as an outcome measure requires improvement, with five of six studies not meeting reporting recommendations based on the ICON consensus as they did not specify a reference period of time over which the patient was rating pain (164). For example, asking patients – what was your worst pain over the past week? Further description regarding return to sport should also be included. Whilst most studies reported on whether patients returned to pre-injury level of sport, additional details on the type of sport athletes returned to, as well as their level of performance, would be meaningful in the interpretation of results.

The variety and diversity of professional backgrounds were a strength of this study. A

limitation of this study was that many of the respondents worked in Australia and were men. This was likely due to the method of recruitment which largely involved recruiting professionals who had previously published on proximal hamstring tendinopathy.

This study has resulted in the progression of four candidate outcome measures for consideration to be part of a core outcome set. Before recommendation, consideration of both the measurement properties and patient views of these measures is required. This will assist in improving confidence that outcome measures are used to capture the range of meaningful health-related core domains in patients with proximal hamstring tendinopathy (164).

7.7 Conclusion

Four candidate measures met consensus for inclusion for truth and feasibility filters: return to sport pre-activity level, return to previous level of activity, global rating of change and the VISA-H. Two outcome measures met the criteria for exclusion: return to sport (level not defined) and the Modified Harris Hip Scale. Data on the measurement properties and patient views are now required to further understand their utility.

7.8 Summary of findings

Thirty experts in proximal hamstring tendinopathy were surveyed on the truth and feasibility of outcome measures used in previous research on the condition. Four outcome measures met consensus for inclusion criteria for both truth and feasibility: return to sport pre-activity level, return to previous level of activity, global rating of change and the VISA-H. Two outcome measures met the criteria for exclusion: return to sport (level not defined) and the Modified Harris Hip Scale. Patient considerations (focus groups) on measures are required to further understand their utility. A key issue with outcome measures was that many were designed to evaluate the impact of other hip conditions and that several measures were not adequately detailed to allow for reliable use in research.

8 Chapter 8: Outcome measures in the management of gluteal tendinopathy: a systematic review of their measurement properties

8.1 Preface

Chapters six and seven identified that there is little information on the measurement properties of outcome measures used to evaluate proximal hamstring tendinopathy. Gluteal tendinopathy is another common tendinopathy that impacts the hip, that has been more widely evaluated than proximal hamstring tendinopathy.

The aim of Chapter eight (study six) was to appraise the measurement properties of outcome measures used in studies on gluteal tendinopathy. To achieve this, a three-phase process was followed. Firstly, a scoping review of the literature was performed to find all outcome measures used in research to evaluate gluteal tendinopathy. Secondly, the measures used were sorted into ICON tendinopathy domains. Finally, the measurement properties of each outcome measure were appraised following COSMIN guidelines.

The following chapter contains a modified version of the paper

Outcome measures in the management of gluteal tendinopathy: a systematic review of their measurement properties

Mr Anthony Nasser, Dr Angie Fearon, Dr Alison Grimaldi, Prof Bill Vicenzino, Dr Rebecca Mellor, Mr Trevor Spencer, Dr Adam Semciw

British Journal of Sports Medicine

URL link to published article: http://dx.doi.org/10.1136/bjsports-2021-104548

8.2 Abstract

Objective

Evaluate properties of outcome measures for gluteal tendinopathy.

Design

Multi-stage scoping/systematic review

Data Sources

Cochrane, PubMed, EMBASE, Scopus, Web of Science, PEDro, CINAHL, SPORTDISCUS were searched (December 2021) to identify measures used to evaluate gluteal tendinopathy. Measures were mapped to the core health domains for tendinopathy. MEDLINE, CINAHL, EMBASE and PubMed were searched (December 2021) for studies evaluating measurement properties of gluteal tendinopathy outcome measures captured in the initial search. Both reviews included studies that evaluated a treatment in participants with gluteal tendinopathy, diagnosed by a professional. Consensus-based-Standards for the Selection of Health Instruments methodology (COSMIN) were followed – including bias assessment and synthesis of findings.

Results

Six studies reported on the Victorian Institute of Sport Assessment – Gluteal Tendinopathy (VISA-G). One study reported on the Hip Outcome Score (HOS) –Activities of Daily Living (ADL) and Sport.

The VISA-G had moderate-quality evidence of sufficient construct validity (known group) and responsiveness (pre-post intervention), low-quality evidence of sufficient reliability, measurement error, comprehensibility and insufficient construct validity (convergent) and very low-level evidence of sufficient comprehensiveness, relevance and responsiveness (comparison with other outcome measures).

The HOS (ADL) and HOS (Sport) had very low-quality evidence for sufficient reliability and relevance, insufficient construct validity and comprehensiveness. The HOS (ADL) provided very low-quality evidence of sufficient comprehensibility and insufficient measurement error.

The HOS (Sport) provided very low-level evidence of inconsistent comprehensibility and sufficient measurement error.

Conclusion

Rigorously validated outcome measures are lacking. The VISA-G is the preferred option to capture the disability associated with gluteal tendinopathy.

8.3 Introduction

Gluteal tendinopathy presents as lateral hip pain that is commonly aggravated by activities such as ascending stairs, walking, running and lying on the side when sleeping (48). It is the most common lower limb tendinopathy in patients attending general practice (6), and is particularly prevalent in post-menopausal women (48).

Over the last decade the incidence and disabling nature of gluteal tendinopathy has been recognised (6, 45). This has plausibly led to an increase in the number of intervention trials for gluteal tendinopathy (38, 39, 200). The outcome of an intervention is best determined by using outcome measures that are feasible, reliable, valid and responsive to change (167, 172). To provide useful evidence for clinical practice guidelines, these outcomes should at least reflect the core health domains for tendinopathy: patient overall rating, participation, pain on activity/loading, disability, function, physical function capacity, quality of life, psychology, and pain over a specified timeframe (164). The group of outcomes that measure these domains are known as the core outcome set (COS) and are established through means outlined in the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN), Core Outcome Measures in Effectiveness Trials (COMET) and Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT) initiatives (167). Core outcome sets are needed to ensure homogeneity in selection and reporting of outcomes across clinical trials, as this allows pooling of data in meta-analyses – a fundamental tenet of evidence-based practice. Selection of outcome measures that have not been developed or validated for use in populations with gluteal tendinopathy, or selection of outcome measures from only a limited number of core health domains may bias outcome reporting and influence results, and therefore recommendations to clinical practice.

The preliminary phase of establishing a core outcome set for gluteal tendinopathy was to comprehensively search the literature for outcomes used in clinical trials, to provisionally map these to the core health domains for tendinopathy (164) and then to evaluate the retrieved outcomes for their measurement properties (167, 172). The aim of this study was to identify and evaluate outcome measures used in studies of gluteal tendinopathy in terms of their clinimetric properties (e.g. reliability and validity) and methodological quality, and then provide recommendations where possible.

8.4 Methods

8.4.1 Study design and registration

The proposal for the development of the core outcome set for gluteal tendinopathy was registered on COMET in 2019. This systematic review was completed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (201) following the COSMIN methodology (165, 172). The protocol for the scoping review (stage one) was preregistered online in December 2019 with the Research Registry (ID: 5287), the protocol for the systematic review (stage three) was preregistered in PROSPERO (CRD42020207452) in October 2020.

Stage one

8.4.2 Search strategy

With the assistance of a librarian Cochrane, PubMed, EMBASE, Scopus, Web of Science, PEDro, CINAHL and SPORTDISCUS were searched (December 2021). In order to capture outcome measures reported in grey literature ProQuest dissertations, Theses Global, NICE, OpenGrey and Google were searched. The search strategy for the PubMed database can be found in Appendix O.

8.4.3 Study selection

Titles and abstracts of articles identified from all databases were uploaded into the online platform COVIDENCE (a systematic review screening tool). After the removal of duplicates, two reviewers independently (AF and AG) screened titles and abstracts against the inclusion criteria. Disagreement was referred to a third reviewer (RM) for consensus. Full texts of the remaining articles were obtained and examined for inclusion independently by two reviewers (AF and AG). Disagreement was referred to a third independent reviewer (RM) for consensus.

8.4.4 Inclusion criteria

- Diagnosed with gluteal tendinopathy by a health professional (with or without radiological confirmation)
- Any study design on an intervention in which a treatment outcome for gluteal tendinopathy had been measured (e.g. papers relating to the development of condition-specific outcome measures, impairments or activity and participation limitations).

8.4.5 Exclusion criteria

We excluded diagnostic utility papers, trial registrations, case reports and case series (n<ten), surgical technique papers, opinion pieces, risk factor studies, papers that only used imaging (e.g. MRI) parameters as outcome measures, and papers that report on trochanteric pain or tendon pathology associated with infective or systemic/rheumatological conditions.
We excluded papers in languages other than English due to limited resources.

8.4.6 Data extraction

Characteristics of the study populations, definitions of the outcomes measured and the outcome measurement instruments used were extracted into an excel document by two independent reviewers (AF and AG). Discrepancies were resolved by an independent third reviewer (RM).

Stage two

8.4.7 Suggested mapping of core health domains

All outcome measures extracted (Part One) were mapped independently by three authors (BV, AF, RM) to the nine core health domains of tendinopathy (Appendix O) (164). If an item could not be mapped to a specific domain the item was labelled – "could not be mapped". Discrepancies were discussed by all authors until a consensus was established.

Stage three

8.4.8 Search strategy

A search was performed to identify all studies that reported on the measurement properties of outcome measures in patients with gluteal tendinopathy identified in part 1, from inception until December 2021 in MEDLINE, CINAHL, EMBASE and PUBMED. In PubMed, a validated search filter for studies on measurement properties was used (Appendix P). In MEDLINE, EMBASE and CINAHL we used search filters recommended by COSMIN (see https://www.cosmin.nl).

8.4.9 Study selection

Titles and abstracts of articles identified from databases were uploaded into COVIDENCE. After removal of duplicates, two independent reviewers screened titles and abstracts against inclusion criteria (AN and TS). Any disagreement was referred to an independent third reviewer for consensus (AF). Two independent reviewers (AN and TS) screened full-text

versions for all remaining studies, with disagreements referred to a third independent reviewer (AF) for consensus. To ensure all relevant articles were included, citation tracking (PubMed/ Google scholar) and reference checking of included studies was performed.

8.4.10 Inclusion criteria:

Studies had to include participants of at least 18 years of age who were clinically diagnosed with gluteal tendinopathy by a healthcare professional (with or without radiological confirmation). The following synonyms for gluteal tendinopathy were also considered: greater trochanteric pain syndrome, gluteus medius/minimus tendinopathy, greater trochanteric bursitis. Studies must have included clinimetric assessment of one or more outcome measures identified in the scoping review in part 1, that could be mapped to the core health domains of tendinopathy. By definition, these studies were to be of gluteal tendinopathy.

8.4.11 Exclusion criteria:

Narrative reviews, commentaries, editorials, conference proceedings, conference posters/abstracts, clinical studies (i.e. controlled trials, cohort studies, case-control studies or case studies), studies in languages other than English and animal and in vitro studies were excluded. Studies that included participants who had (a) congenital or acquired neurological or inflammatory disorders and (b) other hip conditions (e.g. referred pain, fractures, hip replacements) were also excluded.

8.4.12 Type of outcomes

Content validity, structural validity, internal consistency, reliability, measurement error, construct validity, cross-cultural validity, responsiveness, interpretability and feasibility were evaluated.

8.4.13 Deviations from protocol

The focus of this review was refined to only include studies of participants who had gluteal tendinopathy – this deviated from a broader musculoskeletal hip pain inclusion criterion in our protocol. The original intent was to ensure no studies were missed due to ambiguities in nomenclature for gluteal tendinopathy. After completing the search, it became apparent that the number of articles on hip pain broadly focused on other conditions of the hip, and were thereby not relevant, and would be impractical to review in a timely manner with an appropriate

degree of fidelity. A random selection of these papers (20%) revealed no sub-groups of gluteal tendinopathy within papers on hip pain – which conferred a level of confidence in restricting our clinimetric evaluation to those studies of gluteal tendinopathy.

8.4.14 Data extraction

Two independent authors (AN and TS) extracted information on the characteristics of the study populations and all data on measurement properties into an Excel spreadsheet. Any discrepancies in extraction were resolved by an independent third reviewer (AS). In cases of missing data, or to clarify information, authors were contacted. We also extracted all data in studies on interpretability and feasibility.

8.4.15 Methodological quality of included studies

The methodological quality of each study for each measurement property was assessed independently by two reviewers (AN and RM) using the COSMIN Risk of Bias Checklist (178). In this checklist each measurement property is rated on a 4-item checklist as either: very good, adequate, doubtful or of inadequate quality (172). A third independent reviewer (AS) was invited to settle any conflicts. The 'worst score counts' approach was applied to all items regarding a specific measurement property to derive a final rating (e.g. if for internal consistency one item in a box is rated as inadequate, the overall methodological quality of that study for that measurement property will be rated as inadequate) (172, 178). As per the COSMIN guidelines, criterion validity was not evaluated due to the lack of an established gold standard measurement tool for gluteal tendinopathy (172), with the exception of strength capacity, where isokinetic dynamometry was considered to be the gold standard (202).

8.4.16 Data analysis

Results from all outcome measure development studies and any additional studies on content validity were individually rated against the 10 established criteria for good content validity by two independent reviewers (AN & RM), with a third independent reviewer (AS) invited to resolve any disagreement (179). Each of the criteria could be scored as positive (+), negative (-), or indeterminate (?). The same criteria were also scored based on the content of the outcome measure itself (reviewer ratings). An overall sufficient (+), insufficient (-), or inconsistent (±) score was provided for relevance, comprehensiveness, and comprehensibility of each outcome measure, by jointly assessing all the results and reviewer ratings on the same

outcome measure (172, 179). If all studies indicated a "sufficient," "insufficient," or "indeterminate" rating for a specific measurement property, the overall rating of this measurement property was rated accordingly. If there were inconsistencies between studies, explanations were explored (e.g. differences in methodologic quality). The overall rating was scored indeterminate if insufficient information was available.

The results for all other measures were rated against the updated criteria for good measurement properties, by two independent reviewers (AN & RM), with a third independent reviewer (AS) invited to resolve any disagreement (172, 179). Each criterion could be scored as positive (+), negative (-), or indeterminate (?) (178). Pre-formulated hypotheses were used in the analysis of construct validity (decided by consensus) on top of the proposed and specified COSMIN criteria (Table 8.1).

8.4.17 Quality of the body of evidence

The Grading of Recommendations Assessment Development Evaluation (GRADE) were used to determine the confidence in the body of evidence. Quality of evidence was rated, as ''high'', ''moderate'', ''low'', or ''very low'', by considering study quality (risk of bias), consistency of results across studies, and reviewer ratings (content validity only) (165, 172). All studies started at high-quality evidence and could be rated down by up to three levels:
Risk of bias (content validity) – downgraded one level (serious) if the available content validity studies were of doubtful quality, two levels (very serious) if there were no content validity studies (or only of inadequate quality) and the development study was of doubtful quality and three levels (extremely serious) if there were no content validity studies (or only of inadequate quality) and the outcome measure development study was of inadequate quality (172).

Risk of bias (all other measurement properties) - downgraded one level (serious) if there were multiple studies of doubtful quality or only one study of adequate quality, two levels (very serious) if there are multiple studies of inadequate quality or one study of doubtful quality, or three levels (extremely serious) if there was only one study of inadequate quality (172).
and up to two levels for:

- Inconsistency - serious variation between results (e.g. measurement properties rated as sufficient and insufficient in different studies (172)

Table 8.1 Quality criteria for evaluated measurement properties

Measurement property	Definition	Rating	Criteria
Content validity	The degree to which the content of the VISA-G, HOS-Brazil (ADL) or (Sport) subscales are an adequate reflection of the construct to be measured (e.g. reflect the disability caused by GT). In terms of: i) Relevance - all items are relevant for the construct of interest within a GT population and context of use ii) Comprehensiveness - no key aspects of the construct should be missing iii) Comprehensibility - the items should be understood by patients with GT as intended	+	Relevance rating is +, comprehensiveness rating is +, comprehensibility rating is +
		-	Relevance rating -, comprehensiveness rating is -, comprehensibility rating is -
		±	At least one rating is + and at least one rating is –
Internal consistency	Interrelatedness of items within the VISA-G, HOS-Brazil (ADL) or (Sport) subscales	+	At least low evidence for sufficient structural validity AND Cronbach's alpha(s) ≥ 0.7 for each unidimensional scale or subscale. Weighted mean Cronbach's alpha (95% CI)
		?	Criteria for "at least low evidence for sufficient structural validity" not met
		-	At least low evidence for sufficient structural validity AND Cronbach's alpha(s) < 0.70 for each unidimensional scale or subscale
Test-retest reliability	Proportion of total variance in the measurements which is because of "true" differences among patients" in the VISA-G, HOS-Brazil (ADL) or (Sport) subscales	+	ICC or weighted Kappa ≥ 0.7
		?	ICC or weighted Kappa not reported
		-	ICC or weighted Kappa < 0.70
Measurement error	Systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured (e.g. disability for VISA-G, HOS-Brazil ADL and Sport subscales)	+	SDC or LoA < MIC
		?	MIC not defined
		-	SDC or LoA > MIC
Structural validity	Degree to which the scores of the VISA-G, HOS-Brazil (ADL) or (Sport) subscales are an adequate reflection of the dimensionality of the construct to be measured	+	CTT: CFA: CFI or TLI or comparable measure >0.95 OR RMSEA <0.06 OR SRMR <0.082 IRT/Rasch: No violation of unidimensionality: CFI or TLI or comparable measure >0.95 OR RMSEA <0.06 OR SRMR <0.08 AND no violation of local independence: residual correlations among

			the items after controlling for the dominant factor < 0.20 OR Q3's < 0.37 AND no violation of monotonicity: adequate looking graphs OR item scalability >0.30 AND adequate model fit: IRT: χ^2 >0.01 Rasch: infit and outfit mean squares ≥ 0.5 and ≤ 1.5 OR Z standardized values > -2 and <2
		?	CTT: Not all information for '+' reported IRT/Rasch: Model fit not reported
		-	Criteria for '+' not met
Cross-cultural validity\measurement invariance	Degree to which performance of items on a translated version of the VISA-G, HOS-Brazil (ADL) or (Sport) subscales are an adequate reflection of the performance of the items of their respective original versions (e.g. original English version of the VISA-G)	+	No important differences found between group factors (such as age, gender, language) in multiple group factor analysis OR no important DIF for group factors (McFadden's R^2 < 0.02)
		?	No multiple group factor analysis OR DIF analysis performed
		-	Important differences between group factors OR DIF was found
Hypotheses testing for construct validity	Degree to which the scores of the VISA-G, HOS-Brazil (ADL) and (Sport) subscales were consistent with hypotheses based on the assumption that each PROM validly measures the construct to be measured (e.g. being physical properties, more so than mental or emotional properties)	+	We proposed the following hypotheses: i. Strong correlations would be found with instruments measuring similar constructs (≥ 0.50) ii. Correlations with instruments measuring related, but dissimilar constructs would be lower (0.3-0.5) iii. Correlations with instruments measuring unrelated constructs to be <0.3. The construct validity was considered sufficient if at least 75% of our hypotheses are confirmed by analyses.
		?	No hypothesis defined by the review team
		-	The result is not in accordance with the hypothesis
Responsiveness	Ability of the VISA-G, HOS-Brazil (ADL) or (Sport) to detect change over time	+	The result is in accordance with the hypothesis OR AUC ≥ 0.7
		?	No hypothesis defined (by the review team)
		-	The result is not in accordance with the hypothesis OR AUC < 0.7
Interpretability	Distribution of scores in the study population, percentage of missing items and percentage of missing total scores, floor and ceiling effects, scores and change scores available for relevant (sub)groups, minimal important change (MIC) or minimal important difference, information on response shift	Narrative synthesis	

Feasibility	Patient's comprehensibility, clinician's comprehensibility, type and ease of administration, length of the instrument, completion time (minutes), patient's required mental and physical ability level, ease of standardisation, ease of score calculation, copyright, cost of instrument, required equipment, availability in different setting, regulatory agency's requirement for approval	Narrative synthesis

+ = measurement property is sufficient, - = measurement property is insufficient, ? = measurement property is indeterminate
AUC = area under the curve, CFA = confirmatory factor analysis, CFI = comparative fit index, CTT = classical test theory, DIF = differential item functioning, HOS-Brazil (ADL) = Hip Outcome Score-Brazil (Activities of Daily Living), HOS-Brazil (Sport) = Hip Outcome Score-Brazil (Sport), ICC = intraclass correlation coefficient, IRT = item response theory, GT = gluteal tendinopathy, LoA = limits of agreement, MIC = minimal important change, PROM = patient-reported outcome measure, R^2 = measure of model fit, RMSEA = root mean square error of approximation, SDC = smallest detectable change, SRMR = standardised root mean residuals, TLI = Tucker-Lewis index, VISA-G = Victorian Institute of Sport Assessment - Gluteal Tendinopathy
Table adapted from the Consensus-based Standards for the selection of health Measurement Instruments manual for systematic reviews of PROMs (172)

- Imprecision - rated down one level if pooled sample was <100 and two levels if <50. Imprecision was not considered when assessing content validity (172).
- Indirectness (e.g. if studies included in the review were partly performed in participants with another condition (172).

We will consider if different language versions of questionnaires should be treated separately or as a body of evidence (203). We will consider them as a body of evidence if questionnaires are sufficiently similar (number of items, response options) and results and are consistent between versions (no evidence to suggest differential item functioning). As a secondary analysis we have reported findings considering language versions separately (Appendix Q).

8.4.18 Formulating recommendations

Recommendations were based on the COSMIN guidelines: i) if an outcome measure had evidence for sufficient content validity (any level) and at least low-quality evidence for sufficient internal consistency it could be recommended for use; ii) if there was high-quality evidence for an insufficient measurement property the outcome measure it could not be recommended for use (172). If an outcome measure did not fit into i) or ii) the outcome with the best evidence for content validity is provisionally recommended (until further evidence is provided) (172).

8.5 Results

8.5.1 Outcome measures used in gluteal tendinopathy (Stage One)

We identified 8414 articles, of which 3029 were duplicates (Figure 8.1a). After title and abstract screening full texts (n=315) 123 studies met inclusion criteria. Fifty-seven outcome measures were mapped to one of the nine core health domains (Appendix O). The domain disability was the most common category to which outcome measures were mapped (26%), followed by quality of life (11%) and patient rating of condition (9%) (Appendix R). Many outcome measures identified could not be mapped to any of the core health domains (36%), including biomechanical studies or pain ratings of a condition that did not relate to a specific activity (Appendix R).

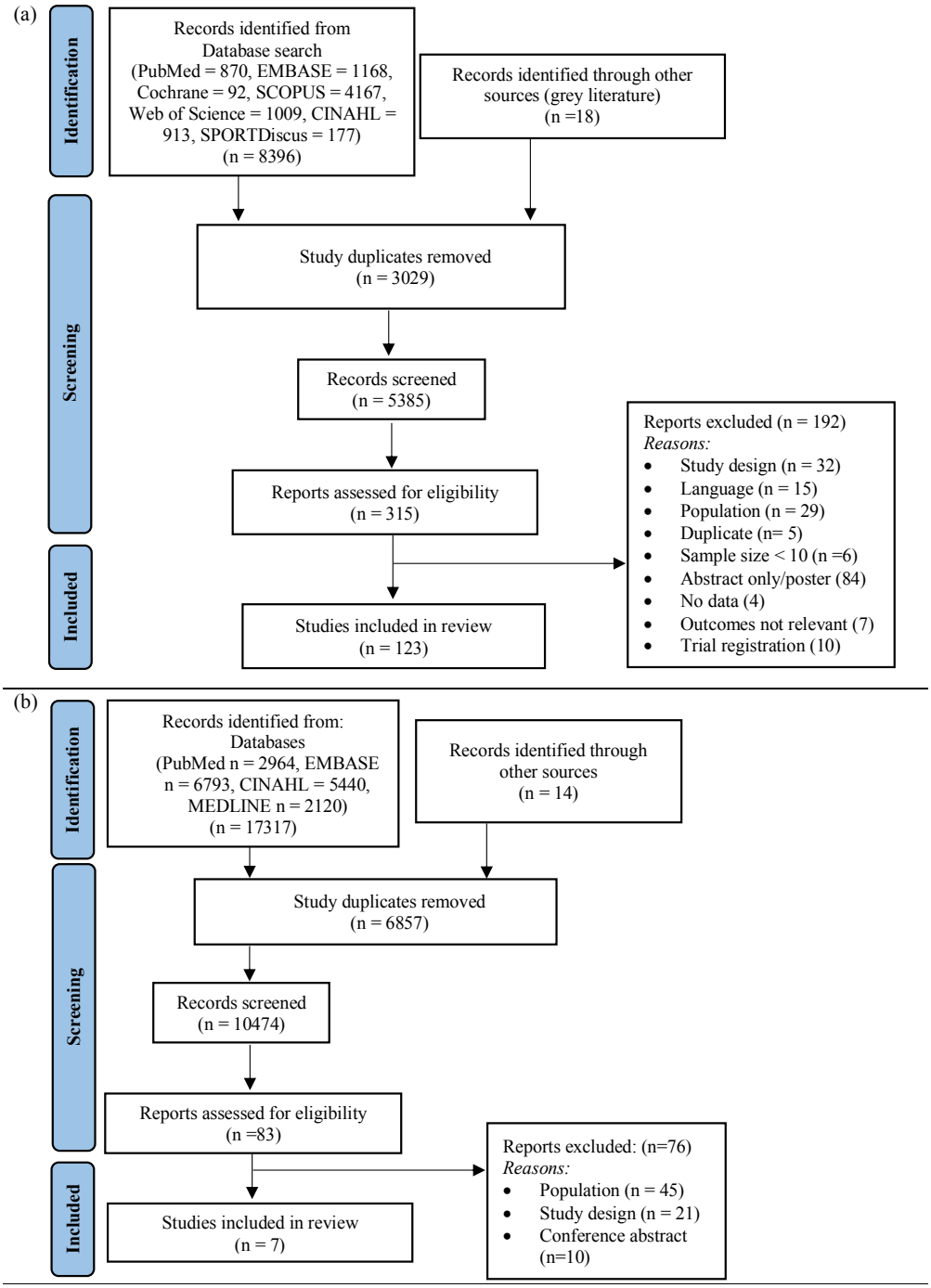


Figure 8.1 Prisma flow diagram

Figure (a) shows the process taken to identify all articles examining outcome measures in patients with gluteal tendinopathy. Figure (b) shows the process taken to identify all articles evaluating measurement properties in patients with gluteal tendinopathy.

GT = gluteal tendinopathy

8.5.2 Study selection measurement properties (Stage 2)

We identified 17331 studies in our search, of which 6857 were duplicates (Figure 8.1b). After title and abstract screening full texts of 83 articles were retrieved, of which seven met inclusion criteria.

8.5.3 Study characteristics

Two outcome measures were identified: the Victorian Institute of Sport Assessment – Gluteal Tendinopathy (VISA-G) and the Hip Outcome Score-Brazil (HOS-Brazil) (Table 8.2). There were six studies on the VISA-G (109, 204-208). One study described the development and validation of the VISA-G (109). Four studies translated the VISA-G into other languages (Danish, French, Italian and Brazilian Portuguese) (204, 206-208) and reported the measurement properties, and one additional study examined the responsiveness of the VISA-G (205). The studies on the VISA-G included a total of 430 patients with gluteal tendinopathy. Of these, 167 patients took part in the original validation of the VISA-G (109). The mean age of the participants in included VISA-G studies varied from 53 (median) to 66 years for patients with gluteal tendinopathy and 50 to 57 years for asymptomatic individuals. One study reported on measurement properties of the HOS-Brazil Activities of Daily Living (ADL) and Sport subscales, which included 70 participants, of which 44 had gluteal tendinopathy (26 had femoroacetabular impingement) and had a mean age of 43 (209).

