Test-Retest Reliability and Validity of Wearable Transdermal Alcohol Monitors

Kelly van Egmond

BSc MSc (Social sciences)

Submitted in total fulfilment of the requirements for the degree of
Doctor of Philosophy

Centre for Alcohol Policy Research

School of Psychology and Public Health

La Trobe University

Victoria, Australia

March 2022

Table of Contents

Abstract

Accurate measurement of alcohol consumption is key to monitoring population drinking patterns and consequences and enabling understanding of the links between alcohol consumption and harms. Researchers have traditionally used retrospective surveys to measure alcohol consumption, but self-reported surveys are prone to bias, such as memory or recall deficits, social desirability and motivation to report. Transdermal alcohol concentration (TAC) monitors offer continuous, objective, and non-invasive data collection, which may overcome these limitations. However, research on the reliability and accuracy of TAC devices is scant. This research was designed to determine the reliability and validity of TAC devices' measurement of alcohol consumption.

Twenty-two participants were invited to the laboratory from about 9.00am-6.00pm for three sessions in total. Each session they were asked to consume four standard drinks (each 10 grams of alcohol). During these sessions participants wore the ION RAP wristband (a new-generation device) and a SCRAM-CAM on both the left and right wrist and ankle respectively. Breath samples were taken in 10-minute intervals. TAC data were then studied to establish the devices' test–retest reliability (consistency of two simultaneously worn devices; reliability and agreement) and validity (correspondence to an external measure, such as a breathalyser).

Results showed lower-than-expected reliability for the SCRAM-CAM, but excellent reliability for the ION RAP; however, the latter experienced a much higher failure rate than the former. Further, both monitors showed substantial disagreement between the TAC values measured on the left and right side of the body. In terms of validation, the monitors showed small correlations to breath alcohol concentration and a lower sensitivity to low levels of alcohol consumption.

This thesis contributes a comprehensive study of transdermal alcohol measurement. The results of the research suggest that wearable TAC monitors are not sufficiently reliable and valid to measure quantities of alcohol consumed in real time.

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.

Kelly van Egmond

Date: 25/03/2022

Declaration for thesis based or partially based on conjointly published or unpublished work

This thesis includes one systematic review and one original paper published in peer-reviewed journals, and three original papers submitted to peer-reviewed journals. The ideas, development and writing up of all the papers in this thesis were the principal responsibility of myself, the candidate, working within the School of Psychology and Public Health under the supervision of Prof. Emmanuel Kuntsche, Dr Cassandra Wright, Dr Benjamin Riordan, and Dr Michael Livingston. The inclusion of the co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.

The undersigned hereby certify that:

1. The author contributions statements listed within this thesis correctly reflect the nature and extent of the candidate's contribution, and the nature of the contribution of each of the co-authors.

2. The co-authors listed below meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise.

3. They take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication.

4. There are no other authors of the publication according to these criteria; and

5. Potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit.

Name	Signature	Date
Kelly van Egmond		
Emmanuel Kuntsche		
Cassandra Wright		
Benjamin Riordan		
Michael Livingston		
Amy Pennay		
Dan Anderson-Luxford		
Gabriel Caluzzi		

Research outcomes during enrolment

Peer-reviewed publications

Articles included in the thesis

van Egmond, K., Wright, C., Livingston, M., & Kuntsche, E. (2020). Wearable transdermal alcohol monitors: a systematic review of detection validity, TAC-to-BrAC relationship and factors influencing both detection and the TAC-to-BrAC relationship. *Alcoholism: Clinical and Experimental Research, 44*(10), 1918-1932.

van Egmond, K., Kuntsche, E., JC Wright, C., Livingston, M. (2021). A parallel test of the SCRAM-CAM transdermal monitors ensuring reliability. *Drug and Alcohol Review, 40*(7), 1122-1130.

Articles outside the thesis

Callinan, S., Leggat, G., **Van Egmond, K.,** & Lindell, A. (2022). The impact of handedness on health risk behaviours and socio-economic outcomes. *Personality and Individual Differences,* 187, 111411.

Wright, C. J., Miller, M., **van Egmond, K.,** Caluzzi, G., Anderson-Luxford, D., Pennay, A., & Kuntsche, E. (2021). Small, slim, sleek, and familiar: user experiences with an ION wearable research alpha prototype transdermal alcohol monitor. *Addiction Research & Theory*, 1-6.

Manuscripts included in the thesis

van Egmond, K., Riordan, B., JC Wright, C., Livingston, M., Kuntsche, E. (submitted for publication). Measurement of transdermal alcohol concentration using a wrist worn enzymatic transdermal monitor.

van Egmond, K., Riordan, B., JC Wright, C., Livingston, M., Kuntsche, E. (submitted for publication). To what degree can we rely transdermal monitors to measure alcohol consumption? A study on agreement and reliability.

van Egmond, K., Anderson-Luxford, D., Kuntsche, E., JC Wright, C., Caluzzi, G., Pennay, A. (submitted for publication). Measuring alcohol consumption while watching sport events: a feasibility study comparing ecological momentary assessments and transdermal alcohol monitors.

Manuscripts outside the thesis

Byrne, J. E. M., Choong S. Y., **Van Egmond, K**., Fulller-Tyszkiewicz, M. & Staiger, P.K. (submitted for publication). Daily associations of the relationships between health behaviours and sleep-wake processes in adults.

Nielsen, L.H., Hayden, M.J., Lawrence, N.S., Kakoschke, N., **Van Egmond, K.,** Lum. J., Staiger, P.K., (submitted for publication). Inhibitory control training to reduce appetitive behaviour: a meta-analytic investigation of effectiveness and potential moderators and mediators.

Oral presentations

van Egmond, K., Wright, C., Livingston, M., Kuntsche, E. (2021, June). Wearable transdermal alcohol monitors: a systematic review of validation studies and a parallel test of the SCRAM-CAM ankle monitors. Oral presentation at the virtual 44th Annual RSA Scientific Meeting / ISBRA Congress, America.

van Egmond, K., Riordan, B., Wright, C., Livingston, M., Kuntsche, E. (2021, June). Testing the validity of a new-generation transdermal alcohol concentration wristband in controlled environments. Oral presentation at the virtual 46th Annual Alcohol Epidemiology Symposium of the Kettil Bruun Society (KBS 2021) in Helsinki, Finland.

van Egmond, K. (2020, July + August). 'Sweat the small stuff' – Can we measure alcohol with a wristband? Oral presentation at 3-Minute Thesis Competition rounds at La Trobe University: 1. School round: Psychology and Health, 2. College round: Science, Health and Engineering. Winner of school round.

van Egmond, K., Kuntsche, E., Wright, C., Livingston, M. (2020, July). Wearable transdermal alcohol monitors: a systematic review of validation studies and a pilot study testing