8.5.4 Content validity

The VISA-G was developed for patients with gluteal tendinopathy and featured interviews of patients to assist with concept elicitation, and pilot tests for question design and comprehensibility (109). It was unclear if patients were asked about the comprehensiveness of the final version (109). The methodological quality of the original study was doubtful (109). Reviewer ratings were sufficient for all items on relevance, comprehensiveness and comprehensibility of the VISA-G. The Danish, French and Italian versions reported on the comprehensibility of the tool but did not provide additional evidence on comprehensiveness or relevance (204, 206, 207). All were of doubtful quality (204, 206, 207). Overall, there was low-quality evidence that content validity was sufficient for comprehensibility, and very low-quality evidence that the comprehensiveness and relevance were sufficient (Table 8.3).

Table 8.2 Characteristics of included studies

Source	Country, Language	N (females), GT, control/other	PROM tested	Mean age (years)	Recruited from	Inclusion criteria	Clinimetric properties										
							1	2	3	4	5	6	7	8	9	10	11
Beaudart (2020)	Liege, Belgium and France, French	106(65), GT: 52, asymptomatic: 54	VISA-G.F	53 (median age)	Community	Medical doctor confirmed GT (history, clinical examination and echography)				✓		✓	✓	✓		✓	✓
Ebert (2019)	Australia, English	56(52), GT:56	VISA-G	65.8	Orthopaedic outpatient clinic	Symptomatic (≥ 6 months) partial or full thickness tears of gluteus minimus, along with the anterior portion of gluteus medius, diagnosed via MRI. All patients had failed non-operative treatment including corticosteroid injections and physical therapy									✓	✓	✓
Fearon (2015)	Regional and urban South Eastern Australia, English	Item generation: 42, NR Pilot testing: 73, NR Clinimetric evaluation: 83 (71), GT: 52(47), asymptomatic: 31 (24)	VISA-G	GT: 58.9 (13.64 SD) asymptomatic: 57.4 (SD 5.59)[1]	Community	Item generation phase: consensus between 2 experienced clinicians as having trochanteric bursitis and/or GT Subsequent stage: history of lateral hip pain + pain on palpation of greater trochanter, and one or both of lateral hip pain with lying on the ipsilateral side, during weight bearing activities or sitting	✓	✓	✓	✓		✓	✓	✓		✓	✓
Jorgensen (2020)	Aalborg (Denmark), Danish	107(88), GT: 49 (47), asymptomatic: 58 (41)	VISA-G.DK	GT: 56, asymptomatic: 50	Community	Hip pain in weight-bearing, inability to lie on affected side				✓		✓	✓			✓	✓

Minetto (2020)	Italy, Italian	76(58), GT: 38 (29), asymptomatic: 38 (29)	VISA-G.I	GT: 64.5, asymptomatic: 56.5	NR	Lateral hip pain aggravated with activity and side lying position for > 3 months + clinical diagnosis of GT by medical doctor + confirmed by imaging (US)						✓	✓	✓		✓	✓
Paiva (2021)	Brazil, Portuguese	68(64), GT:68	VISA-G.BR	57.8	Community private practice	Between 18 and 75 years of age, pain in lateral region of hip, tenderness peri-trochanteric region, MRI showing gluteus medius and/or minimus tendinopathy				✓	✓	✓		✓	✓	✓	✓
Costa (2018)	Brazil, Portuguese	70(46), GT: 44, FAI: 26	HOS-Brazil (ADL) and (Sport)	42.9[2] (range 19-70)	Orthopaedic outpatient clinic	Literate and physically active with hip pain (FAI or GT as confirmed by radiograph, tomography, or MRI)			✓	✓		✓	✓	✓		✓	✓

ADL = activities of daily living, FAI = femoroacetabular impingement, GT = gluteal tendinopathy, HOS = Hip Outcome Score, MRI = magnetic resonance imaging, US = ultrasound, VISA-G = Victorian Institute of Sport Assessment - Gluteal Tendinopathy, VISA-G.BR = Victorian Institute of Sport Assessment - Gluteal Tendinopathy, VISA-G.DK = Victorian Institute of Sport – Gluteal Tendinopathy Danish, VISA-G.F = Victorian Institute of Sport – Gluteal Tendinopathy France, VISA-G.I = Victorian Institute of Sport Assessment – Gluteal Tendinopathy Italian

[1] Data based on mean age for clinimetric testing only. No data reported on item generation or pilot testing.

[2] Data not reported on FAI and gluteal tendinopathy patients separately

Studies received a ✓ if any data was reported on either interpretability of feasibility of the outcome measure

1 = PROM development, 2 = Content validity, 3 = Structural validity, 4 = Internal consistency, 5 = Cross-cultural validity\measurement invariance, 6 = Reliability, 7 = Measurement error, 8 = Hypotheses testing for construct validity, 9 = Responsiveness, 10 = Interpretability, 11 = Feasibility

Table 8.3 Summary of findings

Questionnaire	ROB	Inconsistency	Imprecision	Indirectness	QUALITY OF EVIDENCE (High, moderate, low, very low)	OVERALL RATING + / - / ?
VISA-G						
Content validity - relevance	Extremely serious	No	NA	No	Very Low	+
Content validity - comprehensiveness	Extremely serious	No	NA	No	Very Low	+
Content validity - comprehensibility	Serious	No	NA	Serious[3]	Low	+
Structural validity	Very serious	NA[1]	Serious[2]	No	?	?
Internal consistency	Serious	No	No	Serious[3]	?	?
Cross-cultural validity	NA	NA	NA	NA	?	?
Reliability	Serious	No	No	Serious[3]	Low	+
Measurement error	Serious	No	No	Serious[3]	Low	+
Construct validity (comparison with outcome measures)	No	Serious	No	Serious[3]	Low	-
Construct validity (known group)	No	No	No	Serious[3]	Moderate	+
Responsiveness (before and after intervention)	No	No	No	Serious[4]	Moderate	+
Responsiveness (comparison with other outcome measures)	Serious	NA[1]	Serious[2]	Serious[4]	Very Low	+
HOS-Brazil (ADL)						
Content validity - relevance	Extremely serious	No	NA	Serious[4]	Very Low	+
Content validity - comprehensiveness	Extremely serious	No	NA	Serious[4]	Very Low	-
Content validity - comprehensibility	Extremely serious	No	NA	Serious[4]	Very Low	+
Structural validity	NA	NA	NA	NA	?	?
Internal consistency	Very serious	No	Serious[2]	Serious[4]	?	?

Cross-cultural validity	NA	NA	NA	NA	?	?
Measurement error	Serious	No	Serious[2]	Serious[4]	Very Low	-
Reliability (test-retest)	Serious	No	Serious[2]	Serious[4]	Very Low	+
Construct validity (comparison with other outcome measures)	Serious	No	Serious[2]	Serious[4]	Very Low	-
Responsiveness	NA	NA	NA	NA	?	?
HOS-Brazil (Sport)						
Content validity - relevance	Extremely serious	No	NA	Serious[4]	Very Low	+
Content validity - comprehensiveness	Extremely serious	No	NA	Serious[4]	Very Low	-
Content validity - comprehensibility	Extremely serious	No	NA	Serious[4]	Very Low	+/-
tructural validity	NA	NA	NA	NA	?	?
Internal consistency	Very serious	No	Serious[2]	Serious[4]	?	?
Cross-cultural validity	NA	NA	NA	NA	?	?
Measurement error	Serious	No	Serious[2]	Serious[4]	Very Low	+
Reliability (test-retest)	Serious	No	Serious[2]	Serious[4]	Very Low	+
Construct validity (hypothesis testing)	Serious	No	Serious[2]	Serious[4]	Very Low	-
Responsiveness	NA	NA	NA	NA	?	?

Quality level:

- High –very confident that the true measurement property lies close to that of the estimate of the measurement property
- Moderate –moderately confident in the measurement property estimate: the true measurement property is likely to be close to the estimate of the measurement property, but there is a possibility that it is substantially different
- Low – confidence in the measurement property estimate is limited: the true measurement property may be substantially different from the estimate of the measurement property
- Very low - very little confidence in the measurement property estimate: the true measurement property is likely to be substantially different from the estimate of the measurement property
- NA = no information available

Overall rating:

- + = sufficient, - = insufficient, +/- = inconsistent, ? = indeterminate

HOS-Brazil (ADL) = Hip Outcome Score (activities of daily living), NA = not applicable, VISA-G = Victorian Institute of Sport Assessment – Gluteal Tendinopathy

[1]Inconsistency could not be evaluated because there was only one study available
[2]Imprecision downgraded due to sample size
[3]Indirectness downgraded due to pooling of multiple language versions of the VISA-G
[4]Indirectness downgraded due to varied population

Table 8.4 Measurement properties

PROM, ref (year)	Structural validity			Internal consistency			Cross-cultural validity/measurement invariance			Reliability (test-retest)			Measurement error			Hypotheses testing for construct validity			Responsiveness		
	n	Meth qual	Rating	n	Meth qual	Result (rating)	n	Meth qual	Result (rating)	n	Meth qual	Result (rating)	n	Meth qual	Result (rating)	n	Meth qual	Result (rating)	n	Meth qual	Result (rating)
VISA-G.F, Beaudart (2020)	NT	NT	NT	52	D	0.81 (+)	NT	NT	NT	106 [1]	I	0.99 (+)	52	I	SDC: 4.55 (+)	106	A[4] V[5]	Results in line with 5(5+) hypo[4] Results not in line with 3 hypo (3-)[4] Results in line with 1 hypo (+1)[5]	NT	NT	NT
VISA-G, Ebert (2019)	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	56	V[6] A[7]	AUC using ≥4 GRC): 0.70 (95% CI 0.56-0.81)[6] Results in line with 2 (2+) hypo[7]
VISA-G, Fearon (2015)	52	D	?	52	D	0.52 (-)	NT	NT	NT	26	D	0.83 (+)	26	D	SDC: 5.21 (+)	83	A[4] V[5]	Results not in line with 2(2-) hypo[4] Results in line with 1(+1) hypo[5]	NT	NT	NT
VISA-G.DK, Jorgensen (2020)	NT	NT	NT	49	D	0.98 (+)	NT	NT	NT	49	D	0.96 (+)	107 [2]	D	SDC: 3.17 (+)	NT	NT	NT	NT	NT	NT
VISA-G.I, Minetto (2020)	NT	NT	NT	38	D	0.79 (+)	NT	NT	NT	38	D	0.91 (+)	38	D	SDC: 11.40 (+)	76	A[4, 5]	Results not in line with 1(-1) hypo[4] Results in line with 1(+1) hypo[5]	NT	NT	NT
VISA-G.BR Paiva (2021)	NT	NT	NT	68	D	0.65 (-)	NT	NT	NT	68	D	0.91 (+)	68	D	SDC: 11.60 (+)	68	A[4]	Results not in line with 1(-1) hypo[4]	68	I	ES = 0.19[6]
HOS-Brazil ADL,	NT	NT	NT	70	D	0.95 (+)	NT	NT	NT	70	D	0.99 (+)	70	D	SDC: 4.71 (-)	70	A[4]	Results in line with 2(2+) hypo[4]	NT	NT	NT

Costa (2019)[3]																		Results not in line with 1(-1) hypo[4]			
HOS-Brazil Sport, Costa (2019)[3]	NT	NT	NT	70	D	0.92 (+)	NT	NT	NT	70	D	0.99 (+)	70	D	SDC: 5.27 (+)	70	A[4]	Results in line with 2(+2) hypo[5] Results not in line with 1(-1) hypo[5]	NT	NT	NT

AUC = area under curve, CI = confidence interval, ES = effect size, GRC = global rating of change, hypo = hypothesis, meth qual = methodological quality (V = very good, A = adequate, D = doubtful, I = inadequate), n = sample size, NT = measure property not tested in study, ref = reference, VISA-G = Victorian Institute of Sport Assessment - Gluteal Tendinopathy, VISA-G.BR = Victorian Institute of Sport Assessment - Gluteal Tendinopathy, VISA-G.DK = Victorian Institute of Sport – Gluteal Tendinopathy Danish, VISA-G.F = Victorian Institute of Sport – Gluteal Tendinopathy France, VISA-G.I = Victorian Institute of Sport Assessment – Gluteal Tendinopathy Italian
Result rating (against updated criteria for good measurement properties): + = sufficient, - = insufficient, ? = indeterminate
Rating of measurement error reported as smallest detectable change (SDC)

[1]Reliability value calculated included symptomatic (n =52) and asymptomatic individuals (n =54)

[2]Measurement error value calculated included symptomatic (n=49) and asymptomatic individuals (n=58)

[3] Part of cohort had femoroacetabular impingement (n =26)

[4]Construct validity: comparison with other outcome measures

[5]Construct validity: known group validity

[6]Responsiveness: before and after an intervention

[7]Responsiveness: comparison with other outcome measures

The HOS (Sport and ADL subscales) was originally developed for patients with acetabular labral tears (196). HOS (ADL) provided very low-quality evidence of sufficient relevance and comprehensibility, and insufficient comprehensiveness. The HOS (Sport) provided very low-level evidence of sufficient relevance, inconsistent comprehensibility and insufficient comprehensiveness.

8.5.5 Structural validity

A single study of doubtful quality reported on the structural validity of the VISA-G (Table 8.4) (109). The VISA-G is stated to measure the degree of disability caused by gluteal tendinopathy, which is calculated using a total score (109). This infers that VISA-G measures one construct (disability) (109). In the development paper, however, it suggests that the questionnaire examines more than one domain (disability and activity level) (109), and hence would not conform with uni-dimensionality (172). While factor analysis was performed in the development study there was insufficient data to rate the structural validity of the VISA-G, and it was therefore rated indeterminate. The structural validity of the HOS-Brazil (ADL and Sport) subscales was not investigated and was rated indeterminate (Table 8.3).

8.5.6 Internal consistency

The internal consistency of the VISA-G and the HOS-Brazil (ADL and Sport) was rated indeterminate as COSMIN criteria for at least low-level evidence for sufficient structural validity was not met.

8.5.7 Cross-cultural validity/measurement invariance

None of the translations of the VISA-G (204, 206-208) or HOS-Brazil (ADL) or (Sport) subscales(209) underwent cross-cultural validation, which requires the assessment of measurement variance across different populations (172). Subsequently, the cross-cultural validity of the Danish, French, Italian and Brazilian Portuguese versions of the VISA-G and the HOS-Brazil (Sport and ADL) was rated indeterminate.

8.5.8 Test-retest reliability

Test-retest reliability was reported in the VISA-G, in four studies of doubtful quality (109, 204, 206, 208) and one study of inadequate quality (207). The intraclass correlation coefficient of the VISA-G varied from 0.83-0.99 (Table 8.3). Overall, there was low-quality evidence of

sufficient test-retest reliability of the VISA-G, with the evidence downgraded one level for risk of bias and indirectness (different language versions).

The test-retest reliability was reported in a single study and found to be 0.99 (95% CI 0.99 to 1.00) for the HOS-Brazil(ADL) and 0.99 (95% CI 0.99 to 1.00) for the HOS-Brazil(Sport) (209). The HOS-Brazil subscales (ADL and Sport) were rated down due to indirectness, as the study included patients without gluteal tendinopathy (37% had femoroacetabular impingement), and one level for both risk of bias and imprecision (n <100) (209). Overall, there was very low-quality evidence of sufficient test-retest reliability of the HOS-Brazil (ADL) and the HOS-Brazil (Sport) (Table 8.4).

8.5.9 Measurement error

The smallest detectable change was reported in four studies of doubtful quality (109, 204, 206, 208) and one of inadequate quality (207) and varied from 3.17-11.6. One study reported the minimal important change in people 12 months after gluteal surgery using the VISA-G questionnaire (205). Using a global rating of change of ≥ 4 (i.e. moderately better, on an 11-point scale where -5 indicated very much worse and +5 completely recovered) as an anchor, the minimal important change was 29/100 (205). Using the global rating of change of ≥3 as an anchor (somewhat better at 12 months post gluteal tendon surgery), the minimal important change was 22/100 (205). Overall, there was low-quality evidence of sufficient measurement error of the VISA-G, with the evidence downgraded one level for both risk of bias and indirectness (different language versions).

One study reported the minimal detectable change of the HOS-Brazil (ADL) and HOS-Brazil (Sport), which were 4.71 and 5.27 respectively (Table 8.3) (209). The minimal important change of the HOS-Brazil (ADL) and (Sport) in patients with gluteal tendinopathy was 4.6 and 5.5 respectively (209). Overall, there was very low-quality evidence that the measurement error of the HOS-Brazil (ADL) was insufficient and very low-level quality evidence that the measurement error in the HOS-Brazil (Sport) was sufficient.

8.5.10 Hypothesis testing for construct validity

Comparison of the VISA-G with other outcome measures was assessed in four studies of adequate quality (109, 206-208). When the results were summarised five results were found

to be in line with the hypothesis and seven results were not in line with the hypothesis (Appendix S). Overall, there was low-quality evidence of insufficient construct validity (comparison with other measures) (Table 8.3). Two very good (109, 207) and one adequate quality study (206) assessed known group validity of the VISA-G by comparing the results of the VISA-G between asymptomatic people and those with gluteal tendinopathy. When results were summarised three (out of three) results were found to be in line with the hypothesis and overall provided moderate-quality evidence of sufficient construct validity (known group).

Comparison of the HOS-Brazil (ADL and Sport) with other outcome measures was assessed in one study of adequate quality (209). In the HOS-Brazil (ADL) two results were found to be in line with the hypothesis and one result was not in line with the hypothesis (Appendix S) (209). Overall, there was very low-quality evidence of insufficient construct validity (convergent). In the HOS-Brazil (Sport) two results were found to be in line with the hypothesis and one result was not in line with the hypothesis (Appendix S) (209). Overall, there was very low-quality evidence of insufficient construct validity (convergent).

8.5.11 Responsiveness

Responsiveness was reported in two studies on the VISA-G (205, 208). One study of very good quality reported the responsiveness of VISA-G pre and post gluteal tendon surgery (hip abductor tendon repair) (205). In this surgical study, patients included had symptomatic (minimum six months) partial or full-thickness tears of gluteus minimus and the anterior portion of gluteus medius and had failed non-surgical management (205). A second study of inadequate quality reported on the responsiveness of the VISA-G by testing before and after an uncontrolled treatment for 30 days (208). Overall, there was moderate-quality evidence of sufficient responsiveness (before and after an intervention). One study of adequate-quality provided very low-quality evidence of sufficient responsiveness from the comparison of the VISA-G with other outcome measures (Harris Hip Score, modified Harris Hip Score and Oxford Hip Score) (205). The evidence was downgraded for risk of bias, imprecision (sample size) and indirectness (varied population) (205). There were no studies evaluating the responsiveness of the HOS-Brazil subscales in patients with gluteal tendinopathy.

8.5.12 Interpretability and feasibility

The results for interpretability and feasibility can be found in Appendix T and U respectively. The mean baseline score for VISA-G ranged from 43 (205) to 63/100 (204). No floor or ceiling effects were found in the VISA-G (109) or the Brazilian Portuguese (208), French (207) or Danish (204) language versions. One study reported floor effects were present in items two and seven and ceiling effects were present in items two and six of the Italian version of the VISA-G (206).

A single study reported on the interpretability of the HOS-Brazil (209). The range of baseline scores in this study for the HOS-Brazil (ADL) varied 15.8-65.1 and the HOS-Brazil (Sport) from 2.8-97.2 (209). The questionnaire did not demonstrate floor or ceiling effects (209).

8.6 Discussion

The aim was to evaluate the clinometric properties of outcome measures that have been reported for use in people with gluteal tendinopathy. Clinimetric properties were reported for the VISA-G and the HOS-Brazil (ADL and Sport subscales), both of which lie within a single core health domain (disability). When evaluated as a body of evidence or individually (Appendix Q) – and consistent with the primary findings, there was insufficient evidence to recommend any outcome measure. Future research must investigate clinimetric properties of outcome measures that report on other core health domains or develop new outcome measures, which encompass multiple or all core health domains. In the meantime, the VISA-G may be used to assess the disability associated with gluteal tendinopathy, as they appear better suited from the available evidence.

An outcome measure must cover all key concepts of the construct it is designed to measure for the condition it is used (comprehensiveness) (179). Most patient-reported outcomes used to study gluteal tendinopathy were developed and validated for intra-articular hip conditions, such as the Harris Hip Scale and HOS (ADL and Sport) (196). Results from this study demonstrated that the HOS(ADL and Sport) miss key concepts for patients with gluteal tendinopathy, including activities such as lying on the side at night (HOS-ADL), or running uphill (HOS-Sport) (48, 100). This stresses the importance of patient involvement in the development and validation of measures for use in gluteal tendinopathy prior to use in research and clinical practice. The only measure designed specifically for gluteal tendinopathy was the VISA-G.

The VISA-G, like the VISA-A (Achilles) and VISA-H (Hamstring), was adapted from the original VISA-P (Patella), which was published over two decades ago in 1998 (109, 110, 210, 211). The VISA-P was designed for assessing disability (amongst other constructs) associated with patella tendinopathy (210), a condition that most commonly occurs in young athletes that participate in sport (59). Gluteal tendinopathy is most prevalent in older individuals who are less likely to be participating in sport (212). A challenge for patient-reported outcome measures assessing gluteal tendinopathy is to appropriately score all sub-groups of the condition (e.g. sedentary and sporting groups).

The suitability of two items in the VISA questionnaires for less active populations has come under scrutiny, due to the focus on the performance of sporting activities (items seven and eight) (193, 194, 198). For example, in item eight of VISA-G respondents are asked to answer one of three sections that correspond with their ability to perform activities for certain durations, and consider the influence on pain in the performance of the activity (109). The sporting focus present in the other VISA questionnaires (Achilles, Hamstring and Patella) was amended somewhat in the development of the VISA-G to allow respondents to consider less active weight-bearing activities (e.g. walking or shopping) (109). However, to achieve the maximum score, respondents are still required to be able to perform greater than 30 minutes of such activities (109). Therefore, whilst the VISA-G is not as sport-focused as the other VISA questionnaires, due to the weighting of item eight (30 out of 100 points) it is currently unknown if the point allocations for different activities in this item reflect disease status or participants ideal/normal level of physical activity engagement. For example, a patient who does not wish to undertake any exercise will score zero points on item seven and at best without any pain score six points on item eight - and would be penalised approximately 34 points unrelated to their disease status.

A subscale or scale (when calculated as a total score) should only measure a single construct (uni-dimensionality). Where multiple items measure different constructs, they breach uni-dimensionality. In such cases, items measuring the same construct should be calculated separately (e.g. HOS-ADL and HOS-sport), and the use of a total score avoided. In this study the structural validity of the VISA-G was indeterminate. The VISA-G may be found to cover multiple domains (e.g. disability and activity level) (109), so should perhaps include sub-scores, rather than a total score. Similar recommendations have been proposed for the VISA-

A (213). Currently, the VISA-G is reported as one score. As tendinopathy does not affect a single heterogenous group, amendment of the items measuring the activity level construct as an optional subscale that is scored separately (e.g. sport subscale) may be appropriate – this would necessitate revisiting domain mapping of the VISA-G.

Multi-item outcome measures can be time-consuming and are often not suitable for daily or weekly use (214). Single item measures are often favoured in clinical practice and have been used frequently in studies on gluteal tendinopathy, but we found no clinimetrics on such measures. Limitations of single-item measures include the inability to calculate internal consistency and susceptibility to random measurement errors (214). Single-item measures may be acceptable if an outcome is unambiguous and sufficiently unidimensional, where respondents consider all aspects of the construct (215). Some ICON core health domains lend themselves for use as single-item measures, for example - pain on activity/loading. A qualitative study found expert clinicians favoured single-item question-based monitoring (e.g. patient-specific functional scale) or simple load monitoring tests (e.g. numerical rating scale of pain during a load-based test) for the above reasons, as well as that they felt they were responsive enough to allow self-pacing of rehabilitation by the patient (e.g. regression or progression) (161). Whilst we found single-item question-based outcome measures are frequently used in this population, they require validation to determine their validity, reliability and responsiveness in patients with gluteal tendinopathy.

There are several limitations in interpreting the results of this systematic review. The results for the measurement error of VISA-G should be interpreted with caution as the minimal important change used was established in post-operative patients with gluteal tendinopathy (205). This may not be generalisable to other patients with gluteal tendinopathy as it likely represented more severe cases (mean VISA-G 43/100 at baseline) as all patients had partial or full-thickness tears of the gluteus minimus and anterior portion of the gluteus medius and had failed non-operative treatment (205). In study selection (screening) only studies written in English were considered– which was due to limited resources. This introduces the possibility of language bias. We did not screen all retrieved studies of musculoskeletal hip pain for studies that had sub-groups of gluteal tendinopathy – introducing a risk of selection bias. To counter this, we randomly retrieved studies – finding none had subgroups of gluteal tendinopathy. The scoping review was limited to studies on treatment outcomes where it is the most important to

understand the measurement properties of outcome measures. A limitation of this decision is that we may have missed an outcome measure that is valid, responsive and reliable in measuring outcomes. Whilst the COSMIN is the accepted standard for evaluating clinometric studies, there are several limitations worth mentioning, including a lack of conditions to upgrade the quality of evidence, sample size requirements employed when scoring is based on rule of thumb and the reliability of the tool requires further evaluation (172).

To date, no outcome measure demonstrates sufficient measurement properties to be recommended for use in a core outcome set for gluteal tendinopathy. Until further evidence is provided, we provisionally recommend clinicians use the VISA-G. The HOS-Brazil (ADL) and (Sport) subscales have the potential to be recommended for use but require further research to assess their clinimetric qualities. We encourage clinicians and researchers to use outcome measures judiciously by considering those measures that have the best measurement properties in their specific clinical or research application.

8.7 Conclusion

This review found the VISA-G had a higher level of evidence than the HOS-Brazil (ADL and Sport) in the assessment of gluteal tendinopathy – although it has several clinimetric issues that require addressing. None of the other outcome measures employed in clinical trials can currently be recommended as there is insufficient data to support their use.

Conflict of interest: One of the authors published the original development article of the VISA-G. This author was not involved in the quality assessment, analysis of the data or writing of results regarding this outcome.

8.8 Summary of findings

This systematic review evaluated current evidence regarding the quality of outcome measures used to evaluate treatments for gluteal tendinopathy. This systematic review identified a critical lack of outcome measures with robust measurement properties for use in trials on gluteal tendinopathy. As per the COSMIN methodology, in the absence of further evidence, the outcome measure in a domain with the best content validity is recommended. Consequently, we were able to recommend the VISA-G to capture the disability construct until future research is performed.

9 Chapter 9: Grand discussion and conclusion

9.1 Preface

This thesis aimed to: i) summarise evidence on the effect of interventions in managing proximal hamstring tendinopathy ii) identify how experts currently diagnose and manage the condition and iii) appraise outcome measures that are used to evaluate the impact of two common tendinopathies around the hip.

Chapter one presented an overview of the literature on proximal hamstring tendinopathy and identified existing gaps in the literature that could be addressed in subsequent studies. Chapter two (study one) systematically summarised the efficacy of interventions in managing proximal hamstring tendinopathy. Chapter three (study two) explored expert views on the assessment, management and prevention of proximal hamstring tendinopathy. Chapter four summarised results from the 2018 International Tendinopathy Consensus (ICON), which had direct implications for the remainder of the thesis (80, 164). Chapter five (study three) outlined a protocol to develop a core outcome set for proximal hamstring tendinopathy. Chapter six (study four) mapped outcome measures to ICON core health domains used in studies on proximal hamstring tendinopathy and evaluated the clinimetric properties of outcome measures used in studies on proximal hamstring tendinopathy. Chapter seven (study five) surveyed an international multidisciplinary collaboration to determine the truth and practicality of outcome measures used in proximal hamstring tendinopathy. Chapter eight (study six) gathered outcome measures used in studies evaluating gluteal tendinopathy, mapped outcome measures to core health-related tendinopathy domains and systematically evaluated the clinimetric properties of each outcome measure. This Chapter (Chapter nine) summarises the findings and implications of this thesis.

9.2 Key findings of the thesis

This thesis has several important and novel findings, which are discussed below and summarised in Table 9.1.

9.2.1 There is limited evidence on the effect of interventions

Evidence-based practice incorporates i) therapists' clinical expertise ii) patients' values and iii) research evidence (216). There was no high or moderate-quality research found to guide treatment decisions for proximal hamstring tendinopathy (Chapter two, study one), meaning the confidence in the results of the evidence is low (155). In the absence of direct research evidence, instead of simply relying on the other two pillars of evidence-based practice (clinical expertise and patient values) clinicians can adopt evidence from other tendinopathy research where higher-quality studies have been completed (e.g. gluteal, Achilles and patella tendinopathy) (39, 107, 147). This includes education regarding load management and exercise-based rehabilitation (e.g. heavy slow loading).

The approach taken to rehabilitate proximal hamstring tendinopathy by experts (Chapter three, study two) differed from what had been identified in our systematic review (Chapter two, study one) (155). For example, most studies in the systematic review (11/12) involved passive interventions such as injection therapies, surgery or shockwave therapy, whereas expert physiotherapists primarily manage the condition using education and progressive exercise that targets the hamstring unit and kinetic chain (155, 161).

9.2.2 Core outcome sets are needed

The need for a core outcome set was proposed by researchers involved in the ICON consensus (164). This need was also recognised by the inability to draw meaningful conclusions from intervention studies in the systematic review in Chapter two (study one) due to the inconsistencies in the use of outcome measures across studies (155). To combat this issue Chapter five (study three) outlined a rigorous method to develop a core outcome set for proximal hamstring tendinopathy (176).

9.2.3 High-quality outcome measures are needed

Results in Chapters six (study four) and eight (study six), which evaluated the measurement

Table 9.1 Key findings of the thesis

Study	Title	Key findings
Study 1	Proximal hamstring tendinopathy; a systematic review of interventions.	• There is insufficient evidence to make strong recommendations on management of proximal hamstring tendinopathy. • Surgery and shockwave therapy delivered promising results on physical function and self-reported symptoms, however confidence in the results was low. • No studies provided adequate details of post-surgical rehabilitation protocols. • Patient selection criteria, outcomes and reporting across studies must be standardised.
Study 2	Proximal hamstring tendinopathy; expert physiotherapists' perspectives on diagnosis, management and prevention.	• Experts use findings from the patient interview (onset, pain behavior, location of pain and aggravating factors) and a battery of pain provocation tests to diagnose proximal hamstring tendinopathy – implying no single test is adequate. • Education about proximal hamstring tendinopathy, and how to manage load is essential to ensure athletes self-manage and adhere to the prolonged rehabilitation required. • Exercise prescription should minimise early exposure to provocative ranges of hip flexion. • Exercise that includes progressive loading of the hamstring musculotendinous unit and synergists should be the first line of care for proximal hamstring tendinopathy. • The utility of passive therapies (e.g. injection therapies, surgery and manual therapy) was thought to be limited.
Study 3	Core outcome set development for proximal hamstring tendinopathy (COS-PHT): a study protocol.	• Outcomes are used inconsistently in research on proximal hamstring tendinopathy and limited the ability for synthesis of results • It is critical that outcome measures selected are reliable, valid and responsive to change and cover the range of health domains acknowledged in the ICON consensus process. • This protocol describes the approach to initiate the development of a Core Outcome Set – Proximal Hamstring Tendinopathy (COS-PHT).
Study 4	Outcome measures in the management of proximal hamstring tendinopathy: a systematic review of their measurement properties.	• Published (or current) outcome measures used in studies of proximal hamstring tendinopathy do not have sufficient measurement properties when evaluated against the COSMIN criteria. • Whilst no outcome measure used in studies of proximal hamstring tendinopathy demonstrated sufficient measurement properties to be recommended for inclusion in a core outcome set for proximal hamstring tendinopathy, the VISA-H (and translations) is currently more suitable than other measures of disability and should be used until further research is conducted.
Study 5	Core outcome set for proximal hamstring tendinopathy (COS-PHT); a survey of an international collaboration.	• Four outcome measures met consensus for inclusion for both truth and feasibility: two outcome measures in participation in life activities domain (return to sport pre-activity level, return to previous level of activity) and one outcome measure in each of the patient rating of condition (global rating of change) and disability (VISA-H) domains. • Two outcome measures met criteria for exclusion: return to sport (level not defined) and the Modified Harris Hip Scale. • No outcome measures used previously in research for PHT met consensus for pain on activity /loading, function, psychological factors, physical function, quality of life, or pain over a specified period of time. • A key theme was that measures used to assess intra-articular pathology of the hip (e.g. hip osteoarthritis) did not meet the thresholds for 'truth' and therefore should not be used in research on proximal hamstring tendinopathy.
Study 6	Outcome measures in the management of gluteal tendinopathy: a systematic review of their measurement properties.	• Published (or current) outcome measures used in studies of gluteal tendinopathy do not adequately measure any of the ICON core health related domains of tendinopathy. • No outcome measure used in studies of gluteal tendinopathy demonstrated sufficient measurement properties to be recommended for use. • The VISA-G (and translations) is currently more suitable than other measures and should be used in the absence of other research • Standardised valid outcome measures are urgently required for studies of gluteal tendinopathy.

properties of outcome measures used in proximal hamstring and gluteal tendinopathy found a severe lack of validated measures. Both systematic reviews identified issues of concern with the VISA-G and VISA-H questionnaires. An area of particular concern was the limited evidence of uni-dimensionality, which is critical considering the VISA measures are calculated as total scores when used in clinical trials and clinical settings. As it stands, no outcome measure demonstrates sufficient measurement properties to be recommended for use in a core outcome set for proximal hamstring or gluteal tendinopathy. Until further evidence is provided, the VISA-G or the VISA-H are the best outcome measures available to evaluate disability for gluteal and proximal hamstring tendinopathy respectively. High-quality outcome measures must be developed or revised to accurately capture the impact of tendinopathy across all health domains.

Results in Chapter seven (study five) recognised four 'candidate' outcome measures eligible for progression as part of a core outcome set for proximal hamstring tendinopathy: return to sport (pre-activity level), return to previous level of activity, global rating of change and the VISA-H (disability domain). As it stands, none of these have evidence of adequate measurement properties to meet the requirements for recommendation (172, 179). Further research is required to understand the patient's views on these measures.

9.3 Clinical and research implications

The series of studies in this thesis has several important clinical and research implications for improving the management and evaluation of proximal hamstring tendinopathy and gluteal tendinopathy which follow below.

9.3.1 Inclusion criteria and diagnosis

Inconsistencies in inclusion criteria between studies or incomplete reporting make interpretations of study findings difficult (i.e. are we comparing oranges and apples or is this really an apple?). This was evident in the systematic review (Chapter two, study one) and limited the interpretation of the study findings. To improve reporting in research on tendinopathy a 'standard' for tendinopathy reporting was recently published which includes significant information on patient characteristics (80). Future research on proximal hamstring tendinopathy should follow the aforementioned guideline to allow better interpretation of study findings.

Consensus on the inclusion criteria for proximal hamstring tendinopathy is required. Chapter three (study two) provided key information on the methods experts in clinical practice use to assess proximal hamstring tendinopathy - through the patient interview (onset, pain behaviour, location of pain and aggravating factors) and physical examination, without the need for imaging (161). Reducing the need for unnecessary medical services is important due to the considerable expense of imaging and the use of health resources. Further research investigating the diagnostic utility of items from the patient interview and specific pain provocation tests used in the clinical examination outlined in this qualitative study (e.g. arabesque) will assist in forming a consensus on criteria for the diagnosis of proximal hamstring tendinopathy (161).

9.3.2 Treatment selection for proximal hamstring tendinopathy

There was little evidence to support any treatment investigated in proximal hamstring tendinopathy. This provides a challenge for clinicians managing this condition. The popularity of alternative treatments such as injection therapies and shockwave therapy has soared over the past decade (217, 218). Due to the considerable expense of these interventions (e.g. in purchasing a shockwave therapy machine or receiving platelet-rich plasma injections), their efficacy must be proven (137, 217). Particularly as it is reasonable to hypothesise that desperate patients, not responding to traditional management, are more likely to be open to trying less proven and more expensive treatments. Unfortunately, there was insufficient evidence to make confident conclusions about the effectiveness of any intervention, heralding the need for future trials on frequently utilised interventions such as platelet-rich plasma injection, surgery, rehabilitation, education and shockwave therapy.

9.3.3 Selecting outcome measures

The COSMIN initiative was established in 2005 with the altruistic endeavour of improving the quality of outcome measure development – this is critical for furthering the quality of randomised controlled trials, from which clinical guidelines are developed (180). Over the last 15 years, the recommended benchmark for outcome measures has increased to the point where those developed prior to the COSMIN initiative have little chance of having 'sufficient' measurement properties, due to the rigorous requirements (179, 180). The guidelines have become even more stringent at this time, with those applied in this study updated in 2018 (179). All outcome measures used in studies on proximal hamstring tendinopathy and gluteal

tendinopathy were developed before 2018, and most before the original COSMIN initiative in 2005 (guidelines published in 2007) (180).

There were several issues with outcome measures used to study proximal hamstring and gluteal tendinopathy. None more obvious than the use of outcome measures that were originally designed to assess intra-articular disease, which may have limited relevance in the context of tendinopathy. The specificity of the outcome measure used is important across musculoskeletal healthcare. As such, this finding is also relevant to studies on hamstring ruptures and full-thickness gluteal tendon tears, where the use of outcome measures designed for intra-articular hip disease is also common (219, 220). This is in line with a recent editorial in response to inconclusive results from a systematic review on hamstring avulsion injuries which concluded that the heterogeneity of outcome measures has 'hamstrung' research on hamstring avulsion injuries (220).

The VISA-H had minimal patient involvement in its development (informal interviews only). The COSMIN guidelines instruct that content validity should include an assessment of the relevance, comprehensiveness and comprehensibility following strict qualitative methodology (179). This is essential for adequate content validity (179). Research specifically on the content validity of the VISA-H is required to determine whether patients consider that the items included are relevant, comprehensible and comprehensive enough to properly measure the impact of the condition. Consistent with the findings of Chapter six (study four), a recent systematic review of all VISA measures used in tendinopathy highlighted the need for an update of measures used to evaluate lower limb tendinopathy (194, 198). Key reasons were: i) insufficient content validity (in particular due to a lack of patient involvement), ii) structural validity (lack of conformity to unidimensionality with evidence of assessment of multiple constructs) and iii) scoring system (limitations in use for in-active patients) (194, 198, 199).

Outcome measures must continue to evolve to better inform researchers of the clinical effects of interventions. Researchers need to develop, redesign or validate outcome measures to ensure that they have measurement properties that are adequately robust to withstand use in clinical trials and clinical practice. This must include patient consultation at each stage of development. Future research should aim to develop a core outcome set that includes a measure with robust properties for each of the nine core health domains (164).

9.4 Strengths and limitations of research design

A strength of this thesis was the variety of study designs used. Designs included a systematic review on interventions (using methods outlined in the Cochrane handbook), two systematic reviews on measurement properties of outcome measures applying COSMIN methodology, a qualitative study and finally a multi-staged study involving a scoping review and survey.

9.4.1 Systematic review

The systematic review of interventions for proximal hamstring tendinopathy (study one, Chapter two) provided clinicians with a comprehensive summary of findings of current evidence (155). It is the first systematic review on the effectiveness of interventions in managing proximal hamstring tendinopathy (155). This provided researchers, clinicians and patients with all current data on what is known on outcomes for different interventions and included important considerations such as adverse events and the likelihood of return to sport. The inclusion of the appraisal of bias in included studies, also allows clinicians to understand the level of confidence in the results. The systematic review was pre-registered (Prospero), followed PRISMA guidelines, and had two independent researchers complete the study screening, selection and bias assessment.

A further strength of this study was the inclusion of a variety of study designs (e.g. case series, cohort studies, randomised controlled trials). This allowed for a comprehensive review of all available data in an area that has so far received little attention from researchers. A limitation was that we were unable to determine if the effects are due to factors other than the treatment applied, such as natural recovery, in the studies that did not have a control group. A further limitation was that most included studies contained small sample sizes which affected the overall level of evidence, due to possible imprecision. We were also only able to include studies in English, due to limited funding for translators, introducing potential language bias. Finally, the lack of standardised inclusion criteria between studies and inconsistent use of outcome measures meant the pooling of results between studies was impossible.

9.4.2 Qualitative research

The advantage of using a qualitative design in Chapter three (study two) was that it enabled deep exploration of the complex clinical reasoning process used by experts. However, the findings of qualitative research cannot be extended broadly with the same degree of certainty

as quantitative research (221). Therefore, the results of this study provided a framework for future quantitative studies to explore. Another limitation of this study was that we only included a single profession, physiotherapists. Therefore, our results may be biased towards non-invasive interventions, often delivered by physiotherapists, such as rehabilitation. A further limitation was that most experts worked in Australia (69%). Whilst there were no clear differences in opinions between experts who worked in different countries, this could impact the external validity of the results.

9.4.3 Evaluation of outcome measures

A strength of the systematic review on clinimetric properties was the strict methodology applied. The systematic reviews (Chapter seven, study five and Chapter eight, study six) utilised steps outlined by COSMIN (180), the recommended standard for completing systematic reviews on clinimetric properties. Both systematic reviews followed rigorous methodology including pre-registration (Prospero) and had two independent researchers complete study screening, selection and risk of bias assessment. Further strengths of the two systematic reviews on clinimetric properties were that they provided clinicians and researchers with a comprehensive understanding of the current value of outcome measures that can be used to assess progress in patients with gluteal and proximal hamstring tendinopathy.

There were several limitations of the two systematic reviews on measurement properties (Chapter six, study four and Chapter eight, study six) that are worth considering when interpreting the results. The minimal important change was used in studies to evaluate the measurement error. As the minimal important change is not a fixed value, some caution should be taken when interpreting the results (e.g. the value is context-specific). Anchor-based methods were used in most included studies that assessed responsiveness (Chapter seven, study five and Chapter eight, study six). The most common outcome measure used for this purpose was the Global Rating of Change (GROC). Whilst this is a common method to measure responsiveness, a limitation is that due to the multidimensional nature of recovery the researcher does not know what the patient may consider when making the overall rating of recovery (e.g. pain, functional limitations, quality of life, side effects) (222). The sample size of patients included in studies on gluteal and hamstring tendinopathy was also low for several measurement properties, which meant many of the results of measurement properties were rated down for imprecision. There are also potential limitations of the COSMIN. These

include a lack of conditions to upgrade the quality of evidence (e.g. in outcome measures with 'strong' clinimetric performance), sample size requirements employed when rating measures are based on rule of thumb, and the reliability of the tool (e.g. the risk of bias checklist, requires further establishment).

9.4.4 Impact of covid-19 pandemic

The Covid-19 pandemic had an impact on the direction and duration of this thesis. The scope of this thesis was broadened to include a study on gluteal tendinopathy (study six) because the continuation of the original course was not possible due to a State-wide lockdown and restrictions on accessing university facilities impacting face-to-face data collection. The planned research included a cross-sectional study that was designed to describe the pain presentation of patients presenting with proximal hamstring tendinopathy (using pain mapping), assess the impairments of patients with proximal hamstring tendinopathy and explore associations between physical findings and other health-related factors (e.g. pain self-efficacy). The ethics for these planned studies can be found in Appendix U. Whilst Covid-19 lockdowns had an impact on the duration of the thesis, it created an opportunity to advance worldwide research plans from the recent ICON tendinopathy consensus.

9.5 Conclusion

This thesis aimed to further develop evidence on diagnosis, management and measurement of outcomes in proximal hamstring tendinopathy. A systematic review identified gaps in the literature and methodological limitations of previous research and found that there was at best, low-level evidence to guide management for proximal hamstring tendinopathy (155). Of note, was the dearth of research on exercise-based rehabilitation programs, which is considered best practice for other lower limb tendinopathies (223). Another important finding was the variability between studies of outcome measures used to capture the impact of the condition (155).

The series of studies that followed produced innovative and important information about the measurement and management of proximal hamstring tendinopathy. This included strategies used by experts to assess, manage and prevent the condition, and the clinimetric properties of outcome measures used to evaluate common tendinopathies of the hip and pelvis.

Publications from the International Tendinopathy Consensus (ICON) meeting in 2018 identified key health domains that should be measured in studies reporting on tendinopathy as well as reporting guidelines and consensus on the nomenclature (2, 80, 164). Based on this initiative, we designed studies three (Chapter five), four (Chapter six), five (Chapter seven) and six (Chapter eight). We followed the COSMIN, COMET and OMERACT initiatives in developing core outcome sets to conduct the three-part study: (i) a comprehensive review of literature trials to identify all outcomes used to date, (b) mapping extracted outcomes to the ICON core health-related domains for tendinopathy, and (c) then performing the first critical evaluation of their clinimetric properties (174). This process was completed for both gluteal and proximal hamstring tendinopathy. We found a lack of measures that are valid for capturing the core health domains of proximal hamstring and gluteal tendinopathy but also identified a measure of disability, the VISA-H and VISA-G, that we would recommend clinically in the interim, while further research develops the core outcome set.

As is often the case when conducting research, the findings of this thesis have evoked as many questions as answers. Further research must now be undertaken to standardise inclusion criteria for tendinopathies of the hip and pelvis, and formulate valid, reliable and responsive outcome measures that capture their impact. Standardising this information will pave a path on which to investigate the utility of interventions.

Appendices

The following chapter contains all appendices associated with Chapters one through nine followed by the posters presented at Sports Kongress and Sports Medicine Australia.

Appendix A: Ethical approval (study two)

THE UNIVERSITY OF QUEENSLAND

Institutional Human Research Ethics Approval

Project Title:	Proximal Hamstring Tendinopathy: current practices in assessment and management and prevention
Chief Investigator:	Dr Adam Semciw
Supervisor:	Dr Adam Semciw, Dr Alison Grimaldi, Prof Bill Vicenzino
Co-Investigator(s):	Mr Anthony Nasser, Dr Alison Grimaldi, Prof Bill Vicenzino
School(s):	School of Health and Rehabilitation Sciences, The University of Queensland
Approval Number:	2018001158
Granting Agency/Degree:	MPhil
Duration:	1st July 2019

Comments/Conditions:

- HREA Form, 04/06/2018
- Information and Consent Form - PHT, 04/06/2018
- Negligible Risk Human Ethics Application Form, 07/06/2018
- Schedule of questions – PHT 5th edit, 04/06/2018

Note: if this approval is for amendments to an already approved protocol for which a UQ Clinical Trials Protection/Insurance Form was originally submitted, then the researchers must directly notify the UQ Insurance Office of any changes to that Form and Participant Information Sheets & Consent Forms as a result of the amendments, before action.

Name of responsible Sub-Committee:
University of Queensland Health and Behavioural Sciences, Low & Negligible Risk Ethics Sub-Committee
This project complies with the provisions contained in the *National Statement on Ethical Conduct in Human Research* and complies with the regulations governing experimentation on humans.

Name of Ethics Sub-Committee representative:
Associate Professor Guy Wallis
Chairperson
University of Queensland Health and Behavioural Sciences, Low & Negligible Risk Ethics Sub-Committee

Signature Date **08/06/2018**

Appendix B: Ethical approval (study five)

HEC21210

Information

Record Name	HEC21210	**Ethics Application Title**	Development of a Core Outcome Set for proximal hamstring Tendinopathy (COS-PHT): a Delphi study of healthcare professionals and patients.
Ethics Application Number	HEC21210	**Record Type**	Human Ethics Application
Research Office Contact		**Trim Link**	
Sponsor			

Review and Approval

Ethics Application Approval Date	29/06/2021	**Ethics Application Expiry Date**	29/06/2026
Conditions of Approval		**Ethics Review Committee**	Low Risk Committee
Status	Outcome Communicated	**Ethics Application Review Outcome**	Approved
Ratified	N/A	**Meeting Date**	Research Office Review Friday, 2 July 2021
Assigned Reviewers	This field has intentionally been left blank	**Actions for researchers**	

Time to Decision

Time To Decision	8	**Number of Review Times**	2

Sites, Data and Privacy

Count of Collaborating Organisation/s	0	**Data and Privacy**	Personal Data; Health Data
Site Name	N/A		

Waiver of Consent Section

Waiver of Consent	☐
Waiver of Consent Reasons	N/A

Clinical Trials

Clinical Trials		**Trial Description**	N/A

Appendix C: Search strategy (study one)

Keywords	Hamstring or "Biceps femoris" or Semitendinos* or Semimembranos* or	Tendinopathy Tendinosis Tendinitis Tear
MeSH terms mapped to individual database thesaurus	Hamstring	Tendinopathy

Appendix D: Medline search strategy (study one)

Medline		
Date	30.01.2019	
Set	Results	
# 3	658	#2 AND #1 *Indexes=MEDLINE Timespan=All years*
# 2	56,973	(((TOPIC: (tendinopathy) *OR* TOPIC: (tendinitis)) *OR* TOPIC: (tendinosis)) *OR* TOPIC: (tear)) *Indexes=MEDLINE Timespan=All years*
# 1	13,288	(((TOPIC: (hamstring) *OR* TOPIC: (biceps femoris)) *OR* TOPIC: (semitendinos*)) *OR* TOPIC: (semimembranos*)) *Indexes=MEDLINE Timespan=All years*

Appendix E: Summary of findings (study one)

No. of patients/studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Publication bias	Effect estimate	Certainty
Surgery: VAS (follow-up: 53-71 months; self-reported symptoms)								
61/2	Case series	Serious [a]	Serious [c]	Not serious	Serious [d]	Not serious [e]	Large	⊕◯◯◯ VERY LOW
Surgery: Tegner Score (follow-up: mean 71 months; physical function)								
17/1	Case series	Serious [a]	Serious [b]	Not serious	Serious [d]	Not serious [e]	Large	⊕◯◯◯ VERY LOW
Surgery: subjective weakness score (follow-up: mean 53 months; physical function)								
44/1	Case series	Serious [a]	serious [b]	Not serious	Serious [d]	Not serious [e]	Not estimable	⊕◯◯◯ VERY LOW
Surgery: return to pre-injury level of sport (follow-up: 49-71 months; physical function)								
151/3	Case series	Serious [a]	serious [b]	Not serious	Serious [d]	Not serious [e]	Not estimable	⊕◯◯◯ VERY LOW
Surgery: return to undefined level of sport (follow-up: mean 24 months; physical function)								
59/1	Case series	Serious [a]	serious [b]	Not serious	Serious [d]	Not serious [e]	Not estimable	⊕◯◯◯ VERY LOW
Surgery: adverse effects (follow-up: 24-71 months)								
179/4	Case series	Serious [a]	not serious	Not serious	Serious [d]	Not serious [e]	Not estimable	⊕◯◯◯ VERY LOW
Multi-modal intervention vs multi-modal intervention + delayed PRP + multi-modal intervention: VAS (follow-up: 2 to 4.5 months; self-reported symptoms)								
17/1	Case series	Serious [a]	serious [b]	Not serious	Serious [d]	Not serious [e]	Not estimable	⊕◯◯◯ VERY LOW
PRP vs AWB: HOS ADL (follow-up: mean 6 weeks; physical function)								
17/1	RCT	Not serious	serious [b]	Not serious	Serious [d]	Not serious [e]	No apparent difference between groups	⊕⊕◯◯ LOW
PRP vs AWB: HOS ADL (follow-up: 12 weeks; physical function)								
17/1	RCT	Not serious	Serious [b]	Not serious	Serious [d]	Not serious [e]	No apparent difference between groups	⊕⊕◯◯ LOW
PRP vs AWB: HOS Sport (follow-up: 6 weeks; physical function)								
15/1	RCT	Not serious	Serious [b]	Not serious	Serious [d]	Not serious [e]	No apparent difference between groups	⊕⊕◯◯ LOW
PRP vs AWB: HOS Sport (follow-up: 12 weeks; physical function)								
15/1	RCT	Not serious	Serious [b]	Not serious	Serious [d]	Not serious [e]	No apparent difference between groups	⊕⊕◯◯ LOW
PRP vs AWB: HOS Sport (follow-up: 6 months; physical function)								

15/1	RCT	Not serious	Serious [b]	Not serious	Serious [d]	Not serious [e]	No apparent difference between groups	⊕⊕◯◯ LOW
PRP vs AWB: iHOT 33 (follow-up: 12 weeks; physical function)								
16/1	RCT	Not serious	Serious [b]	Not serious	Serious [d]	Not serious [e]	No apparent difference between groups	⊕⊕◯◯ LOW
PRP vs AWB: iHOT 33 (follow-up: 6 months; physical function)								
16/1	RCT	not serious	serious [b]	not serious	serious [d]	not serious [e]	No apparent difference between groups	⊕⊕◯◯ LOW
PRP vs AWB: MHHS (follow-up: 12 weeks; physical function)								
17/1	RCT	not serious	serious [b]	not serious	serious [d]	not serious [e]	No apparent difference between groups	⊕⊕◯◯ LOW
PRP vs AWB: MHHS (follow-up: 6 months; physical function)								
17/1	RCT	not serious	serious [b]	not serious	serious [d]	not serious [e]	No apparent difference between groups	⊕⊕◯◯ LOW
Multi-modal intervention: vs multi-modal intervention: + delayed PRP + multi-modal intervention: NRPS (follow-up: 4 and 4.5 months; physical function)								
17/1	Case series	Serious	Serious	Not serious	Serious	Not serious	Not estimable	⊕◯◯◯ VERY LOW
PRP: VISA H (follow-up: 8 weeks; physical function)								
29/1	Case series	Serious	Serious [b]	Not serious	Serious [d]	Not serious [e]	Small	⊕◯◯◯ VERY LOW
PRP: LEFS (follow-up: 12 weeks; physical function)								
14/1	Case series	Serious [a]	Serious [b]	Not serious	Serious [d]	Not serious [e]	Large	⊕◯◯◯ VERY LOW
PRP vs AWB: adverse effects (follow-up: 6 months)								
17/1	RCT	Not serious	Serious [b]	Not serious	Serious [d]	Not serious [e]	Not estimable	⊕⊕◯◯ LOW
PRP: adverse effects (follow-up: 8 weeks-46 months)								
47/2	Case series	Serious [a]	Not serious	Not serious	Serious [d]	Not serious [e]	Not estimable	⊕◯◯◯ VERY LOW
PRP: adverse effects (follow-up: 12 weeks)								
14/1	Case series	Serious [a]	Serious [b]	Not serious	Serious [d]	Not serious [e]	Not estimable	⊕◯◯◯ VERY LOW
CSI: VAS (assessed with: mean 21 months self-reported symptoms)								
18/1	Case series	Serious [a]	Serious [b]	Not serious	Serious [d]	Not serious [e]	Not estimable	⊕◯◯◯ VERY LOW
CSI: LEFS (follow-up: mean 21 months; physical function)								
18/1	Case series	Serious [a]	Serious [b]	Not serious	Serious [d]	Not serious [e]	Not estimable	⊕◯◯◯ VERY LOW
CSI: adverse effects (follow-up: 21-48 months)								

83/2	Case series	Serious [a]	Not serious	Not serious	Serious [d]	Not serious [e]	Not estimable	⊕◯◯◯ VERY LOW
SWT vs multi-modal intervention: VAS (follow-up: 1 week; self-reported symptoms)								
40/1	RCT	Not serious	Serious [b]	Not serious	Serious [d]	Not serious [e]	Large (favours SWT)	⊕⊕◯◯ LOW
SWT vs multi-modal intervention: VAS (follow-up: 6 months; self-reported symptoms)								
40/1	RCT	Not serious	Serious [b]	Not serious	Serious [d]	Not serious [e]	Large (favours SWT)	⊕⊕◯◯ LOW
SWT vs multi-modal intervention: VAS (follow-up: 12 months; self-reported symptoms)								
40/1	RCT	Not serious	Serious [b]	Not serious	Serious [d]	Not serious [e]	Large (favours SWT)	⊕⊕◯◯ LOW
SWT vs multi-modal intervention: subjective rating of improvement (follow-up: 1 week; self-reported symptoms)								
40/1	RCT	Not serious	Serious [b]	Not serious	Serious [d]	Not serious [e]	Large (favours SWT)	⊕⊕◯◯ LOW
SWT vs multi-modal intervention: subjective rating of improvement (follow-up: 6 months; self-reported symptoms)								
40/1	RCT	Not serious	Serious [b]	Not serious	Serious [d]	Not serious [e]	Large (favours SWT)	⊕⊕◯◯ LOW
SWT vs multi-modal intervention: subjective rating of improvement (follow-up: 12 months; self-reported symptoms)								
40/1	RCT	Not serious	Serious [b]	Not serious	Serious [d]	Not serious [e]	Large (favours SWT)	⊕⊕◯◯ LOW
SWT vs multi-modal intervention: NRPS (follow-up: 1 week; physical function)								
40/1	RCT	Not serious	Serious [b]	Not serious	Serious [d]	Not serious [e]	Large (favours SWT)	⊕⊕◯◯ LOW
SWT vs multi-modal intervention: NRPS (follow-up: 6 months; physical function)								
40/1	RCT	Not serious	Serious [b]	Not serious	Serious [d]	Not serious [e]	Large (favours SWT)	⊕⊕◯◯ LOW
SWT vs multi-modal intervention: NRPS (follow-up: 12 months; physical function)								
40/1	RCT	Not serious	Serious [b]	Not serious	Serious [d]	Not serious [e]	Large (favours SWT)	⊕⊕◯◯ LOW
SWT vs multi-modal intervention: return to pre-injury level of sport (follow-up: 12 months; physical function)								
40/1	RCT	Not serious	Serious [b]	Not serious	Serious [d]	Not serious [e]	Large (favours SWT)	⊕⊕◯◯ LOW
SWT vs multi-modal intervention: adverse events (follow-up: 12 months)								
40/1	RCT	Not serious	Serious [b]	Not serious	Serious [d]	Not serious [e]	Large (favours SWT)	⊕⊕◯◯ LOW

AWB: autologous whole blood injection, CI: confidence interval, CSI: corticosteroid injection, HOS: Hip Outcome Score, HOS-sport: Hip Outcome Score: sport, HOS-ADL: Hip Outcome Score: activities of daily living, iHOT-33: International Hip Outcome Tool, LEFS: Lower Extremity Functional Scale, MHHS: Modified Hip Harris Score, NA: not applicable, NRPS: Nirschl Phase Rating Scale, PRP: platelet-rich plasma injection, RCT: randomised controlled trial, SMD: standardised mean difference, SWT: shockwave therapy, VAS: Visual Analogue Scale, VISA H: Victorian Institute of Sport Assessment-proximal hamstring tendon

[a] ≥ 25% of the participants from studies with a high risk of bias

[b] Single study (n<300)

[c] $I^2 > 40\%$

[d] Pooled data with < 300 participants for an outcome

[e] The possibility of publication bias is not excluded, but was not considered sufficient to downgrade the quality of evidence

Appendix F: Interview guide of experts (study two)

Interview question guide	
Participant background	What is your current occupation? How long have you been working clinically or in research? How many cases of proximal hamstring tendinopathy do you see (or did you see, if no longer working clinically) per month?
Diagnosis	How do you come to the diagnosis of proximal hamstring tendinopathy?
Management	Could you take me through how you manage the condition, from early to late-stage management?
Monitoring	How do you monitor the patient throughout treatment? How would you describe your treatment outcomes for patients with proximal hamstring tendinopathy?
Treatment results	What are the key factors in achieving a good patient outcome?
Prevention	In your experience, are there any important factors for preventing recurrence of proximal hamstring tendinopathy?
Risk factors	In your experience, are there any risks factors that may lead to initial development of proximal hamstring tendinopathy?
Future research	What should priorities be for research in proximal hamstring tendinopathy?

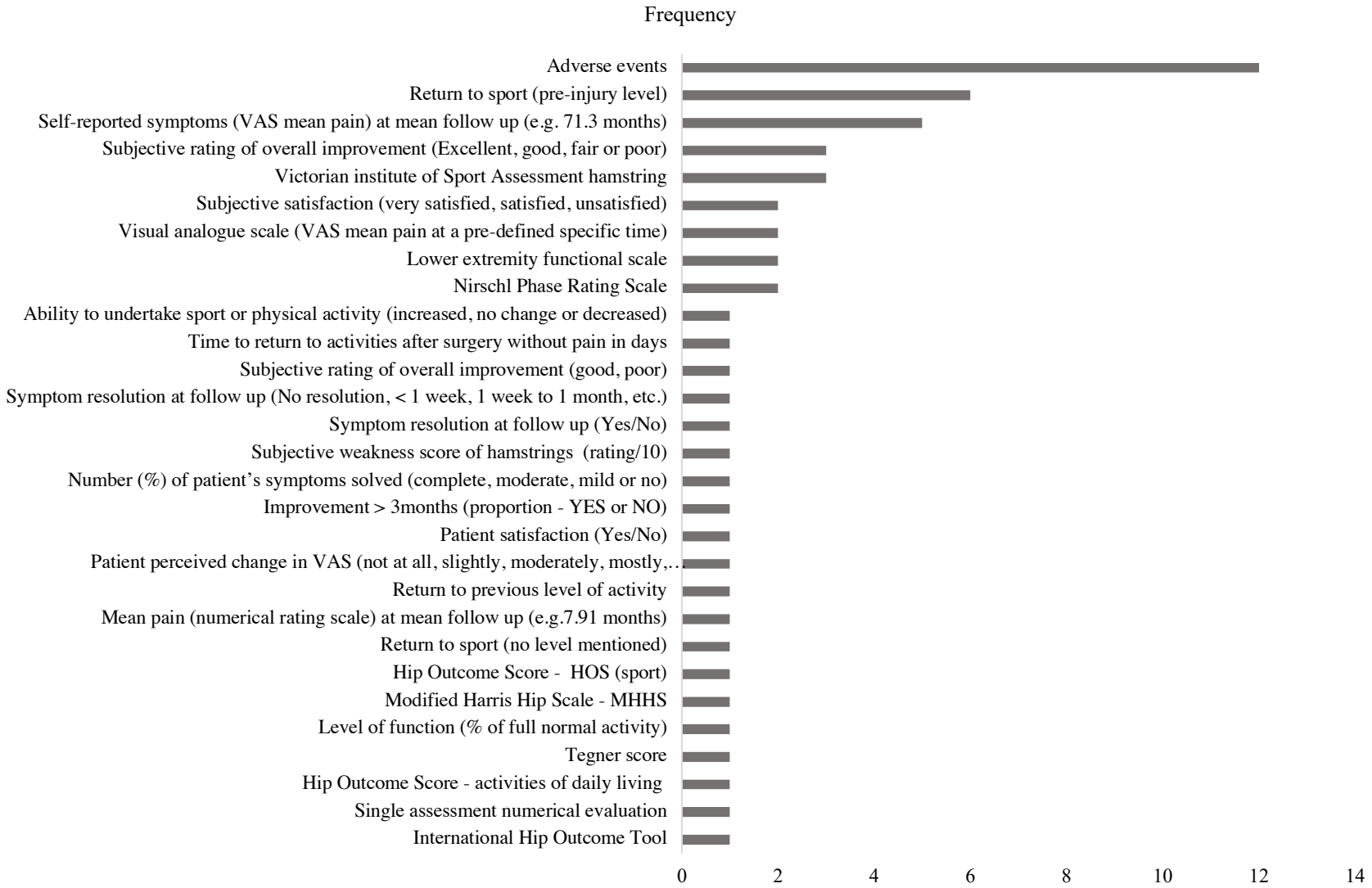
Appendix G: Search strategy PubMed (study four)

			Boolean operator
1	Population	Hamstring[MeSH] OR PHT[Title/Abstract] or "proximal hamstring tendin*"[Title/Abstract] or "hamstring tendin*"[Title/Abstract] or hamstring[Title/Abstract]	1
2	Instrument search	"Nirschl Phase Rating Scale" [Title/Abstract] OR NRPS[Title/Abstract] OR "hip outcome score"[Title/Abstract] OR HOS[Title/Abstract] OR "Modified harris hip score" [Title/Abstract] OR MHHS[Title/Abstract] OR Tegner[Title/Abstract] OR "Lower extremity functional scale"[Title/Abstract] OR LEFS[Title/Abstract] OR function[Title/Abstract] OR 'Victorian Institute of Sport'[Title/Abstract] OR VISA-H[Title/Abstract] OR Rating[Title/Abstract] OR Participation[Title/Abstract] OR Pain[Title/Abstract] OR "Psychological"[Title/Abstract] OR "Physical Function Capacity"[Title/Abstract] OR "Disability"[Title/Abstract] OR "Quality of Life"[Title/Abstract] OR QOL[Title/Abstract] OR "sport"[Title/Abstract] OR "visual analogue scale"[Title/Abstract] or VAS[Title/Abstract]	1 AND 2
3	Measurement properties filter	(instrumentation[sh] OR methods[sh] OR "Validation Studies"[pt] OR "Comparative Study"[pt] OR "psychometrics"[MeSH] OR psychometr*[tiab] OR clinimetr*[tw] OR clinometr*[tw] OR "outcome assessment (health care)"[MeSH] OR "outcome assessment"[tiab] OR "outcome measure*"[tw] OR "observer variation"[MeSH] OR "observer variation"[tiab] OR "Health Status Indicators"[Mesh] OR "reproducibility of results"[MeSH] OR reproducib*[tiab] OR "discriminant analysis"[MeSH] OR reliab*[tiab] OR unreliab*[tiab] OR valid*[tiab] OR "coefficient of variation"[tiab] OR coefficient[tiab] OR homogeneity[tiab] OR homogeneous[tiab] OR "internal consistency"[tiab] OR (cronbach*[tiab] AND (alpha[tiab] OR alphas[tiab])) OR (item[tiab] AND (correlation*[tiab] OR selection*[tiab] OR reduction*[tiab])) OR agreement[tw] OR precision[tw] OR imprecision[tw] OR "precise values"[tw] OR test-retest[tiab] OR (test[tiab] AND retest[tiab]) OR (reliab*[tiab] AND (test[tiab] OR retest[tiab])) OR stability[tiab] OR interrater[tiab] OR inter-rater[tiab] OR intrarater[tiab] OR intra-rater[tiab] OR intertester[tiab] OR inter-tester[tiab] OR intratester[tiab] OR intra-tester[tiab] OR interobserver[tiab] OR inter-observer[tiab] OR intraobserver[tiab] OR intra-observer[tiab] OR intertechnician[tiab] OR inter-technician[tiab] OR intratechnician[tiab]	1 AND 2 AND 3

		OR intra-technician[tiab] OR interexaminer[tiab] OR inter-examiner[tiab] OR intraexaminer[tiab] OR intra-examiner[tiab] OR interassay[tiab] OR inter-assay[tiab] OR intraassay[tiab] OR intra-assay[tiab] OR interindividual[tiab] OR inter-individual[tiab] OR intraindividual[tiab] OR intra-individual[tiab] OR interparticipant[tiab] OR inter-participant[tiab] OR intraparticipant[tiab] OR intra-participant[tiab] OR kappa[tiab] OR kappa's[tiab] OR kappas[tiab] OR repeatab*[tw] OR ((replicab*[tw] OR repeated[tw]) AND (measure[tw] OR measures[tw] OR findings[tw] OR result[tw] OR results[tw] OR test[tw] OR tests[tw])) OR generaliza*[tiab] OR generalisa*[tiab] OR concordance[tiab] OR (intraclass[tiab] AND correlation*[tiab]) OR discriminative[tiab] OR "known group"[tiab] OR "factor analysis"[tiab] OR "factor analyses"[tiab] OR "factor structure"[tiab] OR "factor structures"[tiab] OR dimension*[tiab] OR subscale*[tiab] OR (multitrait[tiab] AND scaling[tiab] AND (analysis[tiab] OR analyses[tiab])) OR "item discriminant"[tiab] OR "interscale correlation*"[tiab] OR error[tiab] OR errors[tiab] OR "individual variability"[tiab] OR "interval variability"[tiab] OR "rate variability"[tiab] OR (variability[tiab] AND (analysis[tiab] OR values[tiab])) OR (uncertainty[tiab] AND (measurement[tiab] OR measuring[tiab])) OR "standard error of measurement"[tiab] OR sensitiv*[tiab] OR responsive*[tiab] OR (limit[tiab] AND detection[tiab]) OR "minimal detectable concentration"[tiab] OR interpretab*[tiab] OR ((minimal[tiab] OR minimally[tiab] OR clinical[tiab] OR clinically[tiab]) AND (important[tiab] OR significant[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR (small*[tiab] AND (real[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR "meaningful change"[tiab] OR "ceiling effect"[tiab] OR "floor effect"[tiab] OR "Item response model"[tiab] OR IRT[tiab] OR Rasch[tiab] OR "Differential item functioning"[tiab] OR DIF[tiab] OR "computer adaptive testing"[tiab] OR "item bank"[tiab] OR "cross-cultural equivalence"[tiab]	
4	Exclusion filter	("addresses"[Publication Type] OR "biography"[Publication Type] OR "case reports"[Publication Type] OR "comment"[Publication Type] OR "directory"[Publication Type] OR "editorial"[Publication Type] OR "festschrift"[Publication Type] OR "interview"[Publication Type] OR "lectures"[Publication Type] OR "legal cases"[Publication Type] OR "legislation"[Publication Type] OR "letter"[Publication Type] OR "news"[Publication Type] OR "newspaper article"[Publication Type] OR "patient education handout"[Publication Type] OR "popular works"[Publication Type] OR "congresses"[Publication Type] OR "consensus development conference"[Publication	1 AND 3 AND NOT 4

	Type] OR “consensus development conference, nih”[Publication Type] OR “practice guideline”[Publication Type]) NOT (“animals”[MeSH Terms] NOT “humans”[MeSH Terms])

Appendix H: Outcome measures used in studies evaluating proximal hamstring tendinopathy (study four)



Appendix I: Coding of outcome measures to core tendinopathy domains (study four)

Domain	Outcome measure
Disability	Nirschl Phase Rating Scale
	Hip outcome score (Activity of Daily Living subscale)
	Victorian Institute of Sport – Hamstring Tendinopathy
	Modified Harris Hip Scale
	Hip outcome score (sport)
	International Hip Outcome Tool (Symptoms and functional limitations)
	Lower extremity functional scale[1]
	Functional Assessment Scale for Acute Hamstring injuries questionnaire
Participation	Tegner score
	Return to previous level of activity (Yes/No)[1]
	Return to sport (pre-injury)
	Return to sport (level not specified)
	International Hip Outcome Tool (Sports and recreational activities subscale)
	International Hip Outcome Tool (Job related concerns)
Function	Lower extremity functional scale[1]
	Level of function (% of full normal activity - 100% being full function in normal activities prior to injury)
	Return to previous level of activity (Yes/No)[1]
Pain over a specified timeframe	Patient-reported pain rating (Visual Analogue Scale) at a specified timepoint
	Mean pain (Visual Analogue Scale) at a specified time point
Patient overall rating	Patient-reported change in symptoms
	Improvement in symptoms in short term (Yes/No)
	Subjective rating of overall outcome (Excellent, good fair or poor)
	Subjective patient rating of improvement in symptoms (improved, same or worse)
	Single assessment numerical evaluation
Quality of life	International Hip Outcome Tool (social emotional and lifestyle concerns)
Unmapped	Patient-reported pain rating (Visual Analogue Scale) at mean follow-up (time not specified)
	Patient subjective satisfaction of treatment
	Number (%) of patient's symptoms solved
	Adverse events
	Numerical rating scale at mean follow-up (time point not specified)
	Subjective weakness score

[1]Outcome measure matched to two domains

Appendix J: Hypothesis testing for construct validity of VISA-H (study four)

Study	Comparison scale	Hypotheses	Pearson's R	Interpretation of Pearson's (strong, moderate, weak)	Validated hypotheses (Yes/No)
Cacchio (2013)	NPRS (non-surgical at baseline)	Strong	0.75	Strong	Yes
Cacchio (2013)	NPRS (surgical at baseline)	Strong	0.75	Strong	Yes
Cacchio (2013)	NPRS (non-surgical at discharge)	Strong	0.89	Strong	Yes
Cacchio (2013)	NPRS (surgical at discharge)	Strong	0.81	Strong	Yes
Cacchio (2013)	TGSCS (non-surgical at baseline)	Strong	0.79	Strong	Yes
Cacchio (2013)	TGSCS (surgical at baseline)	Strong	0.70	Strong	Yes
Cacchio (2013)	TGSCS (non-surgical at discharge)	Strong	0.88	Strong	Yes
Cacchio (2013)	TGSCS (surgical at discharge)	Strong	0.81	Strong	Yes
Locquet (2019)	SF-36 Physical function	Strong	0.78	Strong	Yes
Locquet (2019)	SF-36 Physical role	Strong	0.70	Strong	Yes
Locquet (2019)	SF-36 Body pain	Strong	0.73	Strong	Yes
Locquet (2019)	SF-36 General health	Weak	0.36	Moderate	No
Locquet (2019)	SF-36 Mental health	Weak	0.22	Weak	Yes
Locquet (2019)	SF-36 Social function	Weak	0.63	Strong	No
Locquet (2019)	SF-36 Vitality	Weak	0.32	Moderate	No
Locquet (2019)	FASH	Strong	0.75	Strong	Yes
De-la-Cruz-Torres (2021)	SF-36 Physical functioning	Strong	0.53	Strong	Yes
De-la-Cruz-Torres (2021)	SF-36 Body pain	Strong	0.53	Strong	Yes
De-la-Cruz-Torres (2021)	SF-36 Vitality	Weak	0.01	Weak	Yes
De-la-Cruz-Torres (2021)	SF-36 Emotional role	Weak	0.38	Moderate	No
De-la-Cruz-Torres (2021)	SF-36 Physical role	Strong	0.53	Strong	Yes

De-la-Cruz-Torres (2021)	SF-36 Social functioning	Weak	0.01	Weak	Yes
De-la-Cruz-Torres (2021)	SF-36 Mental health	Weak	0.15	Weak	Yes
De-la-Cruz-Torres (2021)	SF-36 General health	Weak	0.14	Weak	Yes
De-la-Cruz-Torres (2021)	SF-36 Physical component summary	Strong	0.52	Strong	Yes
De-la-Cruz-Torres (2021)	SF-36 Mental component summary	Weak	0.35	Weak	No
De-la-Cruz-Torres (2021)	FASH	Strong	0.77	Strong	Yes
De-la-Cruz-Torres (2021)	Lower limb functional index	Moderate	0.35	Moderate	Yes
Lima (2022)	LEFS	Strong	0.69	Strong	Yes

Hypothesis:

- Strong correlations (≥ 0.50) would be found with instruments measuring similar constructs
- Correlations with instruments measuring related, but dissimilar constructs would be lower (0.3-0.5)
- Correlations with instruments measuring unrelated constructs would be low (<0.3)

FASH = Functional Assessment Scale for Acute Hamstring injuries questionnaire, LEFS = Lower extremity functional scale, LLFI = Lower limb functional index, SF-36 = Short-form-36, NRPS = Nirschl phase rating scale, TGSCS = Generic tendon grading system proposed by Curwin and Stanish

Appendix K: Feasibility (study four)

Source	VISA-H
Patient's comprehensibility	NIA
Clinician's comprehensibility	NIA
Type and ease of administration	NIA
Length of the instrument	8 items
Completion time (minutes)	<5 minutes (VISA-H.Sp), mean 2 minutes and 15 seconds to answer (VISA-H.Br)
Patient's required mental and physical ability level	NIA
Ease of standardisation	NIA
Required equipment	Nil
Availability in different settings	Yes

NIA = no information available, VISA-H = Victorian Institute of Sport – Hamstring tendinopathy, VISA-H.Sp = Victorian Institute of Sport – Hamstring tendinopathy -Spanish

Appendix L: Interpretability (study four)

Source	PROM	Distribution of scores in the study population (mean)	Percentage of missing items and percentage of missing total scores	Floor and ceiling effects	Scores and change scores available for relevant (sub)groups	Minimal important change or minimal important difference	Information on response shift
Cacchio (2013)	VISA-H	Healthy controls: 99.3±1.2 Baseline non-surgical: 56.7±11.6 surgical: 45.8±12.2 Healthy: 99.3±1.2	No missing items	No floor or ceiling effects	PHT non-surgical: 25.3±15.8 PHT surgical: 41.1±18.9	MIC = 22/100	Non-surgical mx: SRM = 1.6 Surgical mx: SRM = 2.2
Locquet (2019)	VISA-H.F	PHT: 58 (38–73) Healthy (at risk): 97 (34–100) Healthy (not at risk): 100 (95–100)	NR	No floor or ceiling effects	NA	NR	NA
De-la-Cruz-Torres (2021) (2021)	VISA-H.Sp	Baseline Healthy: 97.7 ± 4.77 Baseline PHT: 50.2 ±14.5	No missing items	No floor or ceiling effects	Discharge: 90.1±5.1 3-month follow-up: 97.8 ± 1.9	Clinically meaningful improvement: 15/100	Non-surgical mx: 3 months follow-up: SRM = 3.1 Discharge: SRM = 0.59
Lima (2022)	VISA-H.Br	PHT: 59.80±15.01, healthy: 84.82±11.48	NR	No floor or ceiling effects	NA	NA	NA

Minimal important change = MIC, mx = management, NA = not applicable, NR = not reported, PHT = proximal hamstring tendinopathy, SRM = standard response mean, VISA-H = Victorian Institute of Sport – proximal hamstring tendinopathy, VISA-H.Br = Victorian Institute of Sport Assessment - Proximal hamstring tendinopathy – Brazil, VISA-H.F = Victorian Institute of Sport Assessment - Proximal hamstring tendinopathy - French, VISA-H.Sp = Victorian Institute of Sport Assessment - Proximal hamstring tendinopathy – Spanish

Appendix M: Survey (study five)

Q1
Thank you for agreeing to take part in this consensus study.

A recently published systematic review reported on interventions used in managing proximal hamstring tendinopathy (PHT) and identified a range of outcome measures used across individual studies. It is difficult to generate meaningful conclusions from heterogenous data. Adequate reporting and consistency in selection of outcome measures within similar health domains is critical for comparison of interventions and synthesising data across studies, in order to provide meaningful clinical recommendations.

You can find more on Core Outcome Sets at the COMET website (http://www.comet-initiative.org/). We have linked two documents that explain in plain language what a Core Outcome Set is and how it is developed. You can read them here: Comet plain language summary v4 and Delphi plain language summary for comet website

In this current survey you will be asked questions about outcome measures that we have found in the literature. The questions are presented in 9 sections. Each section asks about the measures for a Health-Related Core Domain for tendinopathy (for a quick look at the 9 Core Domains see: 9 core domains infographic). The relevant domain is explained to you in each section. If you would like more information on the Core Domains for Tendinopathy you can read the paper (Table 2 in this paper has key information) by clicking here - Icon domains. It might be worthwhile to save this paper for reference throughout the survey. For a simple breakdown of the domains click here: Domain summary

In the survey you will be asked if you agree, disagree or are unsure about an outcome measure being in the COS-PHT. After this first survey we will collect all responses. The responses will be used to reduce the number of outcomes that might be in the COS-PHT. The consensus process means that we count your responses. This is very much like an election vote, but instead of a politician we are voting for an outcome measure for each Health-Related Core Domain. The number of votes that an outcome measure needs to be included in the COS-PHT is described in: Consensus criteria

This survey will take approximately 30 minutes to complete if you are familiar with the outcome measures and domains. It may take longer if you are not familiar with the measures or the domains. If you do not have a single block of time to do the survey, you are able to return to the survey at a later time. To do this, you will need to do use the survey link you were sent by email. Always click on the >> at the bottom of the page to save that page's responses. Please remember to complete the survey within 14 working days of receiving the email with the survey link.

The nature of the survey means that it will be better completed on a computer and not a mobile device.

Qualtrics, the survey software, automatically generates a mobile device version of the survey, but we recommend you use a computer.

Q2 This is the consent form Information regarding this survey: Project Title: Development of a core outcome set for proximal hamstring tendinopathy (COS-PHT): A consensus process involving a scoping review of the literature and a Delphi study of healthcare professionals and patients. Investigator(s): Mr Anthony Nasser, La Trobe University, School of Allied Health Contact details: Phone: +61 (02) 9514 7337 E-mail: a.nasser@latrobe.edu.au Co-investigators: Dr Adam Semciw, Prof Bill Vicenzino, Dr Alison Grimaldi, Dr Ebonie Rio, and Dr Tania Pizzari. We aim to determine the Core Outcome Set for proximal hamstring tendinopathy. The core outcome set for proximal hamstring tendinopathy is the agreed set of outcomes that should be reported in clinical trials of proximal hamstring tendinopathy. Keep in mind it is the minimum set of outcomes. It does not preclude researchers from taking other measures. The important thing is that the measures in the Core Outcome Set for proximal hamstring tendinopathy are standardized and valid measures of proximal hamstring tendinopathy. We will determine the Core Outcome Set for proximal hamstring tendinopathy with a consensus process that involves two online surveys and a follow-up meeting for those who are able to attend. The first survey will help reduce the number of possible outcome measures that might fit the core outcome set for proximal hamstring Tendinopathy. The second survey will be is important to help finalize the most likely measures in the core outcome set. In the online meeting will then make the final recommendations, based on the results of our online surveys. In the surveys and online meeting, you will be asked questions about outcome measures that we have found in the literature. The questions are presented in 9 sections. Each section asks about the measures for a Health-Related Core Domain. The Health-Related Core Domain is explained to you in each section. If you would like more information on the Core Domains for Tendinopathy you can read the reference paper at this link [or email Anthony Nasser and he will send it to you]. You will be asked if you agree (yes), not agree (no) or are unsure that a measure: (a) is a truthful measure of the domain (valid) and (b) is practical to use in clinic (feasible). There is also a space for you to type in comments that you feel are important for us to consider. We encourage you to provide as much detail as you feel important. We will use agreed criteria for including or not including a measure in the Core Outcome Set for proximal hamstring tendinopathy. We will include those that at least 70% of responders agree with as long as no more than 15% disagree. We will not include those measures that at least 70% of responders disagree with, as long as there are no more than 15% that agree to it being included. What are the risks of participating? There are no foreseeable risks or discomfort associated with providing the required information on the surveys or with the discussions and voting at the online consensus meeting. *Please click on >> even if you do not agree to do this survey. This saves your decision in the system. It will record you have decided not to do the survey and will prevent us from inviting you again to do the survey.*

- ○ I AGREE to take part in this survey (9)
- ○ I DO NOT AGREE to take part in this survey (10)

Skip To: End of Survey If This is the consent form Information regarding this survey: Project Title: Development of a... = I DO NOT AGREE to take part in this survey

End of Block: Preamble

Start of Block: Survey

Q3 Qualtrics does not collect any identifying data, so it is important for our participant and authorship requirements that you leave your name and email here, and to allow us to contact you for the consensus process. Remember the data will only ever be reported as anonymized data, where you will not be identified. The data you provide on this survey is password protected and secure.

- Last name: (4) _____
- First name: (5) _____
- Email: (6) _____
- Age: (7) _____

Q4 Sex

- Female (1)
- Male (2)

Q5 Professional discipline. Note that you can select more than one.

- Orthopaedic Surgeon (1)
- Physiotherapist (2)
- Clinician Researcher (3)
- Sports Medicine Physician (4)
- Chiropractor (5)
- Rheumatologist (6)

☐ General practitioner (8)

☐ Other (please list) (9) _____

Q6

At this stage we hope to gather your thoughts on whether each outcome records things about PHT that you believe are important and is feasible.

Whilst rating outcome measures remember that it is completely acceptable for us not to select any of these measures, and that there is an expectation that once selected the measure will be used in all clinical trials.

Q7 Domain 1: Patient Rating of Condition The Patient Rating of Condition Domain is typically measured with a single question that asks how the patient feels about their condition. For the outcome measures listed below, please select yes (agree), no (disagree) or unsure for the following: (a). Truth: Does the outcome measure record the things about proximal hamstring tendinopathy that you believe are important? (b). Feasibility: Are the following outcome measures practical to use in a clinic? This outcome was measured in 3 different ways, in 8 different studies. Please rate the outcome measures below for (a) truth and (b) feasibility that have been used in research on PHT.

Please click on the link below to read more in-depth information about the outcome(s) used: Domain 1: patient rating of condition - Click on the link for more in depth information

	Truth			Practical		
	Yes (1)	No (2)	Unsure (3)	Yes (1)	No (2)	Unsure (3)
Global rating of change scale (e.g. improved, the same or worse) (1)	○	○	○	○	○	○
Rating of symptom improvement (e.g. Yes, no) (2)	○	○	○	○	○	○
Overall rating of outcome (e.g. poor,	○	○	○	○	○	○

fair, good or excellent) (3)

Q8 Domain 1: Patient rating of condition domain comment: We are very keen to hear what you feel is important about the outcome measures listed above. Especially if you would map the outcome measure to another domain. Please suggest outcomes measures that you think belong in this domain that could be used in an interim COS-PHT. Add any other comments about your decisions there as well.

Q9 Domain 2: Participation The Participation in Life Activities Domain is usually measured with a questionnaire. The questionnaires typically ask the patients to rate their level of participation in activities such as sport, work or usual life activities. These measures are presented below for you to consider. For the outcome measures listed below, please select yes (agree), no (disagree) or unsure for the following: (a). Truth: Does the outcome measure record the things about proximal hamstring tendinopathy that you believe are important? (b). Feasibility: Are the following outcome measures practical to use in a clinic? This outcome measured was reported in 4 different ways, in 8 different studies. Please rate the outcome measures below for (a) truth and (b) feasibility that have been used in research on PHT.

Please click on the link below to read more in-depth information about the outcome(s) used:

Domain 2 - Participation in Life Activities - click on the link for more in depth information

	Truth			Practical		
	Yes (1)	No (2)	Unsure (3)	Yes (1)	No (2)	Unsure (3)
Return to Sport (level not defined) (1)	○	○	○	○	○	○
Return to Sport (Pre-injury level) (2)	○	○	○	○	○	○
Return to previous level of activity (3)	○	○	○	○	○	○
Tegner Score (4)	○	○	○	○	○	○
iHOT-33 (Sports and recreational	○	○	○	○	○	○

activities subscale) (5) iHOT-33 (Job related concerns sub scale) (6)	○	○	○	○	○	○

Q10 Domain 2: Participation domain comment: We are very keen to hear what you feel is important about the outcome measures listed above. Especially if you would map the outcome measure to another domain. Please suggest outcomes measures that you think belong in this domain that could be used in an interim COS-PHT. Add any other comments about your decisions there as well.

Q11 Domain 3: Pain on activity/loading The Pain on Activity or Loading Domain is typically measured on a pain rating scale. For example, the patient is asked to mark their level of pain on a 100 mm line on a piece of paper. This is usually called a pain visual analogue scale (or VAS for short). Alternatively, the patient is asked to rate their level of pain by nominating a number between 0 and 10. This is usually called a pain numerical rating scale (or NRS for short). Both of these scales for measuring pain have 0 being no pain and the other end (100 mm or 10 point) being worst pain imaginable.

The important part of this domain is that the patient is asked to think of the level of pain that occurs when they do an activity that loads or stresses the tendon. That is, they do one of the activities that are usually considered to be linked to proximal hamstring tendinopathy. For example, they might be asked to rate their pain when sitting, lunging, or running.

There were no studies that measured pain with an activity or a loading task.

Whilst there were no measures in studies on proximal hamstring tendinopathy that capture this domain, examples in the original ICON paper included: Participant/patient-reported intensity of pain on performing a task/activity that loads the tendon. (e.g. VAS or NRS for pain intensity when the patient performs a tendon-specific pain-provocative task).

Please suggest outcomes measures (e.g. loading tasks) that you think belong in this domain that could be used in an interim COS-PHT. Add any other comments about your decisions there as well.

Q12 Domain 4: Function The Function Domain measures how much difficulty the patient is having with certain activities. The activities might be ones they nominate, or there might be a predefined list of activities for the lower limb. The patient typically rates the amount of difficulty on a numerical rating scale.
The lowest end of the scale is the patient having extreme difficulty doing the activity. The highest end of the scale is selected when the patient has no difficulties doing the activity.

For the outcome measures listed below, please select yes (agree), no (disagree) or unsure for the following: (a). Truth: Does the outcome measure record the things about Proximal Hamstring Tendinopathy that you believe are important? (b). Feasibility: Are the following outcome measures practical to use in a clinic?
We found 2 different ways in which function is measured, across 2 different studies. Please rate the outcome measures below for (a) truth and (b) feasibility that have been used in research on PHT.

Please click on the link below to read more in-depth information about the outcome(s) used. This includes copies of the outcome measures. Domain 4: function - Click on the link for more in depth information

	Truth			Practical		
	Yes (1)	No (2)	Unsure (3)	Yes (1)	No (2)	Unsure (3)
Lower Extremity Functional Scale (LEFS) (1)	○	○	○	○	○	○
Level of Function (% of full activity) (2)	○	○	○	○	○	○

Q13 Domain 4: Function domain comment: We are very keen to hear what you feel is important about the outcome measures listed above. Especially if you would map the outcome measure to another domain. Please suggest outcomes measures that you think belong in this domain that could be used in an interim COS-PHT. Add any other comments about your decisions there as well.

Q14 Domain 5: Psychology The Psychological Factors Domain is usually measured with questionnaires focusing on different psychological aspects. Some questions relate to anxiety and depression and others to a person's mental response to a condition. This includes
confidence, exaggerated negative thinking or fear of movement or activity.

There were no outcome measures that were mapped to the psychology domain. Whilst there were no measures in studies on proximal hamstring tendinopathy that capture this domain, examples in the original ICON paper included: Pain self-efficacy scale, Pain Catastrophizing Scale (PCS), Kinesiophobia scales (e.g. Tampa Scale of Kinesiophobia (TSK)), and anxiety and depression scales (e.g. Hospital Anxiety and Depression Scale (HADS) or Patient Health Questionnaire 9 (PHQ9)). Domain 5 psychological - click on the link for more information
Please use the text box below to indicate if any of these outcome measures, or any other outcome measures could be used in a COS-PHT within the psychology domain. Add any other comments about your decisions there as well.

Q15 Domain 6: Physical Function Capacity

The Physical Function Capacity Domain is about measuring a patient's capacity to do a physical task. The following will provide an idea of what this domain is about with three examples

a.) A stopwatch can
be used to measure the time taken to walk a number of stairs or a set distance on a flat surface.
b). A strength testing machine, which is called a dynamometer, would
measure the force that muscles can produce
c). Counting the
number of hops or single limb squats that a patient can perform is another physical function capacity measure.

There were no outcome measures that were mapped to the physical function domain. Whilst there were no measures in studies on proximal hamstring tendinopathy that capture this domain, examples in the original ICON paper included: number of hops, timed stair walk, number of single leg squats, dynamometry (strength) and wearable technology. *Please suggest outcomes measures that you think belong in this domain that could be used in a COS-PHT. Add any other comments about your decisions there as well.*

Q16 Domain 7: Disability

The Disability Domain is usually measured with a questionnaire. The questionnaire usually contains a number of questions that ask the patient about how their pain interferes with various activities in their life.

For the outcome measures listed below, please select yes (agree), no (disagree) or unsure for the following:
(a). Truth: Does the outcome measure record the things about proximal hamstring
tendinopathy that you believe are important?
(b). Feasibility: Are the following outcome measures practical to use in a clinic?

The Disability domain was recorded in six different ways, in six different studies. Please rate the outcome measures below for (a) truth and (b) feasibility that have been used in research on PHT.

Please click on the link below to read more in-depth information about the outcome(s). This includes copies of the outcome measures:
Domain 7 Disability - Click on the link for more in depth information

	Truth			Practical		
	Yes (1)	No (2)	Unsure (3)	Yes (1)	No (2)	Unsure (3)
Victorian Institute of Sport	○	○	○	○	○	○

Questionnaire - Proximal Hamstring Tendinopathy (VISA-H) (1)						
Nirschl Phase Rating Scale (NPRS) (2)	○	○	○	○	○	○
Hip Outcome Score - Activities of Daily Living (HOS-ADL) (3)	○	○	○	○	○	○
Hip Outcome Score (HOS-Sport) (4)	○	○	○	○	○	○
Modified Harris Hip Scale (MHHS) (5)	○	○	○	○	○	○
Lower Extremity Functional Scale (LEFS) (6)	○	○	○	○	○	○
iHOT-33 (Symptoms and functional limitations) (7)	○	○	○	○	○	○

Q17 Domain 7: Disability domain comment:

We are very keen to hear what you feel is important about the outcome measures listed above. Especially if you would map the outcome measure to another domain.

Please suggest outcomes measures that you think belong in this domain that could be used in an interim COS-PHT. Add any other comments about your decisions there as well.

Q18 Domain 8: Quality of life

The Quality of Life (QoL) Domain is usually measured with a multi-item questionnaire. The questionnaires typically ask the patients to rate their general well-being.

For the outcome measures listed below, please select yes (agree), no (disagree) or unsure for the following:
(a). Truth: Does the outcome measure record the things about proximal hamstring tendinopathy that you believe are important?
(b). Feasibility: Are the following outcome measures practical to use in a clinic?

The Quality of Life Domain was measured in 1 way, in a single study. Please rate the outcome measures below for (a) truth and (b) feasibility that have been used in research on PHT.

Please click on the link below to read more in depth information about the outcome(s). This includes copies of the outcome measures: Domain 8 - quality of life - click on the link for more in-depth information

	Truth			Practical		
	Yes (1)	No (2)	Unsure (3)	Yes (1)	No (2)	Unsure (3)
iHOT-33 (social emotional and lifestyle concerns subscale) (1)	○	○	○	○	○	○

Q19 Domain 8: Quality of Life domain comment:

We are very keen to hear what you feel is important about the outcome measure listed above. Especially if you would map the outcome measure to another domain. Please suggest outcomes measures that you think belong in this domain that could be used in an interim COS-PHT. Add any other comments about your decisions there as well.

Q20 Domain 9: Pain Over a Specified Time

The Pain Over a Specified Time Domain is typically measured on a pain rating scale. For example, the patient is asked to mark their level of pain on a 100 mm line on a piece of paper. This is usually called a pain visual analogue scale (or VAS for short). Alternatively, the patient is asked to rate their level of pain by nominating a number between 0 and 10. This is usually called a pain numerical rating scale (or NRS for short). Both of these scales for measuring pain have 0 being no pain and the other end (100 mm or 10 point) being worst pain imaginable.

The important part of this domain is that patient is asked to think of a period of time and rate their level of pain at that time.

For example, they might be asked for their pain now (at this moment); this morning or last night; over the past 24 hours, week or month). Less often the patient is just asked if they had pain, with the reply being either yes or no. For example, do you have pain at night. This outcome measure is typically only one question.

For the outcome measures listed below, please select yes (agree), no (disagree) or unsure for the following:

(a). Truth: Does the outcome measure record the things about proximal hamstring tendinopathy that you believe are important?
(b). Feasibility: Are the following outcome measures practical to use in a clinic?

The Domain: Pain Over a Specified Time was reported 1 way, in two studies. Please rate the outcome measures below for (a) truth and (b) feasibility that have been used in research on PHT.

Measurement properties: There have been no clinimetric studies done on this measure in participants with Proximal Hamstring Tendinopathy.

Please click on the link below to read more in depth information about the outcome(s) used to assist you in making your decision. Domain 9: pain over a specified time. Click on the link for more in depth information

	Truth			Practical		
	Yes (1)	No (2)	Unsure (3)	Yes (1)	No (2)	Unsure (3)
Mean pain at a specified time point (e.g. 12 weeks) (1)	○	○	○	○	○	○

Q21 Domain 9: Pain over a specific period of time domain comment:
We are very keen to hear what you feel is important about the outcome measure listed above. Especially if you would map the outcome measure to another domain.
Please suggest outcomes measures that you think belong in this domain that could be used in an interim COS-PHT. Add any other comments about your decisions there as well.

Q24 The study working group could not map the outcome measures listed below that have been used in studies on proximal hamstring tendinopathy to any of the core domains.
If you think a measure belongs to a domain, please indicate by ticking yes on no. In the free text space provided next to the outcome measure please indicate domain the measure belongs to.
For more information on the domains click here (Domain summary).

Yes = belongs in a domain

	Yes (1)	No (2)
Adverse events (1)	○	○
Patient satisfaction of treatment (2)	○	○
Number (%) of patient's symptoms solved (3)	○	○
Patient-reported pain rating (VAS) at mean follow-up (no specific time point) (5)	○	○
Mean pain (numerical rating scale) at mean follow-up (no specific time point) (6)	○	○

Q25 Thank you for taking the time to complete this survey. We look forward to providing you with a list of measures that should be included in the final COS-PHT or interim COS-PHT. Measures requiring further discussion will be addressed in the Delphi consensus meeting online. Please contact us if you have any suggestions or feedback on how the process is going so far.

'Next' will take you to the end and save this last page of responses.

End of Block: Survey

Appendix N: Search results PubMed search one (study six)

PubMed		Search number combined with Boolean operator
1	(“Gluteal tendinopathy" OR "Greater trochanteric pain syndrome" OR abductor tendinopathy OR "lateral hip pain" OR “trochanteric bursitis” OR (GTPS NOT (Guanosine-triphosphate OR "green tea" OR "grapevine trunk pathogens")) OR ((gluteal OR "gluteus medius" OR "gluteus minimus") AND (tendinitis OR tendonitis OR tendinosis OR tendinopathy OR “tendon pathology” OR enthesis OR enthesopathy OR tear OR tendon)))	1
2	("therapy"[Subheading] OR "therapy"[All Fields] OR treat*[All Fields] OR "interventions" OR "intervention" OR "managements" OR "management" OR "manages" OR "managed" OR "manage" OR "rehabilitation"[Subheading] OR "rehabilitation"[All Fields] OR "rehabilitation"[MeSH Terms] OR "rehab"[All Fields] OR physiotherapy OR "physical therapy")	1 AND 2
Results	870	

Appendix O: Outcome measures used in gluteal tendinopathy (study six)

Domain	Outcome measure
Patient rating of condition	Global rating of satisfaction
	Global rating of change
	Patient acceptable symptom state
	Single assessment numeric evaluation
Participation in life activities	Lower Extremity Activity Scale
	University of California at Los Angeles Activity Score
	Work Participation
Pain on loading	Single leg stance
	Walking a distance
	Strength test pain
	Unspecified activity
	Sleep
Function	Lower Extremity Functional Score
	Patient Specific Functional Score
Psychological factors	Pain Catastrophizing Scale
	Pain Self-Efficacy Questionnaire
	Patient Health Questionnaire 9
	Tampa Scale of Kinesiophobia
	Hospital Anxiety and Depression Scale
Physical Functional Capacity	Strength
	Six-minute walk test
	Ten-minute walk test 10mtw
	Balance
	Timed up and Go
	Lag
Disability	Harris Hip Score
	Modified Harris Hip Score
	Oxford Hip Score
	Hip Outcome Score (Sport)
	Hip Outcome Score (Activities of Daily Living)
	Victorian Institute of Sport Assessment – Gluteal Tendinopathy
	Western Ontario and McMaster Universities Arthritis Index
	Non-Arthritis Hip Score
	Liquesce Index
	Lateral Hip Pain Questionnaire
	Chronic Pain Grade Score
	Hip Disability and Osteoarthritis Outcome Score
	Roles and Maudsley Score
	Merle d'Aubigne and Postal Score
	Copenhagen Hip and Groin Outcome Score
	Greater Trochanteric Pain Syndrome Patient-reported Outcome
Quality of Life	International Hip Outcome Tool-33
	International Hip Outcome Tool-12
	Assessment of Quality of life-8D
	Assessment of Quality of life-6D
	Assessment of Quality of life-4D
	EuroQol five-dimensional questionnaire-5L
	EuroQol five-dimensional questionnaire-3L
	Musculoskeletal Health Questionnaire
	Short form-36
	Short form-12
Pain over specific time	Pain at night (numerical rating scale - pain)
	Pain today (numerical rating scale - pain)
	Pain now (numerical rating scale - pain)
	Pain at times over the past 2 weeks (numerical rating scale - pain)
	Pain over the past month (numerical rating scale - pain)

	Pain over the past week (numerical rating scale - pain)
Could not be mapped to any domain (examples)	Movement analysis (e.g. Vicon motion systems) Clinical examination Pain (over no specified time-point or without loading) Physical activity General questions Satisfaction with treatment

Appendix P: Search results of PubMed Search two (study six)

PubMed Search			Search number combined with Boolean operator
1	Population	Hip[MeSH Terms] OR 'hip injuries'[MeSH Terms] OR groin[MeSH Terms] OR (gluteal[Title/Abstract] AND (tendin*[Title/Abstract] OR tendon*[Title/Abstract] OR 'tendon pathology'[Title/Abstract] OR enthesis[Title/Abstract] OR enthesopathy[Title/Abstract] OR tear*[Title/Abstract] OR avul*[Title/Abstract] OR rupture[Title/Abstract] OR teno*[Title/Abstract])) OR ('gluteus minimus'[Title/Abstract] AND (tendin*[Title/Abstract] OR tendon*[Title/Abstract] OR 'tendon pathology'[Title/Abstract] OR enthesis[Title/Abstract] OR enthesopathy[Title/Abstract] OR tear*[Title/Abstract] OR avul*[Title/Abstract] OR rupture[Title/Abstract] OR teno*[Title/Abstract])) OR ('gluteus medius'[Title/Abstract] AND (tendin*[Title/Abstract] OR tendon*[Title/Abstract] OR 'tendon pathology'[Title/Abstract] OR enthesis[Title/Abstract] OR enthesopathy[Title/Abstract] OR tear*[Title/Abstract] OR avul*[Title/Abstract] OR rupture[Title/Abstract] OR teno*[Title/Abstract])) OR (abductor[Title/Abstract] AND (tendin*[Title/Abstract] OR tendon*[Title/Abstract] OR 'tendon pathology'[Title/Abstract] OR enthesis[Title/Abstract] OR enthesopathy[Title/Abstract] OR tear*[Title/Abstract] OR avul*[Title/Abstract] OR rupture[Title/Abstract] OR teno*[Title/Abstract])) OR GTPS[Title/Abstract] OR 'trochanteric pain'[Title/Abstract] OR 'lateral hip pain'[Title/Abstract] OR 'trochanteric bursitis'[Title/Abstract]	1
2	Instrument search	'Global rating of satisfaction'[Title/Abstract] OR 'global rating of change'[Title/Abstract] OR GROS[Title/Abstract] OR GROC[Title/Abstract] OR 'single assessment numeric evaluation'[Title/Abstract] OR SANE[Title/Abstract] OR PASS[Title/Abstract] OR 'Patient acceptable symptom state'[Title/Abstract] OR satisfaction [Title/Abstract] OR 'Lower extremity activity scale'[Title/Abstract] OR LEAS[Title/Abstract] OR 'University of California at Los Angeles Activity Score'[Title/Abstract] OR 'UCLA Activity score'[Title/Abstract] OR 'Lower extremity functional score'[Title/Abstract] OR LEFS[Title/Abstract] OR 'single leg stance'[Title/Abstract] OR SLS[Title/Abstract] OR Sleep[Title/Abstract] OR walking[Title/Abstract] OR stair[Title/Abstract] OR GROS[Title/Abstract] OR GROC[Title/Abstract] OR 'single assessment numeric evaluation'[Title/Abstract] OR SANE[Title/Abstract] OR PASS[Title/Abstract] OR 'Patient acceptable symptom state'[Title/Abstract] OR satisfaction[Title/Abstract] OR 'Patient specific functional score'[Title/Abstract] OR PSFS[Title/Abstract] OR 'pain catastrophi?ing scale'[Title/Abstract] OR PCS[Title/Abstract] OR 'pain self-efficacy questionnaire'[Title/Abstract] OR 'patient health questionnaire 9'[Title/Abstract] OR PSEQ[Title/Abstract] OR PHQ[Title/Abstract] OR 'Tampa scale of Kinesiophobia'[Title/Abstract] OR TSK[Title/Abstract] OR 'Healthy Anxiety and depression scale'[title/abstract] HADS[Title/Abstract] OR TAMPA[Title/Abstract] OR Strength[Title/Abstract] OR 'six minute walk test'[Title/Abstract] OR 6mwt[Title/Abstract] OR 'walk test'[Title/Abstract] OR 10mtw[Title/Abstract] OR balance[Title/Abstract] OR ' timed-up and go'[Title/Abstract] OR TUG[Title/Abstract] OR Lag[Title/Abstract] OR 'modified Harris hip'[Title/Abstract] OR mHHS[Title/Abstract] OR 'oxford hip score'[Title/Abstract] OR OHS[Title/Abstract] OR 'Hip outcome score'[Title/Abstract] OR HOS[Title/Abstract] OR VISA-G[Title/Abstract] OR 'Victorian Institute of sport assessment'[Title/Abstract] OR 'Oswestry disability index'[Title/Abstract] OR ODI[Title/Abstract] OR WOMAC[Title/Abstract] OR 'Western Ontario and McMaster Universities Arthritis index'[Title/Abstract] OR 'Non-arthritic hip score'[Title/Abstract] OR NAHS[Title/Abstract] OR 'Chronic Pain Grade Score'[Title/Abstract] OR CPGS[Title/Abstract] OR 'Lateral	1 AND 2

		hip pain questionnaire'[Title/Abstract] OR LHPQ[Title/Abstract] OR 'Hip disability and osteoarthritis outcome score'[Title/Abstract] OR HOOS[Title/Abstract] OR 'Roles and Maudsley Score'[Title/Abstract] OR RMS[Title/Abstract] OR 'Merle d'Aubigne and Postal Score'[Title/Abstract] OR MdA[Title/Abstract] OR 'Copenhagen hip and groin outcome score'[Title/Abstract] OR HAGOS[Title/Abstract] OR GTPS-PRO[Title/Abstract] OR 'Lequesne Index'[Title/Abstract] OR L-index[Title/Abstract] OR 'Quality of life'[Title/Abstract] OR QOL[Title/Abstract] OR 'International hip outcome tool'[Title/Abstract] OR 'iHot-33'[Title/Abstract] OR iHot-12[Title/Abstract] OR 'Assessment of Quality of life'[Title/Abstract] OR AQoL-8d[Title/Abstract] OR AQoL-6D[Title/Abstract] OR AQoL-4d[Title/Abstract] OR 'EuroQol five-dimensional questionnaire'[Title/Abstract] OR EQ-5D-5L[Title/Abstract] OR EQ-5D-3L[Title/Abstract] OR 'Musculoskeletal health questionnaire'[Title/Abstract] MSK-HQ[Title/Abstract] OR 'Short form 36'[Title/Abstract] OR SF-36[Title/Abstract] OR SF-12[Title/Abstract] OR Pain[Title/Abstract]	
3	Validated measurement properties filter	(instrumentation[sh] OR methods[sh] OR 'Validation Studies'[pt] OR 'Comparative Study'[pt] OR 'psychometrics'[MeSH] OR psychometr*[tiab] OR clinimetr*[tw] OR clinometr*[tw] OR 'outcome assessment (health care)'[MeSH] OR 'outcome assessment'[tiab] OR 'outcome measure*'[tw] OR 'observer variation'[MeSH] OR 'observer variation'[tiab] OR 'Health Status Indicators'[Mesh] OR 'reproducibility of results'[MeSH] OR reproducib*[tiab] OR 'discriminant analysis'[MeSH] OR reliab*[tiab] OR unreliab*[tiab] OR valid*[tiab] OR 'coefficient of variation'[tiab] OR coefficient[tiab] OR homogeneity[tiab] OR homogeneous[tiab] OR 'internal consistency'[tiab] OR (cronbach*[tiab] AND (alpha[tiab] OR alphas[tiab])) OR (item[tiab] AND (correlation*[tiab] OR selection*[tiab] OR reduction*[tiab])) OR agreement[tw] OR precision[tw] OR imprecision[tw] OR 'precise values'[tw] OR test-retest[tiab] OR (test[tiab] AND retest[tiab]) OR (reliab*[tiab] AND (test[tiab] OR retest[tiab])) OR stability[tiab] OR interrater[tiab] OR inter-rater[tiab] OR intrarater[tiab] OR intra-rater[tiab] OR intertester[tiab] OR inter-tester[tiab] OR intratester[tiab] OR intra-tester[tiab] OR interobserver[tiab] OR inter-observer[tiab] OR intraobserver[tiab] OR intra-observer[tiab] OR intertechnician[tiab] OR inter-technician[tiab] OR intratechnician[tiab] OR intra-technician[tiab] OR interexaminer[tiab] OR inter-examiner[tiab] OR intraexaminer[tiab] OR intra-examiner[tiab] OR interassay[tiab] OR inter-assay[tiab] OR intraassay[tiab] OR intra-assay[tiab] OR interindividual[tiab] OR inter-individual[tiab] OR intraindividual[tiab] OR intra-individual[tiab] OR interparticipant[tiab] OR inter-participant[tiab] OR intraparticipant[tiab] OR intra-participant[tiab] OR kappa[tiab] OR kappa's[tiab] OR kappas[tiab] OR repeatab*[tw] OR ((replicab*[tw] OR repeated[tw]) AND (measure[tw] OR measures[tw] OR findings[tw] OR result[tw] OR results[tw] OR test[tw] OR tests[tw])) OR generaliza*[tiab] OR generalisa*[tiab] OR concordance[tiab] OR (intraclass[tiab] AND correlation*[tiab]) OR discriminative[tiab] OR 'known group'[tiab] OR 'factor analysis'[tiab] OR 'factor analyses'[tiab] OR 'factor structure'[tiab] OR 'factor structures'[tiab] OR dimension*[tiab] OR subscale*[tiab] OR (multitrait[tiab] AND scaling[tiab] AND (analysis[tiab] OR analyses[tiab])) OR 'item discriminant'[tiab] OR 'interscale correlation*'[tiab] OR error[tiab] OR errors[tiab] OR 'individual variability'[tiab] OR 'interval variability'[tiab] OR 'rate variability'[tiab] OR (variability[tiab] AND (analysis[tiab] OR values[tiab])) OR (uncertainty[tiab] AND (measurement[tiab] OR measuring[tiab])) OR 'standard error of measurement'[tiab] OR sensitiv*[tiab] OR responsive*[tiab] OR (limit[tiab] AND detection[tiab]) OR 'minimal detectable concentration'[tiab] OR interpretab*[tiab] OR ((minimal[tiab] OR minimally[tiab] OR clinical[tiab] OR clinically[tiab]) AND (important[tiab] OR significant[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR (small*[tiab] AND (real[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR 'meaningful change'[tiab] OR 'ceiling effect'[tiab] OR 'floor effect'[tiab] OR 'Item response model'[tiab] OR IRT[tiab] OR Rasch[tiab] OR 'Differential item functioning'[tiab] OR DIF[tiab] OR 'computer adaptive testing'[tiab] OR 'item bank'[tiab] OR 'cross-cultural equivalence'[tiab])	1 AND 2 AND 3

4	Exclusion filter	(‘addresses’[Publication Type] OR ‘biography’[Publication Type] OR ‘case reports’[Publication Type] OR ‘comment’[Publication Type] OR ‘directory’[Publication Type] OR ‘editorial’[Publication Type] OR ‘festschrift’[Publication Type] OR ‘interview’[Publication Type] OR ‘lectures’[Publication Type] OR ‘legal cases’[Publication Type] OR ‘legislation’[Publication Type] OR ‘letter’[Publication Type] OR ‘news’[Publication Type] OR ‘newspaper article’[Publication Type] OR ‘patient education handout’[Publication Type] OR ‘popular works’[Publication Type] OR ‘congresses’[Publication Type] OR ‘consensus development conference’[Publication Type] OR ‘consensus development conference, nih’[Publication Type] OR ‘practice guideline’[Publication Type]) NOT (‘animals’[MeSH Terms] NOT ‘humans’[MeSH Terms])	1 AND 2 AND 3 AND NOT 4
Results: 2964			

Appendix Q: Summary of findings per language version (study six)

Questionnaire	ROB	Inconsistency	Imprecision	Indirectness	QUALITY OF EVIDENCE (High, moderate, low, very low)	OVERALL RATING + / - / ?
VISA-G						
Content validity - relevance	Extremely serious	No	NA	No	Very low	+
Content validity - comprehensiveness	Extremely serious	No	NA	No	Very low	+
Content validity - comprehensibility	Very serious	No	NA	No	Low	+
Structural validity	Very serious	NA[1]	Serious	No	?	?
Internal consistency	Very serious	NA[1]	Serious	No	?	?
Cross-cultural validity	NT	NT	NT	NT	?	?
Reliability	Very serious	NA[1]	Very serious	No	Very low	+
Measurement error	Very serious	NA[1]	Very serious	No	Very low	+
Construct validity (comparison with outcome measures)	Serious	NA[1]	Serious	No	Low	-
Construct validity (known group)	No	NA[1]	Serious	No	Moderate	+
Responsiveness (before and after intervention)	No	NA[1]	Serious	Serious	Low	+
Responsiveness (comparison with other outcome measures)	Serious	NA[1]	Serious	Serious	Very low	+
VISA-G.BR						
Content validity - relevance	Extremely serious	No	NA	No	Very low	+
Content validity - comprehensiveness	Extremely serious	No	NA	No	Very low	+
Content validity - comprehensibility	Extremely serious	No	NA	No	Very low	+
Structural validity	NT	NT	NT	NT	?	?
Internal consistency	Very serious	NA[1]	Serious	No	?	?
Cross-cultural validity	NT	NT	NT	NT	?	?

Reliability	Very serious	NA[1]	Serious	No	Very low	+
Measurement error	Very serious	NA[1]	Serious	No	Very low	+
Construct validity (comparison with outcome measures)	Serious	NA[1]	Serious	No	Low	-
Construct validity (known group)	NT	NT	NT	NT	?	?
Responsiveness (before and after intervention)	Extremely serious	NA[1]	Serious	No	Very low	-
Responsiveness (comparison with other outcome measures)	NT	NT	NT	NT	?	?
VISA-G.F						
Content validity - relevance	Extremely serious	No	NA	No	Very low	+
Content validity - comprehensiveness	Extremely serious	No	NA	No	Very low	+
Content validity - comprehensibility	Very serious	No	NA	No	Low	+
Structural validity	NT	NT	NT	NT	?	?
Internal consistency	Very serious	NA[1]	Serious	No	?	?
Cross-cultural validity	NT	NT	NT	NT	?	?
Reliability	Extremely serious	NA[1]	No	Very serious	Very low	+
Measurement error	Extremely serious	NA[1]	Serious	No	Very low	+
Construct validity (comparison with outcome measures)	Serious	NA[1]	No	No	Moderate	-
Construct validity (known group)	No	NA[1]	No	No	High	+
Responsiveness (before and after intervention)	NT	NT	NT	NT	?	?
Responsiveness (comparison with other outcome measures)	NT	NT	NT	NT	?	?
VISA-G.I						
Content validity - relevance	Extremely serious	No	NA	No	Very low	+
Content validity - comprehensiveness	Extremely serious	No	NA	No	Very low	+

Content validity - comprehensibility	Very serious	No	NA	No	Low	+
Structural validity	NT	NT	NT	NT	?	?
Internal consistency	Very serious	NA[1]	Very serious	No	?	?
Cross-cultural validity	NT	NT	NT	NT	?	?
Reliability	Very serious	NA[1]	Very serious	No	Very Low	+
Measurement error	Very serious	NA[1]	Very serious	No	Very Low	+
Construct validity (comparison with outcome measures)	Serious	NA[1]	Serious	No	Low	-
Construct validity (known group)	Serious	NA[1]	Serious	No	Low	+
Responsiveness (before and after intervention)	NT	NT	NT	NT	?	?
Responsiveness (comparison with other outcome measures)	NT	NT	NT	NT	?	?
VISA-G.DK						
Content validity - relevance	Extremely serious	No	NA	No	Very low	+
Content validity - comprehensiveness	Extremely serious	No	NA	No	Very low	+
Content validity - comprehensibility	Very serious	No	NA	No	Low	+
Structural validity	NT	NT	NT	NT	?	?
Internal consistency	Very serious	NA[1]	Very serious	No	?	?
Cross-cultural validity	NT	NT	NT	NT	?	?
Reliability	Very serious	NA[1]	Very Serious	No	Very low	+
Measurement error	Very serious	NA[1]	No	Very serious	Very low	+
Construct validity (comparison with outcome measures)	NT	NT	NT	NT	?	?
Construct validity (known group)	NT	NT	NT	NT	?	?
Responsiveness (before and after intervention	NT	NT	NT	NT	?	?

Responsiveness (comparison with other outcome measures)	NT	NT	NT	NT	?	?
HOS-Brazil (ADL)						
Content validity - relevance	Extremely serious	No	NA	Serious	Very low	+
Content validity - comprehensiveness	Extremely serious	No	NA	Serious	Very low	-
Content validity - comprehensibility	Extremely serious	No	NA	Serious	Very low	+
Structural validity	NT	NT	NT	NT	?	?
Internal consistency	Very serious	NA[1]	Serious	Serious	?	?
Cross-cultural validity	NT	NT	NT	NT	?	?
Measurement error	Serious	No	Serious	Serious	Very low	-
Reliability (test-retest)	Serious	No	Serious	Serious	Very low	+
Construct validity (comparison with other outcome measures)	Serious	No	Serious	Serious	Very low	-
Responsiveness	NT	NT	NT	NT	?	?
HOS-Brazil (Sport)						
Content validity - relevance	Extremely serious	No	NA	Serious	Very low	+
Content validity - comprehensiveness	Extremely serious	No	NA	Serious	Very low	-
Content validity - comprehensibility	Extremely serious	No	NA	Serious	Very low	+/-
Structural validity	NT	NT	NT	NT	?	?
Internal consistency	Very serious	NA[1]	Serious	Serious	?	?
Cross-cultural validity	NT	NT	NT	NT	?	?
Measurement error	Serious	No	Serious	Serious	Very low	+
Reliability (test-retest)	Serious	No	Serious	Serious	Very low	+
Construct validity (hypothesis testing)	Serious	No	Serious	Serious	Very low	-
Responsiveness	NT	NT	NT	NT	?	?

Quality level:

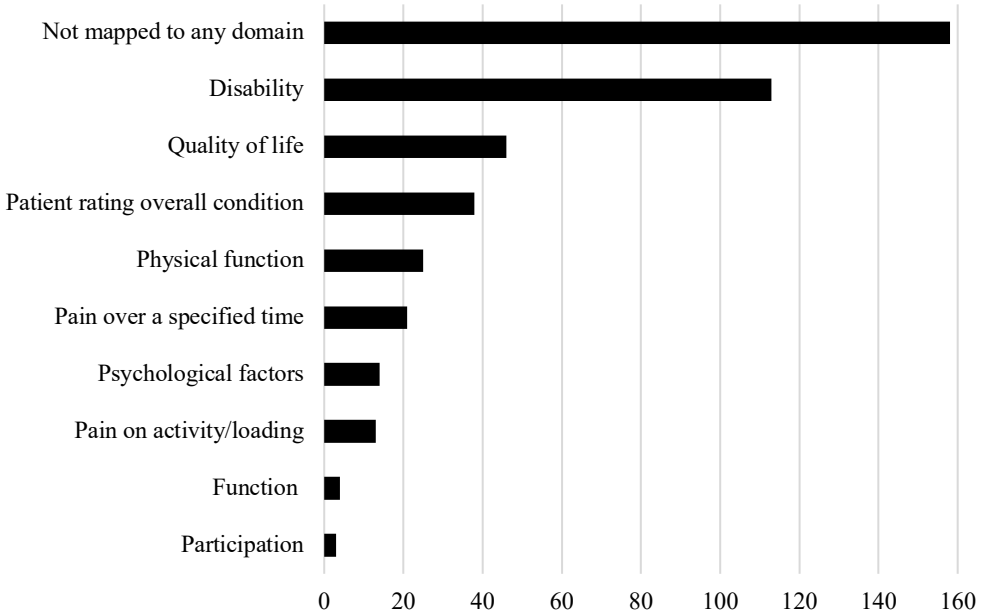
- High –very confident that the true measurement property lies close to that of the estimate of the measurement property
- Moderate –moderately confident in the measurement property estimate: the true measurement property is likely to be close to the estimate of the measurement property, but there is a possibility that it is substantially different
- Low – confidence in the measurement property estimate is limited: the true measurement property may be substantially different from the estimate of the measurement property
- Very low - very little confidence in the measurement property estimate: the true measurement property is likely to be substantially different from the estimate of the measurement property
- [1]inconsistency could not be evaluated because there was only one study available

Overall rating:

- + = sufficient, - = insufficient, +/- = inconsistent, ? = indeterminate

HOS-Brazil (ADL) = Hip Outcome Score (activities of daily living), NA = not applicable, NT = not tested, VISA-G = Victorian Institute of Sport Assessment – Gluteal Tendinopathy, VISA-G.BR = Victorian Institute of Sport Assessment – Gluteal Tendinopathy Brazil, VISA-G.DK = Victorian Institute of Sport Assessment – Gluteal Tendinopathy Danish, VISA-G.F = Victorian Institute of Sport Assessment – Gluteal Tendinopathy French, VISA-G.I = Victorian Institute of Sport Assessment – Gluteal Tendinopathy Italian

Appendix R: Outcome measures mapped to tendinopathy suggested domains (study six)



Illustrating the frequency of outcome measures used in each of the core domains of tendinopathy

Appendix S: Hypothesis testing for construct validity (study six)

Study	Outcome measure	Hypotheses	Pearson's R	Interpretation of Pearson's R	Validated hypotheses (Yes/No)
Beaudart (2020)	SF-36 (physical functioning) vs VISA-G.F	Strong	0.77	Strong	Yes
Beaudart (2020)	SF-36 (social functioning) vs VISA-G.F	Weak	0.27	Weak	Yes
Beaudart (2020)	SF-36 (role limitation physical) vs VISA-G.F	Strong	0.77	Strong	Yes
Beaudart (2020)	SF-36 (role limitation emotional) vs VISA-G.F	Weak	0.35	Moderate	No
Beaudart (2020)	SF-36 (mental health) vs VISA-G.F	Weak	0.38	Moderate	No
Beaudart (2020)	SF-36 (vitality) vs VISA-G.F	Weak	0.23	Weak	Yes
Beaudart (2020)	SF-36 (body pain) vs VISA-G.F	Strong	0.72	Strong	Yes
Beaudart (2020)	SF-36 (general health) vs VISA-G.F	Weak	0.55	Strong	No
Fearon (2015)	VISA-G vs HHS	Moderate	0.02	Weak	No
Fearon (2015)	VISA-G vs ODI	Moderate	0.21	Weak	No
Minetto (2020)	VISA-G.I vs ODI	Moderate	-0.80	Strong	No
Paiva (2021)	VISA-G.BR vs ODI	Moderate	-0.77	Strong	**No**
Costa (2019)	HOS-Brazil ADL vs NAHS total score	Strong	0.87	Strong	Yes
Costa (2019)	HOS-Brazil ADL vs SF-12 Physical subscale score	Strong	0.74	Strong	Yes
Costa (2019)	HOS-Brazil ADL vs SF-12 mental subscale score	Weak	0.35	Moderate	No
Costa (2019)	HOS-Brazil Sport vs SF-12 physical subscale score	Strong	0.69	Strong	Yes
Costa (2019)	HOS-Brazil Sport vs NAHS total score	Strong	0.79	Strong	Yes
Costa (2019)	HOS-Sport vs SF-12 Mental subscale	Weak	0.34	Moderate	No

Hypothesis:
Strong correlations (≥ 0.50) would be found with instruments measuring similar constructs
Correlations with instruments measuring related, but dissimilar constructs would be lower (0.3-0.5)
Correlations with instruments measuring unrelated constructs would be low (<0.3)
ADL = activities of daily living, HHS = Harris Hip Score, HOS = Hip Outcome Score, NAHS = Non-arthritic Hip Score, ODI = Oswestry Disability Index, SF = short form-36, VISA-G = Victorian Institute of Sport Assessment - Gluteal Tendinopathy, VISA-G.BR = Victorian Institute of Sport Assessment – Gluteal Tendinopathy Brazil, VISA-G.F = Victorian Institute of Sport Assessment - Gluteal Tendinopathy French, VISA-G.I = Victorian Institute of Sport Assessment - Gluteal Tendinopathy Italian
[1] Data not reported on femoroacetabular impingement and gluteal tendinopathy patients separately

Appendix T: Interpretability (study six)

Source	Outcome measure	Language version	Distribution of scores in the study population	Percentage of missing items and percentage of missing total scores	Floor and ceiling effects	Scores and change scores available for relevant (sub)groups	Minimal important change or minimal important difference	Information on response shift
Beaudart (2020)	VISA-G.F	French	Mean (baseline): 60.5 Range: 43-71	No missing items	None of the GTPS participants obtained the lowest or the highest score to the scale	NR	NR	NR
Ebert (2019)	VISA-G	English	Mean (baseline): 43 (SD 15.0) 12-month post intervention: mean: 78.7 (SD 14.1)	NR	Only one patient attained (1.8%) the maximum VISA-G score at 12 months	Change in VISA-G score from pre-surgery (mean (SD)) 3 months: 22.1 (SD17.2) 6 months:28.0 (15.6SD) 12 months: 35.2 (SD 16.8)	MIC: 29/100 (GRC ≥4) 22/100 (GRC≥3)	SRM: mean (95% CI) 3 months:1.28 (0.88 to 1.65) 6 months 1.80 (95% CI :1.47 to 2.16) 12 months:2.13 (95% CI 1.75 to 2.63)
Fearon (2015)	VISA-G	English	Mean (baseline): 47.00 (95% CI 42.62 to 50.18)	No missing items	No floor or ceiling effects identified.	NR	NR	NA
Jorgensen (2020)	VISA-G.DK	Danish	Mean (baseline): 61.94 (SD +/- 5.78) Range: 48- 77	NR	No floor and ceiling effects identified	NR	NR	NR
Minetto (2020)	VISA-G.I	Italian	Median: 54 (34.2-63.0)[1]	NR	No floor or ceiling effects	NR	NR	NR

					identified for the total score and for five of the eight items (items number 1, 3, 4, 5, 8). Floor effect identified for items 2 and 7. Ceiling effect identified for items 2 and 6			
Paiva (2021)	VISA-G.BR	Brazilian Portuguese	Mean (baseline): 62.82 (15.75SD) Range: 47.07-78.57	NR	No floor or ceiling effect identified	Change in VISA-G.BR score post 30 days of uncontrolled intervention[2] = 2.86 (/100)	NR	NR
Costa (2018)	HOS-Brazil (ADL)	Brazilian Portuguese	HOS-Brazil (ADL): 15.8-65.1	NR	No questionnaire exhibited scores of zero or 100 (maximum score)	NR	MCID: HOS-Brazil (ADL): 4.6	NR
Costa (2018)	HOS-Brazil (Sport)	Brazilian Portuguese	HOS-Brazil (Sport): 2.8-97.2	NR	No questionnaire exhibited scores of zero or 100 (maximum score)	NR	MCID: HOS-Brazil (Sport): 5.5	NR

[1]Median 1st and 3rd quartile scores of VISA-G.I
[2]Uncontrolled intervention
HOS-Brazil (ADL) = Hip Outcome Score-Brazil (Activities of Daily Living), HOS-Brazil (Sport) = Hip Outcome Score-Brazil (Sport), GTPS = Greater Trochanteric Pain Syndrome, MIC = minimally important change, MCID = minimal clinically important difference, NR = not reported, SD = standard deviation, SRM = standard response mean, VISA-G = Victorian Institute of Sport – Gluteal Tendinopathy, VISA-G.BR = Victorian Institute of Sport Assessment – Gluteal Tendinopathy Brazil, VISA-G = Victorian Institute of Sport – Gluteal Tendinopathy Danish, VISA-G = Victorian Institute of Sport – Gluteal Tendinopathy French, VISA-G = Victorian Institute of Sport – Gluteal Tendinopathy Italian

Appendix U: Feasibility (study six)

Source	HOS-Brazil	VISA-G, VISA-G.BR, VISA-G.DK, VISA-G.F, VISA-G.I
Patient's comprehensibility	NIA	Assessed in English developmental study for question clarity Item 8 – difficulty interpreting noted in Danish version. Added clarification during development.
Clinician's comprehensibility	NIA	Assessed for clarity in English version
Type and ease of administration	Paper based or online No additional information available	Paper based or online No additional information available
Length of the instrument	28 items (ADL:19 items, Sport: 9 items)	8 items
Completion time (minutes)[1]	NIA	Not reported in VISA-G, VISA-G.F, VISA-G.I, VISA-G.BR Reported in Danish version only: First completion: VISA-G.DK: 2:36–9:56 (±4:45) Second completion: 2:12–8:38 (±3:49)
Patient's required mental and physical ability level	NIA	NIA
Ease of standardisation	NIA	NIA
Ease of score calculation	19 items ADL scale 9 items Sport scale Each section scored on Likert scale from 4-0 or N/A (with 4 being no difficulty at all and 0 being unable to perform). Calculated as separate sub scores or a total score and expressed as a percentage	Addition of items 1 to 8 Item 8 involves 3 sections (section A, section B or section C). Only one of section A, B or C should be completed. Section A is scored out of 10, Section B out of 20 and section C out of 30. Total score possible = 100 (all versions)
Copyright	No	No
Cost of an instrument	Free	Free
Required equipment	No equipment	No equipment
Availability in different settings	Yes	Yes
Regulatory agency's requirement for approval	Not required	Not required

ADL = activity of daily living, HOS-Brazil = Hip Outcome Score Brazil (Activities of Daily Living), NIA = no information available, VISA-G = Victorian Institute of Sport – Gluteal tendinopathy, VISA-G.BR = Victorian Institute of Sport Assessment – Gluteal Tendinopathy Brazil, VISA-G.DK = Victorian Institute of Sport – Gluteal Tendinopathy Danish, VISA-G.F = Victorian Institute of Sport – Gluteal Tendinopathy French, VISA-G.I = Victorian Institute of Sport – Gluteal Tendinopathy Italian.

Appendix V: Ethical approval for Covid-19 impacted studies

THE UNIVERSITY OF QUEENSLAND

Institutional Human Research Ethics Approval

Project Title:	Structural and Physical Impairments in Proximal Hamstring Tendinopathy
Chief Investigator:	Mr Anthony Nasser
Supervisor:	Dr Adam Semciw, Prof Bill Vicenzino, Dr Alison Grimaldi
Co-Investigator(s):	None
School(s):	School of Health and Rehabilitation Sciences
Approval Number:	2017001715
Granting Agency/Degree:	MPhil
Duration:	5th November 2019

Comments/Conditions:

HREA Application Form, AN00124, 16/01/2018
Supervisor Support Email, 22/01/2018
Participant information Sheet, 16/01/2018
Participant Information – Healthy Controls, 16/01/2018
Appendix 2 – Testing Protocol, 16/01/2018
Appendix 3 – VISA – H, 16/01/2018
Appendix 4 – Questionnaire, 16/01/2018
Appendix 5 – Lower Extremity Functional Scale (LEFS), 16/01/2018
Appendix 6 – International Physical Activity Questionnaire, 16/01/2018
Consent Form, Version 3 10/02/2018
Protocol, 16/01/2018
Screening procedure, 16/01/2018

Note: if this approval is for amendments to an already approved protocol for which a UQ Clinical Trials Protection/Insurance Form was originally submitted, then the researchers must directly notify the UQ Insurance Office of any changes to that Form and Participant Information Sheets & Consent Forms as a result of the amendments, before action.

Name of responsible Committee:
University of Queensland Human Research Ethics Committee A
This project complies with the provisions contained in the *National Statement on Ethical Conduct in Human Research* and complies with the regulations governing experimentation on humans.

Name of Ethics Committee representative:
Professor Emerita Gina Geffen
Chairperson
University of Queensland Human Research Ethics Committee A
Registration: EC00456

12/02/2018

Signature Date _____

Appendix W: Sports Kongress poster

Proximal Hamstring Tendinopathy: A Systematic Review of Interventions

Mr Anthony Nasser, Prof Bill Vicenzino, Dr Alison Grimaldi, Dr Adam Semciw

The University Of Queensland AUSTRALIA

BACKGROUND

INTRODUCTION

Proximal hamstring tendinopathy (PHT)

- Is a recognised cause of persistent buttock pain in the athletic and non-athletic population.
- Was first described as "the hamstring syndrome" by Puranen in 1988. Other synonyms include hamstring tendinitis and high hamstring tendinopathy.
- PHT is classified as an insertional tendinopathy, with the cardinal feature being pain at the proximal attachment of the hamstring.
- The evidence of the effect of different treatments is currently unknown, but must be evaluated to assist in treatment selection.

AIM

To evaluate the effect of non-surgical and surgical management on:

- Symptoms
- Physical function
- Quality of life

METHODS

DESIGN

Search Strategy

- MEDLINE, CINAHL, EMBASE, SPORTDiscus and PUBMED

PHT	Tendinopathy
"Hamstring"	"Tendinopathy"
"Biceps femoris"	"Tendinosis"
"Semitendinos*"	"Tendinitis"
"Semimembranos*"	Tear

Figure- 1 - Summarising search strategy

SELECTION CRITERIA

Included

- Studies of level I-IV evidence:
- Randomised controlled trials, prospective comparative cohort trials, case-control trials and retrospective case series with 2 or more subjects.

Excluded

- level V evidence: letters to the editor, isolated single case reports, expert opinion reviews, systematic reviews, biomechanical reports, narrative reviews

RISK OF BIAS

Epidemological appraisal

- An appraisal instrument was developed to critically appraise included studies.

Key findings included

- Limited use of validated outcome measures in surgical studies
- Blinding of assessors of data was an issue across studies

Reporting

- Reporting of interventions and control groups was assessed using the TIDieR item checklist.
- Results can be seen in figure 3 below.

RESULTS

SEARCH RESULTS

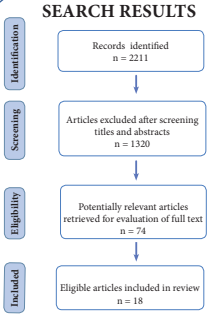
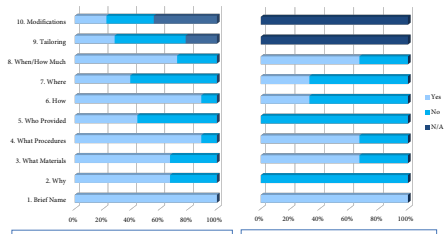


Figure 2- Flow diagram outlining study selection

Summary of included studies	
Studies included	18
Total participants	582 (227F)
Age (range)	13-71 yrs

Figure- 3 - table- summarising included studies

REPORTING OF INTERVENTIONS



Percentage of trials achieving each TIDieR item in intervention groups

Percentage of trials achieving each TIDieR item in control groups

Figure 4 - Summary of reporting of interventions in experimental group and control groups

Key Points

- Completeness of reporting interventions was investigated using the TIDieR tool.
- The completeness of reporting was higher in the intervention groups than control groups.
- The percentage of studies that satisfactorily reported each item ranged from 100% to 38% in the intervention group.
- The percentage of studies that satisfactorily reported each item ranged from 100% to 33% in the control group.

TREATMENT TYPE

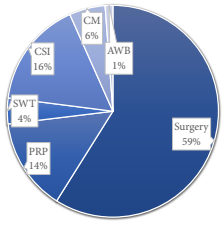


Figure 5 – Summary of number of patients receiving each treatment type in included studies. AWB = autologous whole blood, CSI = corticosteroid injection, SWT = shockwave therapy, PRP = platelet-rich plasma, CM = conservative management

Key Points

- Surgical management was the most common treatment type with 344 patients undergoing surgery.
- 67 patients received PRP treatment.
- Only a small portion of data exists on conservative management.

PRP: STANDARDISED MEAN CHANGE IN PAIN

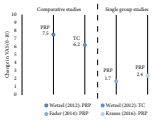


Figure 6 – Graph demonstrating the change in pain after PRP (mean difference in VAS score), PRP = platelet-rich plasma, TC = traditional care.

Key Points

- 3 studies investigating PRP reported on change in pain (VAS) over time.
- Only 1 study had a comparison group (Wetzel 2012). It showed only minor differences between PRP and the traditional care group at 12 weeks (0.7/10)
- Randomised comparitive studies, with larger sample sizes are required to determine the effect of PRP on pain.

INTERVENTIONS: CHANGE IN PAIN

Author (year)	Follow up
Wetzel (2012)	12 weeks
Cacchio (2011)	6 months
Cacchio (2011)*	12 months

Author	Follow up (mean)
Young (2008)	53 months
Benazzo (2013)	71.3 months
Krauss (2016)	12 weeks

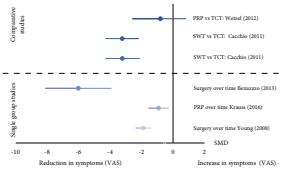


Figure 7 - Effect size (SMD) of comparative studies and single group studies on pain (VAS). In comparative studies the treatment listed first favours the intervention. PRP = platelet-rich plasma, TCT = traditional care, SWT = shockwave therapy.

Key Points

- Effect size on pain (VAS) ranged from -0.8 to -3.23 in groups with a comparison intervention.
- In groups without a comparison group, effect sizes ranged from -6.02 to -0.90.
- All treatments had a positive effect on pain (VAS), however the magnitude varied.
- Comparitive studies with larger sample sizes will assist in understanding the optimal treatment.

PATIENT SATISFACTION OF SURGERY

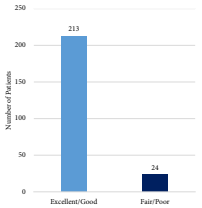


Figure 8– Summary of patient satisfaction of surgery

Key Points

- 4 of 8 surgical studies reported on subjective patient success
- Almost 90% had either a excellent or a good results
- Different surgical methods were used across studies
- No surgical studies included had a control group

SUMMARY

KEY MESSAGES

- While 18 studies were included in the systematic review, most studies were found to be of low level evidence.
- Surgery was the most reported on intervention (8 studies). All surgical studies demonstrated improvement in symptoms and physical function over time. Unfortunately, all were retrospective case series (level 4 evidence). Comparative studies are required to distinguish the effect of surgery against time, and against other more conservative treatment options.
- Adverse outcomes occurred in 15% of patients undergoing surgery.
- There was 1 high quality study on conservative management, which demonstrated a poor effect on pain and physical function compared to shockwave therapy. Results must be interpreted with caution due to the small sample size and lack of long term follow-up. Futher studies also need to be conducted to determine the optimal conservative management plan.
- A lack of use of validated outcome measures limited comparison of results between studies.
- Results of the systematic review highlight the need for further high quality research in people with PHT, to assist health practitioners in treating this debilitating condition.

References

Aldridge (2012). J Bone Joint Surg Br, 94(5), 660-662.
Benazzo (2013). Journal Of Orthopaedics And Traumatology, 14(2), 83-88.
Boeuman (2013). Am J Sports Med, 41(6), 1363-1371.
Cacchio (2011). Am J Sports Med, 39(1), 146-153.
Davenport (2015). J Ultrasound Med, 34(8), 1455-1463.
Fadier (2014). Muscles Ligaments Tendons J, 4(4), 461-466.
Jacobson (2015). J Ultrasound Med, 34(11), 2029-2035.
Lefevre (2013). Ortho & Trauma, Surg & Res, 97(5S, 0S2), 235-240.
Lempainen (2006). British Journal of Sports Medicine, 40(8), 688-691.
Lempainen (2009). Am J Sports Med, 37(4), 727-734.
Maintner (2013). Pm r, 5(3), 169-175.
Nicholson (2014). Orthop J Sports Med, 2(3), 2325967114526135.
Wetzel (2013) Orthopedics, 36(1), e64-70.
White (2011). J Chiropr Med, 10(2), 93-99
Young (2008). Am J Sports Med, 36(12), 2372-2378.
Zissen, (2010) American Journal of Roentgenology, 195(4), 993-998.

Scan QR code for additional information on included studies

@TonyNasserPT

University of Queensland
School of Health and Rehabilitation Sciences

Correspondence:
anthony.nasser@uq.edu.au

Appendix X: Sports Medicine Australia poster

Proximal Hamstring Tendinopathy: a Systematic Review of Interventions

Mr Anthony Nasser, Prof Bill Vicenzino, Dr Alison Grimaldi, Dr Adam Semciw

The University Of Queensland AUSTRALIA

BACKGROUND

PROXIMAL HAMSTRING TENDINOPATHY (PHT)

- Is a recognised cause of persistent buttock pain in the athletic and non-athletic population
- Was first described as "the hamstring syndrome" in 1988
- It's primary feature is localised pain at the proximal attachment of the hamstring
- Therefore evidence of the effect of different interventions must be evaluated to assist treatment selection

AIMS

To evaluate the effect of interventions on:
1. Symptoms
2. Physical Function
3. Quality of Life

METHODS

SELECTION CRITERIA

Search Strategy
- MEDLINE, CINAHL, EMBASE, SPORTDiscus and PUBMED

Included
- Studies of level I-IV evidence:
- Randomised controlled trials, prospective comparative cohort trials, case-control trials and retrospective case series with 10 or more subjects

Excluded
- Level V evidence: letters to editor, isolated single case reports, expert opinion reviews, systematic reviews, biomechanical reports, narrative reviews

RISK OF BIAS

Epidemological appraisal
- Appraisal instrument was developed to appraise included studies.

Reporting
- Reporting was assessed with the TIDieR checklist

RESULTS

SEARCH

Records identified
n = 1692

↓

Articles included after screening titles and abstracts
n = 882

↓

Potentially relevant articles retrieved in full text
n = 34

↓

Studies included	16
Total participants	537(213F)
Age (range)	13-71 yrs

Figure 3 - summarising included studies

REPORTING

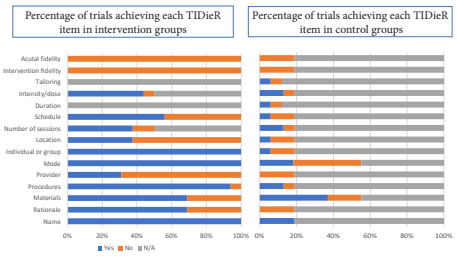


Figure 4 - Summary of reporting of interventions in experimental and control groups

- Higher quality studies (RCTs) reported on interventions more completely
- Post-surgical protocols are not described adequately (protocol and adherence)
- Future research should prospectively report post-surgical protocols and devise methods to report adherence of post-surgical rehabilitation

TREATMENT TYPE

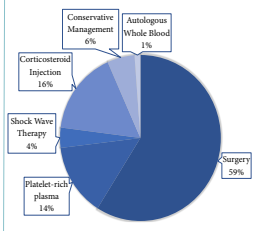


Figure 5 – Summary of number of patients receiving each treatment type in included studies.

- The most common intervention reported on was surgery (n=8)
- While 5 studies reported on platelet-rich plasma injections, sample sizes tended to be small (n<15)

RESULTS: COMPARATIVE STUDIES

Figure 6 - Effect sizes (SMD) of comparative studies on physical function. Intervnetion listed first favours that treatment PRP = platelet-rich plasma, CM = conservative management, SWT = shockwave therapy, NRPS = Nirsch Phase Rating Scale, MHHS = Modified Hip Harris Score

Figure 7 – Effect sizes (SMD) of comparative studies on symptoms (VAS). Intervention listed first favours that treatment. SWT = Shockwave therapy CM = conservative management, VAS = visual analogue scale

- Of the 16 included studies, only 2 had comparison groups
- While SWT had a significant effect on physical function and symptoms over conservative management, the overall quality of evidence (GRADE) was classified as low, as there was only 1 study on this intervention, with a small sample size

SATISFACTION OF SURGERY

Figure 8– Summary of patient satisfaction of surgery. 50% of studies on surgery reported on satisfaction

- Whilst almost 90% of participants reported excellent or good results, the quality of studies was low
- No surgical studies included had a comparison group

SUMMARY

KEY MESSAGES

The main findings that arise from this systematic review are:
1. There is a serious lack of high quality trials in proximal hamstring tendinopathy, particularly on conservative management
2. There is a lack of consistency in use of validated patient reported outcome measures between studies
3. Reporting on interventions in studies, particularly post-surgical protocols, is poor

As there is only low level evidence to guide treatment for PHT, until high quality evidence exists, a pragmatic approach would be to trial rehabilitation programs that have been successful in other lower limb tendinopathies

References

Aldridge (2012). J Bone Joint Surg Br, 94(5), 660-662.
Benazzo (2013). Journal Of Orthopaedics And Traumatology, 14(2), 83-89.
Bowman (2013). Am J Sports Med, 41(6), 1363-1371.
Cacchio (2011). Am J Sports Med, 39(1), 146-153.
Davenport (2015). J Ultrasound Med, 34(8), 1455-1463.
Fader (2014). Muscles Ligaments Tendons J, 4(4), 461-466.
Jacobson (2015). J Ultrasound Med, 34(11), 2029-2035.
Lefevre (2013). Ortho & Trauma, Surg & Res OTSR, 99(2), 235-240.
Lempainen (2006). British Journal of Sports Medicine, 40(8), 688-691.
Lempainen (2009) Am J Sports Med, 37(4), 727-734.
Mautner (2013). Pm r, 5(3), 169-175
Nicholson (2014). Orthop J Sports Med, 2(3), 2325967114526135.
Wetzel (2013) Orthopedics, 36(1), e64-70
White (2011). J Chiropr Med, 10(2), 93-99
Young (2008). Am J Sports Med, 36(12), 2372-2378.
Zissen, (2010) American Journal of Roentgenology, 195(4), 993-998.

@TonyNasserPT

University of Queensland
School of Health and Rehabilitation Sciences

anthony.nasser@uq.edu.au

References

1. Maffulli N, Khan KM, Puddu G. Overuse tendon conditions: time to change a confusing terminology. Arthroscopy. 1998;14(8):840-43.
2. Scott A, Squier K, Alfredson H, Bahr R, Cook JL, Coombes B, et al. ICON 2019: International Scientific Tendinopathy Symposium Consensus: Clinical Terminology. British Journal of Sports Medicine. 2020;54(5):260-62.
3. Dean BJF, Dakin SG, Millar NL, Carr AJ. Review: Emerging concepts in the pathogenesis of tendinopathy. Surgeon. 2017;15(6):349-54.
4. Rio E, Moseley L, Purdam C, Samiric T, Kidgell D, Pearce AJ, et al. The Pain of Tendinopathy: Physiological or Pathophysiological? Sports Medicine. 2014;44(1):9-23.
5. Millar NL, Silbernagel KG, Thorborg K, Kirwan PD, Galatz LM, Abrams GD, et al. Tendinopathy. Nat Rev Dis Primers. 2021;7(1):1-21.
6. Albers IS, Zwerver J, Diercks RL, Dekker JH, Van den Akker-Scheek I. Incidence and prevalence of lower extremity tendinopathy in a Dutch general practice population: a cross sectional study. BMC Musculoskeletal Disorders. 2016;17(1):1-6.
7. Cardoso TB, Pizzari T, Kinsella R, Hope D, Cook JL. Current trends in tendinopathy management. Best Practice & Research Clinical Rheumatology. 2019;33(1):122-40.
8. Puranen J, Orava S. The hamstring syndrome: a new diagnosis of gluteal sciatic pain. American Journal of Sports Medicine. 1988;16(5):517-21.
9. Zissen MH, Wallace G, Stevens KJ, Fredericson M, Beaulieu CF. High hamstring tendinopathy: MRI and ultrasound imaging and therapeutic efficacy of percutaneous corticosteroid injection. American Journal of Roentgenology. 2010;195(4):993-98.
10. Lipman K, Wang C, Ting K, Soo C, Zheng Z. Tendinopathy: injury, repair, and current exploration. Drug Des Devel Ther. 2018;12:591-603.
11. Florit D, Pedret C, Casals M, Malliaras P, Sugimoto D, Rodas G. Incidence of Tendinopathy in Team Sports in a Multidisciplinary Sports Club Over 8 Seasons. J Sports Sci Med. 2019;18(4):780-88.
12. Kjaer M, Langberg H, Miller BF, Boushel R, Crameri R, Koskinen S, et al. Metabolic activity and collagen turnover in human tendon in response to physical activity. J Musculoskelet Neuronal Interact. 2005;5(1):41-52.
13. Loiacono C, Palermi S, Massa B, Belviso I, Romano V, Gregorio AD, et al. Tendinopathy: Pathophysiology, Therapeutic Options, and Role of Nutraceutics. A Narrative Literature Review. Medicina (Kaunas). 2019;55(8):447.
14. Vogel K. Tendon structure and response to changing mechanical load. Journal of musculoskeletal & neuronal interactions. 2004;3:323-25.
15. Sharma P, Maffulli N. The Future: Rehabilitation, Gene Therapy, Optimization of Healing. Foot and Ankle Clinics. 2005;10(2):383-97.
16. Magnusson SP, Langberg H, Kjaer M. The pathogenesis of tendinopathy: balancing the response to loading. Nat Rev Rheumatol. 2010;6(5):262-68.
17. Sharma P, Maffulli N. Biology of tendon injury: healing, modeling and remodeling. J Musculoskelet Neuronal Interact. 2006;6(2):181-90.
18. Rees JD, Wilson AM, Wolman RL. Current concepts in the management of tendon disorders. Rheumatology. 2006;45(5):508-21.
19. Zwerver J, Bredeweg SW, van den Akker-Scheek I. Prevalence of Jumper's knee among nonelite athletes from different sports: a cross-sectional survey. Am J Sports Med. 2011;39(9):1984-88.

20. Visnes H, Hoksrud A, Cook J, Bahr R. No effect of eccentric training on jumper's knee in volleyball players during the competitive season: a randomized clinical trial. Clin J Sport Med. 2005;15(4):227-34.
21. Kujala UM, Sarna S, Kaprio J. Cumulative incidence of achilles tendon rupture and tendinopathy in male former elite athletes. Clin J Sport Med. 2005;15(3):133-35.
22. O'Neill S, Watson PJ, Barry S. A Delphi study of risk factors for Achilles Tendinopathy - opinions of world tendon experts. Int J Sports Phys Ther. 2016;11(5):684-97.
23. van der Vlist AC, Breda SJ, Oei EHG, Verhaar JAN, de Vos R-J. Clinical risk factors for Achilles tendinopathy: a systematic review. British Journal of Sports Medicine. 2019;53(21):1352-61.
24. Plinsinga ML, Coombes BK, Mellor R, Nicolson P, Grimaldi A, Hodges P, et al. Psychological factors not strength deficits are associated with severity of gluteal tendinopathy: A cross-sectional study. Eur J Pain. 2018;22(6):1124-33.
25. Leadbetter WB. Cell-matrix response in tendon injury. Clin Sports Med. 1992;11(3):533-78.
26. Schubert TEO, Weidler C, Lerch K, Hofstädter F, Straub RH. Achilles tendinosis is associated with sprouting of substance P positive nerve fibres. Ann Rheum Dis. 2005;64(7):1083-86.
27. Cook JL, Rio E, Purdam CR, Docking SI. Revisiting the continuum model of tendon pathology: what is its merit in clinical practice and research? Br J Sports Med. 2016;50(19):1187-91.
28. Yuan J, Wang MX, Murrell GA. Cell death and tendinopathy. Clin Sports Med. 2003;22(4):693-701.
29. Yuan J, Murrell GA, Wei AQ, Wang MX. Apoptosis in rotator cuff tendonopathy. J Orthop Res. 2002;20(6):1372-79.
30. Cook JL, Purdam CR. Is tendon pathology a continuum? A pathology model to explain the clinical presentation of load-induced tendinopathy. Br J Sports Med. 2009;43(6):409-16.
31. Almekinders LC, Weinhold PS, Maffulli N. Compression etiology in tendinopathy. Clin Sports Med. 2003;22(4):703-10.
32. Cook JL, Purdam C. Is compressive load a factor in the development of tendinopathy? Br J Sports Med. 2012;46(3):163-68.
33. Benjamin M, Ralphs JR. Entheses--the bony attachments of tendons and ligaments. Ital J Anat Embryol. 2001;106(2):151-57.
34. Benjamin M, McGonagle D. The anatomical basis for disease localisation in seronegative spondyloarthropathy at entheses and related sites. J Anat. 2001;199(Pt 5):503-26.
35. Soslowsky LJ, Thomopoulos S, Tun S, Flanagan CL, Keefer CC, Mastaw J, et al. Neer Award 1999. Overuse activity injures the supraspinatus tendon in an animal model: a histologic and biomechanical study. J Shoulder Elbow Surg. 2000;9(2):79-84.
36. Soslowsky LJ, Thomopoulos S, Esmail A, Flanagan CL, Iannotti JP, Williamson JD, 3rd, et al. Rotator cuff tendinosis in an animal model: role of extrinsic and overuse factors. Ann Biomed Eng. 2002;30(8):1057-63.
37. Cowan RM, Ganderton CL, Cook J, Semciw AI, Long DM, Pizzari T. Does Menopausal Hormone Therapy, Exercise, or Both Improve Pain and Function in Postmenopausal Women With Greater Trochanteric Pain Syndrome? A 2 × 2 Factorial Randomized Clinical Trial. Am J Sports Med. 2022;50(2):515-25.
38. Ganderton C, Semciw A, Cook J, Moreira E, Pizzari T. Gluteal Loading Versus Sham Exercises to Improve Pain and Dysfunction in Postmenopausal Women with Greater Trochanteric Pain Syndrome: A Randomized Controlled Trial. J Womens Health. 2018;27(6):815-29.

39. Mellor R, Bennell K, Grimaldi A, Nicolson P, Kasza J, Hodges P, et al. Education plus exercise versus corticosteroid injection use versus a wait and see approach on global outcome and pain from gluteal tendinopathy: prospective, single blinded, randomised clinical trial. Br J Sports Med. 2018;52(22):1464-72.
40. Malliaras P, O'Neill S. Potential risk factors leading to tendinopathy. Apunts Medicina de l'Esport. 2017;52(194):71-7.
41. Ranger TA, Wong AMY, Cook JL, Gaida JE. Is there an association between tendinopathy and diabetes mellitus? A systematic review with meta-analysis. British Journal of Sports Medicine. 2016;50(16):982.
42. Tilley BJ, Cook JL, Docking SI, Gaida JE. Is higher serum cholesterol associated with altered tendon structure or tendon pain? A systematic review. British Journal of Sports Medicine. 2015;49(23):1504-09.
43. Frizziero A, Vittadini F, Gasparre G, Masiero S. Impact of oestrogen deficiency and aging on tendon: concise review. Muscles Ligaments Tendons J. 2014;4(3):324-28.
44. Cross M, Smith E, Hoy D, Nolte S, Ackerman I, Fransen M, et al. The global burden of hip and knee osteoarthritis: estimates from the Global Burden of Disease 2010 study. Ann Rheum Dis. 2014;73(7):1323-30.
45. Fearon AM, Cook JL, Scarvell JM, Neeman T, Cormick W, Smith PN. Greater trochanteric pain syndrome negatively affects work, physical activity and quality of life: a case control study. J Arthroplasty. 2014;29(2):383-86.
46. Goom TS, Malliaras P, Reiman MP, Purdam CR. Proximal Hamstring Tendinopathy: Clinical Aspects of Assessment and Management. J Orthop Sports Phys Ther. 2016;46(6):483-93.
47. Benazzo F, Marullo M, Zanon G, Indino C, Pelillo F. Surgical management of chronic proximal hamstring tendinopathy in athletes: a 2 to 11 years of follow-up. Journal Of Orthopaedics And Traumatology: Official Journal Of The Italian Society Of Orthopaedics And Traumatology. 2013;14(2):83-9.
48. Grimaldi A, Mellor R, Hodges P, Bennell K, Wajswelner H, Vicenzino B. Gluteal Tendinopathy: A Review of Mechanisms, Assessment and Management. Sports Med. 2015;45(8):1107-19.
49. Docking SI, Ooi CC, Connell D. Tendinopathy: Is Imaging Telling Us the Entire Story? J Orthop Sports Phys Ther. 2015;45(11):842-52.
50. Sato K, Nimura A, Yamaguchi K, Akita K. Anatomical study of the proximal origin of hamstring muscles. J Orthop Sci. 2012;17(5):614-18.
51. Miller SL, Gill J, Webb GR. The proximal origin of the hamstrings and surrounding anatomy encountered during repair. A cadaveric study. J Bone Joint Surg Am. 2007;89(1):44-8.
52. Beltran L, Ghazikhanian V, Padron M, Beltran J. The proximal hamstring muscle-tendon-bone unit: a review of the normal anatomy, biomechanics, and pathophysiology. Eur J Radiol. 2012;81(12):3772-79.
53. Chumanov ES, Heiderscheit BC, Thelen DG. Hamstring musculotendon dynamics during stance and swing phases of high-speed running. Med Sci Sports Exerc. 2011;43(3):525-32.
54. Schache AG, Dorn TW, Blanch PD, Brown NA, Pandy MG. Mechanics of the human hamstring muscles during sprinting. Med Sci Sports Exerc. 2012;44(4):647-58.
55. Kalema RN, Schache AG, Williams MD, Heiderscheit B, Siqueira Trajano G, Shield AJ. Sprinting Biomechanics and Hamstring Injuries: Is There a Link? A Literature Review. Sports (Basel). 2021;9(10):141-63.

56. Van Hooren B, Bosch F. Is there really an eccentric action of the hamstrings during the swing phase of high-speed running? part I: A critical review of the literature. J Sports Sci. 2017;35(23):2313-21.
57. Lempainen L, Sarimo J, Mattila K, Vaittinen S, Orava S. Proximal hamstring tendinopathy: results of surgical management and histopathologic findings. Am J Sports Med. 2009;37(4):727-34.
58. Lempainen L, Johansson K, Banke IJ, Ranne J, Makela K, Sarimo J, et al. Expert opinion: diagnosis and treatment of proximal hamstring tendinopathy. Muscles Ligaments Tendons J. 2015;5(1):23-8.
59. Kountouris A, Cook J. Rehabilitation of Achilles and patellar tendinopathies. Best Pract Res Clin Rheumatol. 2007;21(2):295-316.
60. Lima POP, Lourenco BM, Valentim de Freitas L, Nazareno TS, Maffulli N, de Oliveira RR. The Brazilian version of the Victorian Institute of Sport Assessment - Hamstring (VISA-H) Questionnaire: Translation, cross-cultural adaptation and measurement properties. Musculoskelet Sci Pract. 2022;58:102516.
61. Cacchio A, Borra F, Severini G, Foglia A, Musarra F, Taddio N, et al. Reliability and validity of three pain provocation tests used for the diagnosis of chronic proximal hamstring tendinopathy. Br J Sports Med. 2012;46(12):883-87.
62. Martin RL, Schroder RG, Gomez-Hoyos J, Khoury AN, Palmer IJ, McGovern RP, et al. Accuracy of 3 Clinical Tests to Diagnose Proximal Hamstrings Tears With and Without Sciatic Nerve Involvement in Patients With Posterior Hip Pain. Arthroscopy. 2018;34(1):114-21.
63. Reiman M, Burgi C, Strube E, Prue K, Ray K, Elliott A, et al. The utility of clinical measures for the diagnosis of achilles tendon injuries: a systematic review with meta-analysis. J Athl Train. 2014;49(6):820-29.
64. Grimaldi A, Mellor R, Nicolson P, Hodges P, Bennell K, Vicenzino B. Utility of clinical tests to diagnose MRI-confirmed gluteal tendinopathy in patients presenting with lateral hip pain. British Journal of Sports Medicine. 2017;51(6):519-24.
65. Young IJ, van Riet RP, Bell SN. Surgical release for proximal hamstring syndrome. Am J Sports Med. 2008;36(12):2372-78.
66. Heiderscheit BC, Sherry MA, Silder A, Chumanov ES, Thelen DG. Hamstring strain injuries: recommendations for diagnosis, rehabilitation, and injury prevention. The Journal of orthopaedic and sports physical therapy. 2010;40(2):67-81.
67. Reiman MP, Loudon JK, Goode AP. Diagnostic accuracy of clinical tests for assessment of hamstring injury: a systematic review. J Orthop Sports Phys Ther. 2013;43(4):223-31.
68. Severini G, Holland D, Drumgoole A, Delahunt E, Ditroilo M. Kinematic and electromyographic analysis of the Askling L-Protocol for hamstring training. Scand J Med Sci Sports. 2018;28(12):2536-46.
69. Askling CM, Tengvar M, Saartok T, Thorstensson A. Acute first-time hamstring strains during slow-speed stretching: clinical, magnetic resonance imaging, and recovery characteristics. Am J Sports Med. 2007;35(10):1716-24.
70. Fletcher AN, Cheah JW, Nho SJ, Mather RC, 3rd. Proximal Hamstring Injuries. Clin Sports Med. 2021;40(2):339-61.
71. Bowman KF, Jr., Cohen SB, Bradley JP. Operative management of partial-thickness tears of the proximal hamstring muscles in athletes. Am J Sports Med. 2013;41(6):1363-71.
72. Battaglia PJ, D'Angelo K, Kettner NW. Posterior, Lateral, and Anterior Hip Pain Due to Musculoskeletal Origin: A Narrative Literature Review of History, Physical Examination, and Diagnostic Imaging. J Chiropr Med. 2016;15(4):281-93.

73. Caliesch R, Sattelmayer M, Reichenbach S, Zwahlen M, Hilfiker R. Diagnostic accuracy of clinical tests for cam or pincer morphology in individuals with suspected FAI syndrome: a systematic review. BMJ Open Sport & Exercise Medicine. 2020;6(1):e000772.
74. Hatem MA, Palmer IJ, Martin HD. Diagnosis and 2-year outcomes of endoscopic treatment for ischiofemoral impingement. Arthroscopy. 2015;31(2):239-46.
75. Harrast MA, Colonno D. Stress fractures in runners. Clin Sports Med. 2010;29(3):399-416.
76. Saueressig T, Owen PJ, Diemer F, Zebisch J, Belavy DL. Diagnostic Accuracy of Clusters of Pain Provocation Tests for Detecting Sacroiliac Joint Pain: Systematic Review With Meta-analysis. J Orthop Sports Phys Ther. 2021;51(9):422-31.
77. Kaur J, Singh P. Pudendal Nerve Entrapment Syndrome. StatPearls. Treasure Island (FL): StatPearls Publishing
Copyright © 2022, StatPearls Publishing LLC.; 2022.
78. Poddubnyy D, Callhoff J, Spiller I, Listing J, Braun J, Sieper J, et al. Diagnostic accuracy of inflammatory back pain for axial spondyloarthritis in rheumatological care. RMD Open. 2018;4(2):e000825.
79. Sieper J, van der Heijde D, Landewé R, Brandt J, Burgos-Vagas R, Collantes-Estevez E, et al. New criteria for inflammatory back pain in patients with chronic back pain: a real patient exercise by experts from the Assessment of SpondyloArthritis international Society (ASAS). Ann Rheum Dis. 2009;68(6):784-88.
80. Rio EK, Mc Auliffe S, Kuipers I, Girdwood M, Alfredson H, Bahr R, et al. ICON PART-T 2019–International Scientific Tendinopathy Symposium Consensus: recommended standards for reporting participant characteristics in tendinopathy research (PART-T). British Journal of Sports Medicine. 2019;54(11):627-30.
81. Lempainen L, Sarimo J, Mattila K, Orava S. Proximal hamstring tendinopathy-overview of the problem with emphasis on the surgical treatment. Operative Techniques in Sports Medicine. 2009;17(4):225-28.
82. Fredericson M, Moore W, Guillet M, Beaulieu C. High hamstring tendinopathy in runners: meeting the challenges of diagnosis, treatment, and rehabilitation. Phys Sportsmed. 2005;33(5):32-43.
83. Barton CJ, Bonanno DR, Carr J, Neal BS, Malliaras P, Franklyn-Miller A, et al. Running retraining to treat lower limb injuries: a mixed-methods study of current evidence synthesised with expert opinion. British Journal of Sports Medicine. 2016;50(9):513-26.
84. Thompson SM, Fung S, Wood DG. The prevalence of proximal hamstring pathology on MRI in the asymptomatic population. Knee Surg Sports Traumatol Arthrosc. 2017;25(1):108-11.
85. De Smet AA, Blankenbaker DG, Alsheik NH, Lindstrom MJ. MRI appearance of the proximal hamstring tendons in patients with and without symptomatic proximal hamstring tendinopathy. AJR Am J Roentgenol. 2012;198(2):418-22.
86. Bowden DJ, Byrne CA, Alkhayat A, Kavanagh EC, Eustace SJ. Differing MRI appearances of symptomatic proximal hamstring tendinopathy with ageing: a comparison of appearances in patients below and above 45 years. Clin Radiol. 2018;73(11):922-27.
87. Hernando MF, Cerezal L, Perez-Carro L, Abascal F, Canga A. Deep gluteal syndrome: anatomy, imaging, and management of sciatic nerve entrapments in the subgluteal space. Skeletal Radiol. 2015;44(7):919-34.
88. Martin HD, Kivlan BR, Palmer IJ, Martin RL. Diagnostic accuracy of clinical tests for sciatic nerve entrapment in the gluteal region. Knee Surg Sports Traumatol Arthrosc. 2014;22(4):882-88.

89. Erickson LN, Sherry MA. Rehabilitation and return to sport after hamstring strain injury. J Sport Health Sci. 2017;6(3):262-70.
90. Gidwani S, Bircher MD. Avulsion injuries of the hamstring origin - a series of 12 patients and management algorithm. Ann R Coll Surg Engl. 2007;89(4):394-99.
91. Grange S, Reurink G, Nguyen AQ, Riviera-Navarro C, Foschia C, Croisille P, et al. Location of Hamstring Injuries Based on Magnetic Resonance Imaging: A Systematic Review. Sports Health. 2022.
92. Brukner P, Cook JL, Purdam CR. Does the intramuscular tendon act like a free tendon? British Journal of Sports Medicine. 2018;52(19):1227.
93. Askling C, Saartok T, Thorstensson A. Type of acute hamstring strain affects flexibility, strength, and time to return to pre-injury level. Br J Sports Med. 2006;40(1):40-4.
94. Obey MR, Broski SM, Spinner RJ, Collins MS, Krych AJ. Anatomy of the Adductor Magnus Origin: Implications for Proximal Hamstring Injuries. Orthop J Sports Med. 2016;4(1):1-9.
95. Serner A, Weir A, Tol JL, Thorborg K, Roemer F, Guermazi A, et al. Characteristics of acute groin injuries in the adductor muscles: A detailed MRI study in athletes. Scand J Med Sci Sports. 2018;28(2):667-76.
96. Harper BE, Reveille JD. Spondyloarthritis: clinical suspicion, diagnosis, and sports. Curr Sports Med Rep. 2009;8(1):29-34.
97. Paavola M, Järvinen TAH. Paratendinopathy. Foot and Ankle Clinics. 2005;10(2):279-92.
98. Board TN, Hughes SJ, Freemont AJ. Trochanteric bursitis: the last great misnomer. Hip Int. 2014;24(6):610-5.
99. Lange J, Tvedesøe C, Lund B, Bohn MB. Low prevalence of trochanteric bursitis in patients with refractory lateral hip pain. Dan Med J. 2022;69(7).
100. Woodley SJ, Nicholson HD, Livingstone V, Doyle TC, Meikle GR, Macintosh JE, et al. Lateral hip pain: findings from magnetic resonance imaging and clinical examination. J Orthop Sports Phys Ther. 2008;38(6):313-28.
101. Ganderton C, Semciw A, Cook J, Pizzari T. Demystifying the Clinical Diagnosis of Greater Trochanteric Pain Syndrome in Women. Journal of Women's Health. 2017;26(6):633-43.
102. Freitas A, Rosa TM, Macedo Neto SL, Bandeira VC, Souto DRM, Barin FR. FABREX: A NEW CLINICAL TEST FOR DIAGNOSIS GLUTEAL TENDINOPATHY. Acta Ortop Bras. 2022;30(2):e241045.
103. Ladurner A, Fitzpatrick J, O'Donnell JM. Treatment of Gluteal Tendinopathy: A Systematic Review and Stage-Adjusted Treatment Recommendation. Orthop J Sports Med. 2021;9(7):23259671211016850.
104. Torriani M, Souto SC, Thomas BJ, Ouellette H, Bredella MA. Ischiofemoral impingement syndrome: an entity with hip pain and abnormalities of the quadratus femoris muscle. AJR Am J Roentgenol. 2009;193(1):186-90.
105. Liong SY, Whitehouse RW. Lower extremity and pelvic stress fractures in athletes. Br J Radiol. 2012;85(1016):1148-56.
106. Challoumas D, Clifford C, Kirwan P, Millar NL. How does surgery compare to sham surgery or physiotherapy as a treatment for tendinopathy? A systematic review of randomised trials. BMJ Open Sport & Exercise Medicine. 2019;5(1):e000528.
107. Kongsgaard M, Kovanen V, Aagaard P, Doessing S, Hansen P, Laursen AH, et al. Corticosteroid injections, eccentric decline squat training and heavy slow resistance training in patellar tendinopathy. Scand J Med Sci Sports. 2009;19(6):790-802.

108. Macdermid JC, Silbernagel KG. Outcome Evaluation in Tendinopathy: Foundations of Assessment and a Summary of Selected Measures. J Orthop Sports Phys Ther. 2015;45(11):950-64.
109. Fearon AM, Ganderton C, Scarvell JM, Smith PN, Neeman T, Nash C, et al. Development and validation of a VISA tendinopathy questionnaire for greater trochanteric pain syndrome, the VISA-G. Man Ther. 2015;20(6):805-13.
110. Cacchio A, De Paulis F, Maffulli N. Development and validation of a new visa questionnaire (VISA-H) for patients with proximal hamstring tendinopathy. Br J Sports Med. 2014;48(6):448-52.
111. Davenport KL, Campos JS, Nguyen J, Saboeiro G, Adler RS, Moley PJ. Ultrasound-Guided Intratendinous Injections With Platelet-Rich Plasma or Autologous Whole Blood for Treatment of Proximal Hamstring Tendinopathy: A Double-Blind Randomized Controlled Trial. Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine. 2015;34(8):1455-63.
112. Cacchio A, Rompe JD, Furia JP, Susi P, Santilli V, De Paulis F. Shockwave therapy for the treatment of chronic proximal hamstring tendinopathy in professional athletes. Am J Sports Med. 2011;39(1):146-53.
113. Murphy MC, Travers MJ, Chivers P, Debenham JR, Docking SI, Rio EK, et al. Efficacy of heavy eccentric calf training for treating mid-portion Achilles tendinopathy: a systematic review and meta-analysis. Br J Sports Med. 2019;53(17):1070-77.
114. Challoumas D, Pedret C, Biddle M, Ng NYB, Kirwan P, Cooper B, et al. Management of patellar tendinopathy: a systematic review and network meta-analysis of randomised studies. BMJ Open Sport & Exercise Medicine. 2021;7(4):e001110.
115. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. 2014;348:1687.
116. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ (Clinical research ed). 2009;339:b2700.
117. van der Eng DM, Schepers T, Goslings JC, Schep NW. Rerupture rate after early weightbearing in operative versus conservative treatment of Achilles tendon ruptures: a meta-analysis. The Journal of foot and ankle surgery : official publication of the American College of Foot and Ankle Surgeons. 2013;52(5):622-28.
118. White KE. High hamstring tendinopathy in 3 female long distance runners. J Chiropr Med. 2011;10(2):93-9.
119. Jayaseelan DJ, Moats N, Ricardo CR. Rehabilitation of proximal hamstring tendinopathy utilizing eccentric training, lumbopelvic stabilization, and trigger point dry needling: 2 case reports. J Orthop Sports Phys Ther. 2014;44(3):198-205.
120. Jacobson JA, Rubin J, Yablon CM, Kim SM, Kalume-Brigido M, Parameswaran A. Ultrasound-Guided Fenestration of Tendons About the Hip and Pelvis: Clinical Outcomes. Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine. 2015;34(11):2029-35.
121. Lawrenson PR, Crossley KM, Vicenzino BT, Hodges PW, James G, Croft KJ, et al. Muscle size and composition in people with articular hip pathology: a systematic review with meta-analysis. Osteoarthritis Cartilage. 2019;27(2):181-95.
122. Borenstein M, Hedges LV, Higgins JPT, Rothstein H. Introduction to meta-analysis2011. 150-90 p.

123. Higgins JPT, Green S, (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. 2011.
124. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale, N.J.: L. Erlbaum Associates; 1988.
125. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ (Clinical research ed). 2003;327(7414):557-60.
126. Aromataris E, Munn Z. Joanna Briggs Institute Reviewer's Manual. The Joanna Briggs Institute. 2017.
127. Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. Journal of clinical epidemiology. 2011;64(4):401-06.
128. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, et al. Grading quality of evidence and strength of recommendations. BMJ (Clinical research ed). 2004;328(7454):1490.
129. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ (Clinical research ed). 2008;336(7650):924-26.
130. Thoomes EJ, Scholten-Peeters W, Koes B, Falla D, Verhagen AP. The effectiveness of conservative treatment for patients with cervical radiculopathy: a systematic review. The Clinical journal of pain. 2013;29(12):1073-86.
131. Fader RR, Mitchell JJ, Traub S, Nichols R, Roper M, Mei Dan O, et al. Platelet-rich plasma treatment improves outcomes for chronic proximal hamstring injuries in an athletic population. Muscles Ligaments Tendons J. 2014;4(4):461-66.
132. Wetzel RJ, Patel RM, Terry MA. Platelet-rich plasma as an effective treatment for proximal hamstring injuries. Orthopedics. 2013;36(1):e64-70.
133. Levy GM, Lucas P, Hope N. Efficacy of a platelet-rich plasma injection for the treatment of proximal hamstring tendinopathy: A pilot study. Journal of science and medicine in sport. 2019;22(3):247-52.
134. Krauss J, Bodor M, Nugent R, Fredericson M. Platelet rich plasma injections in high hamstring tendinopathy. PM and R. 2014;6(9):S276.
135. Nicholson LT, DiSegna S, Newman JS, Miller SL. Fluoroscopically Guided Peritendinous Corticosteroid Injection for Proximal Hamstring Tendinopathy: A Retrospective Review. Orthop J Sports Med. 2014;2(3):1-7.
136. Challoumas D, Clifford C, Kirwan P, Millar NL. How does surgery compare to sham surgery or physiotherapy as a treatment for tendinopathy? A systematic review of randomised trials. BMJ Open Sport & Exercise Medicine. 2019;5(1):e000528.
137. Orchard JW, Saw R, Masci LJCRR. The Use of Ultrasound-Guided Injections for Tendinopathies. 2018;6(10):38.
138. Coombes BK, Bisset L, Vicenzino B. Efficacy and safety of corticosteroid injections and other injections for management of tendinopathy: a systematic review of randomised controlled trials. Lancet. 2010;376(9754):1751-67.
139. Cook T, Minns Lowe C, Maybury M, Lewis JS. Are corticosteroid injections more beneficial than anaesthetic injections alone in the management of rotator cuff-related shoulder pain? A systematic review. Br J Sports Med. 2018;52(8):497-504.
140. Bisset L, Beller E, Jull G, Brooks P, Darnell R, Vicenzino B. Mobilisation with movement and exercise, corticosteroid injection, or wait and see for tennis elbow: randomised trial. BMJ (Clinical research ed). 2006;333(7575):1751-67.

141. Smidt N, van der Windt DA, Assendelft WJ, Deville WL, Korthals-de Bos IB, Bouter LM. Corticosteroid injections, physiotherapy, or a wait-and-see policy for lateral epicondylitis: a randomised controlled trial. Lancet. 2002;359(9307):657-62.
142. Dean BJ, Lostis E, Oakley T, Rombach I, Morrey ME, Carr AJ. The risks and benefits of glucocorticoid treatment for tendinopathy: a systematic review of the effects of local glucocorticoid on tendon. Seminars in arthritis and rheumatism. 2014;43(4):570-76.
143. Zhang JY, Fabricant PD, Ishmael CR, Wang JC, Petrigliano FA, Jones KJ. Utilization of Platelet-Rich Plasma for Musculoskeletal Injuries: An Analysis of Current Treatment Trends in the United States. Orthop J Sports Med. 2016;4(12):1-7.
144. Miller LE, Parrish WR, Roides B, Bhattacharyya S. Efficacy of platelet-rich plasma injections for symptomatic tendinopathy: systematic review and meta-analysis of randomised injection-controlled trials. BMJ Open Sport & Exercise Medicine. 2017;3(1):e000237.
145. Furia JP, Rompe JD, Cacchio A, Del Buono A, Maffulli N. A single application of low-energy radial extracorporeal shock wave therapy is effective for the management of chronic patellar tendinopathy. Knee Surg Sports Traumatol Arthrosc. 2013;21(2):346-50.
146. Rompe JD, Segal NA, Cacchio A, Furia JP, Morral A, Maffulli N. Home training, local corticosteroid injection, or radial shock wave therapy for greater trochanter pain syndrome. Am J Sports Med. 2009;37(10):1981-90.
147. Rompe JD, Furia J, Maffulli N. Eccentric loading compared with shock wave treatment for chronic insertional achilles tendinopathy. A randomized, controlled trial. J Bone Joint Surg Am. 2008;90(1):52-61.
148. Rompe JD, Furia J, Maffulli N. Eccentric loading versus eccentric loading plus shock-wave treatment for midportion achilles tendinopathy: a randomized controlled trial. Am J Sports Med. 2009;37(3):463-70.
149. Furia JP, Rompe JD, Maffulli N. Low-energy extracorporeal shock wave therapy as a treatment for greater trochanteric pain syndrome. Am J Sports Med. 2009;37(9):1806-13.
150. Rompe JD, Cacchio A, Furia JP, Maffulli N. Low-energy extracorporeal shock wave therapy as a treatment for medial tibial stress syndrome. Am J Sports Med. 2010;38(1):125-32.
151. Li W, Pan Y, Yang Q, Guo Z-G, Yue Q, Meng Q-G. Extracorporeal shockwave therapy for the treatment of knee osteoarthritis: A retrospective study. Medicine. 2018;97(27):e11418.
152. Bisset L, Beller E, Jull G, Brooks P, Darnell R, Vicenzino B. Mobilisation with movement and exercise, corticosteroid injection, or wait and see for tennis elbow: randomised trial. BMJ (Clinical research ed). 2006;333(7575):939.
153. Mellor R, Bennell K, Grimaldi A, Nicolson P, Kasza J, Hodges P, et al. Education plus exercise versus corticosteroid injection use versus a wait and see approach on global outcome and pain from gluteal tendinopathy: prospective, single blinded, randomised clinical trial. BMJ (Clinical research ed). 2018;361:k1662-k.
154. Herbert R. Practical evidence-based physiotherapy. Edinburgh: Elsevier, Churchill Livingstone; 2015.
155. Nasser AM, Vicenzino B, Grimaldi A, Anderson J, Semciw AI. Proximal hamstring tendinopathy; a systematic review of interventions. Int J Sports Phys Ther. 2021;16(2):288-305.
156. McCambridge AB, Nasser AM, Mehta P, Stubbs PW, Verhagen AP. Has Reporting on Physical Therapy Interventions Improved in 2 Decades? An Analysis of 140 Trials Reporting on 225 Interventions. J Orthop Sports Phys Ther. 2021;51(10):503-09.
157. Draper C. Role of qualitative research in exercise science and sports medicine. South African Journal of Sports Medicine. 2009;21(1).

158. Bolling C, Delfino Barboza S, van Mechelen W, Pasman HR. Letting the cat out of the bag: athletes, coaches and physiotherapists share their perspectives on injury prevention in elite sports. British Journal of Sports Medicine. 2019;54(14):871-77.
159. Scott A, Squier K, Alfredson H, Bahr R, Cook JL, Coombes B, et al. ICON 2019: International Scientific Tendinopathy Symposium Consensus: Clinical Terminology. British Journal of Sports Medicine. 2019;54(5):260-62.
160. Silbernagel KG, Thomeé R, Eriksson BI, Karlsson J. Continued sports activity, using a pain-monitoring model, during rehabilitation in patients with Achilles tendinopathy: a randomized controlled study. Am J Sports Med. 2007;35(6):897-906.
161. Nasser AM, Pizzari T, Grimaldi A, Vicenzino B, Rio E, Semciw AI. Proximal hamstring tendinopathy; expert physiotherapists' perspectives on diagnosis, management and prevention. Phys Ther Sport. 2021;48:67-75.
162. de Vos R-J, van der Vlist AC, Winters M, van der Giesen F, Weir A. Diagnosing Achilles tendinopathy is like delicious spaghetti carbonara: it is all about key ingredients, but not all chefs use the same recipe. British Journal of Sports Medicine. 2021;55(5):247.
163. Zwerver J, Mc Auliffe S, Rio EK, Scott A, Vicenzino BT, Weir A. ICON 2019—international scientific tendinopathy symposium: building an ICONic tendon tower—launching a new era in clinical tendinopathy research. British Journal of Sports Medicine. 2020;54(8):442-43.
164. Vicenzino B, de Vos R-J, Alfredson H, Bahr R, Cook JL, Coombes BK, et al. ICON 2019—International Scientific Tendinopathy Symposium Consensus: There are nine core health-related domains for tendinopathy (core domains): Delphi study of healthcare professionals and patients. British Journal of Sports Medicine. 2020;54(8):444-51.
165. Consensus-based Standards for the selection of health Measurement Instruments [Available from: http://www.cosmin.nl].
166. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, et al. The COMET Handbook: version 1.0. Trials. 2017;18(Suppl 3):280.
167. Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino M-A, et al. Developing Core Outcome Measurement Sets for Clinical Trials: OMERACT Filter 2.0. Journal of clinical epidemiology. 2014;67(7):745-53.
168. Gargon E, Gurung B, Medley N, Altman DG, Blazeby JM, Clarke M, et al. Choosing Important Health Outcomes for Comparative Effectiveness Research: A Systematic Review. PLOS ONE. 2014;9(6):e99111.
169. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. J Adv Nurs. 2000;32(4):1008-15.
170. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res. 2004;6(3):e34.
171. Mokkink LB PC, Patrick DL, et al. COSMIN Study Design checklist for Patient-reported outcome measurement instruments. 2019 [Available from: https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf].
172. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018;27(5):1147-57.
173. Hughes KL, Clarke M, Williamson PR. A systematic review finds Core Outcome Set uptake varies widely across different areas of health. Journal of clinical epidemiology. 2021;129:114-23.
174. Prinsen CA, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, et al. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" - a practical guideline. Trials. 2016;17(1):449.

175. Boric K, Boric M, Dosenovic S, Jelicic Kadic A, Batinic M, Cavar M, et al. Authors' lack of awareness and use of core outcome set on postoperative pain in children is hindering comparative effectiveness research. Journal of Comparative Effectiveness Research. 2018;7(5):463-70.
176. Nasser A, Semciw A, Grimaldi A, Rio E, Pizzari T, Vicenzino B. Core outcome set development for proximal hamstring tendinopathy (COS-PHT): a study protocol. Physical Therapy Reviews. 2022:1-7.
177. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ (Clinical research ed). 2021;372:n71.
178. Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. Qual Life Res. 2018;27(5):1171-79.
179. Terwee CB, Prinsen CAC, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Qual Life Res. 2018;27(5):1159-70.
180. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. Journal of clinical epidemiology. 2007;60(1):34-42.
181. Lee CP, Chiu YW, Chu CL, Chen Y, Jiang KH, Chen JL, et al. A reliability generalization meta-analysis of coefficient alpha and test-retest coefficient for the aging males' symptoms (AMS) scale. Aging Male. 2016;19(4):244-53.
182. Longo UG, Stelitano G, Berton A, Candela V, Barneschi G, Marescalchi M, et al. 2019 Marathon of Rome. Hamstring injuries in long distance runners: influence of age, gender, weight, height, numbers of marathon and impact profile. J Sports Med Phys Fitness. 2021;61(12).
183. Rich A, Cook JL, Hahne AJ, Rio EK, Ford J. Randomised, cross-over trial on the effect of isotonic and isometric exercise on pain and strength in proximal hamstring tendinopathy: trial protocol. BMJ Open Sport Exerc Med. 2021;7(1):e000954.
184. Kirschner JS, Cheng J, Hurwitz N, Santiago K, Lin E, Beatty N, et al. Ultrasound-guided percutaneous needle tenotomy (PNT) alone versus PNT plus platelet-rich plasma injection for the treatment of chronic tendinosis: A randomized controlled trial. Pm r. 2021;13(12):1340-49.
185. Fletcher AN, Pereira GF, Lau BC, Mather RC, 3rd. Endoscopic Proximal Hamstring Repair Is Safe and Efficacious with High Patient Satisfaction at a Minimum of Two-Year Follow-Up. Arthroscopy. 2021;37(11):3275-85.
186. Maldonado DR, Annin S, Lall AC, Krych AJ, Athey AG, Spinner RJ, et al. Outcomes of Open and Endoscopic Repairs of Chronic Partial- and Full-Thickness Proximal Hamstring Tendon Tears: A Multicenter Study With Minimum 2-Year Follow-up. Am J Sports Med. 2021;49(3):721-28.
187. Locquet M, Bornheim S, Colas L, Beveren JV, BruyËre O, Reginster J, et al. Traduction en langue franÁaise, adaptation transculturelle et validation du questionnaire ´ Victorian Institute of Sport Assessment Scale for Proximal Hamstring Tendinopathy ª (VISA-H). Journal De Traumatologie Du Sport. 2019;36:217-21.
188. De-la-Cruz-Torres B, Maffulli N, Hernández-Sánchez S. Cross-cultural adaptation of the Victorian institute of sports assessment - Hamstrings (VISA-H) questionnaire for Spanish speaking athletes with proximal hamstring tendinopathy. Physical Therapy in Sport. 2021;51:50-7.

189. Locquet M, Willems T, Specque C, Beaudart C, Bruyère O, Van Beveren J, et al. Cross-cultural adaptation, translation, and validation of the functional assessment scale for acute hamstring injuries (FASH) questionnaire for French-speaking patients. Disability & Rehabilitation. 2020;42(14):2076-82.
190. Lohrer H, Nauck T, Korakakis V, Malliaropoulos N. Validation of the FASH (Functional Assessment Scale for Acute Hamstring Injuries) questionnaire for German-speaking football players. J Orthop Surg Res. 2016;11(1):130.
191. Malliaropoulos N, Korakakis V, Christodoulou D, Padhiar N, Pyne D, Giakas G, et al. Development and validation of a questionnaire (FASH--Functional Assessment Scale for Acute Hamstring Injuries): to measure the severity and impact of symptoms on function and sports ability in patients with acute hamstring injuries. Br J Sports Med. 2014;48(22):1607-12.
192. Hernández-Sanchez S, Korakakis V, Malliaropoulos N, Moreno-Perez V. Validation study of the Functional Assessment Scale for Acute Hamstring injuries in Spanish professional soccer players. Clin Rehabil. 2019;33(4):711-23.
193. Mallows A, Littlewood C, Malliaras P. Measuring patient-reported outcomes (PROs/PROMs) in people with Achilles tendinopathy: how useful is the VISA-A? British Journal of Sports Medicine. 2018;52(19):1221.
194. Korakakis V, Kotsifaki A, Stefanakis M, Sotiralis Y, Whiteley R, Thorborg K. Evaluating lower limb tendinopathy with Victorian Institute of Sport Assessment (VISA) questionnaires: a systematic review shows very-low-quality evidence for their content and structural validity-part I. Knee Surg Sports Traumatol Arthrosc. 2021;29(9):2749–64.
195. Kamper SJ. Interpreting Outcomes 3-Clinical Meaningfulness: Linking Evidence to Practice. J Orthop Sports Phys Ther. 2019;49(9):677-78.
196. Martin RL, Kelly BT, Philippon MJ. Evidence of validity for the hip outcome score. Arthroscopy. 2006;22(12):1304-11.
197. Harris WH. Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. An end-result study using a new method of result evaluation. J Bone Joint Surg Am. 1969;51(4):737-55.
198. Korakakis V, Whiteley R, Kotsifaki A, Stefanakis M, Sotiralis Y, Thorborg K. A systematic review evaluating the clinimetric properties of the Victorian Institute of Sport Assessment (VISA) questionnaires for lower limb tendinopathy shows moderate to high-quality evidence for sufficient reliability, validity and responsiveness-part II. Knee Surg Sports Traumatol Arthrosc. 2021;29(9):2765-88.
199. Korakakis V, Whiteley R, Kotsifaki A, Thorborg K. Tendinopathy VISAs have expired—is it time for outcome renewals? Knee Surgery, Sports Traumatology, Arthroscopy. 2021;29(9):2745–48.
200. Fitzpatrick J, Bulsara MK, O'Donnell J, McCrory PR, Zheng MH. The Effectiveness of Platelet-Rich Plasma Injections in Gluteal Tendinopathy: A Randomized, Double-Blind Controlled Trial Comparing a Single Platelet-Rich Plasma Injection With a Single Corticosteroid Injection. Am J Sports Med. 2018;46(4):933-39.
201. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. Open Med. 2009;3(3):e123-30.
202. Chamorro C, Armijo-Olivo S, De la Fuente C, Fuentes J, Javier Chirosa L. Absolute Reliability and Concurrent Validity of Hand Held Dynamometry and Isokinetic Dynamometry in the Hip, Knee and Ankle Joint: Systematic Review and Meta-analysis. Open Med. 2017;12(17):359-75.
203. Sattler MC, Jaunig J, Watson ED, van Poppel MNM, Mokkink LB, Terwee CB, et al. Physical Activity Questionnaires for Pregnancy: A Systematic Review of Measurement Properties. Sports Medicine. 2018;48(10):2317-46.

204. Jorgensen JE, Fearon AM, Mølgaard CM, Kristinsson J, Andreasen J. Translation, validation and test-retest reliability of the VISA-G patient-reported outcome tool into Danish (VISA-G.DK). PeerJ. 2020;8:e8724.
205. Ebert JR, Fearon AM, Smith A, Janes GC. Responsiveness of the Victorian Institute for Sport Assessment for Gluteal Tendinopathy (VISA-G), modified Harris hip and Oxford hip scores in patients undergoing hip abductor tendon repair. Musculoskelet Sci Pract. 2019;43(1):1-5.
206. Minetto MA, Busso C, Giannini A, Meiburger K, Massazza G, Maffulli N. Cross-cultural adaptation and validation of the Victorian Institute of Sports Assessment for Gluteal Tendinopathy questionnaire in Italian and investigation of the association between tendinopathy-related disability and pain. Eur J Phys Rehabil Med. 2020;56(6):764-70.
207. Beaudart C, Gillier M, Bornheim S, Beveren J, Bruyère O, Kaux J-F. French Translation and Validation of the Victorian Institute of Sports Assessment for Gluteal Tendinopathy Questionnaire. PM&R. 2020:137-43.
208. Paiva EB, Azevedo DC, Pereira AL, Garcia AN, Percope de Andrade MA. Translation, cross-cultural adaptation and validation of the Brazilian Portuguese version of the Victorian Institute of Sports Assessment for Gluteal Tendinopathy patient reported-outcome measure (VISA-G.BR). Musculoskelet Sci Pract. 2021;52:1-6.
209. Costa RMdP, Cardinot TM, Mathias LNCDC, Leporace G, Oliveira LPd. Validation of the Brazilian version of the Hip Outcome Score (HOS) questionnaire. Advances in Rheumatology. 2018;58(1):4-11.
210. Visentini PJ, Khan KM, Cook JL, Kiss ZS, Harcourt PR, Wark JD. The VISA score: an index of severity of symptoms in patients with jumper's knee (patellar tendinosis). Victorian Institute of Sport Tendon Study Group. Journal of science and medicine in sport. 1998;1(1):22-8.
211. Robinson JM, Cook JL, Purdam C, Visentini PJ, Ross J, Maffulli N, et al. The VISA-A questionnaire: a valid and reliable index of the clinical severity of Achilles tendinopathy. British Journal of Sports Medicine. 2001;35(5):335-41.
212. Segal NA, Felson DT, Torner JC, Zhu Y, Curtis JR, Niu J, et al. Greater trochanteric pain syndrome: epidemiology and associated factors. Arch Phys Med Rehabil. 2007;88(8):988-92.
213. Comins J, Siersma V, Couppe C, Svensson RB, Johansen F, Malmgaard-Clausen NM, et al. Assessment of content validity and psychometric properties of VISA-A for Achilles tendinopathy. PLOS ONE. 2021;16(3):e0247152.
214. Jeffries AC, Wallace L, Coutts AJ, McLaren SJ, McCall A, Impellizzeri FM. Athlete-Reported Outcome Measures for Monitoring Training Responses: A Systematic Review of Risk of Bias and Measurement Property Quality According to the COSMIN Guidelines. Int J Sports Physiol Perform. 2020:1-13.
215. de Boer AGEM, van Lanschot JJB, Stalmeier PFM, van Sandick JW, Hulscher JBF, de Haes JCJM, et al. Is a single-item visual analogue scale as valid, reliable and responsive as multi-item scales in measuring quality of life? Quality of Life Research. 2004;13(2):311-20.
216. Hush JM, Alison JA. Evidence-based practice: lost in translation? J Physiother. 2011;57(3):143-44.
217. Hussain N, Johal H, Bhandari M. An evidence-based evaluation on the use of platelet rich plasma in orthopedics - a review of the literature. SICOT J. 2017;3:57.
218. Crevenna R, Mickel M, Schuhfried O, Gesslbauer C, Zdravkovic A, Keilani M. Focused Extracorporeal Shockwave Therapy in Physical Medicine and Rehabilitation. Current Physical Medicine and Rehabilitation Reports. 2021;9(1):1-10.

219. Kirby D, Fried JW, Bloom DA, Buchalter D, Youm T. Clinical Outcomes After Endoscopic Repair of Gluteus Medius Tendon Tear Using a Knotless Technique With a 2-Year Minimum Follow-Up. Arthroscopy. 2020;36(11):2849-55.
220. Blakeney WG. Editorial Commentary: Proximal Hamstring Tendon Injuries: Is the Research Hamstrung by the Use of Poor Outcome Measures? Arthroscopy. 2020;36(5):1308-10.
221. Hunger I, Sparkes A, Stelter R. Qualitative Methods in Sport Sciences: A Special FQS Issue. Social Science Open Access Repository. 2003;29(3):192-204.
222. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. J Man Manip Ther. 2009;17(3):163-70.
223. Irby A, Gutierrez J, Chamberlin C, Thomas SJ, Rosen AB. Clinical management of tendinopathy: A systematic review of systematic reviews evaluating the effectiveness of tendinopathy treatments. Scand J Med Sci Sports. 2020;30(10):1810-26.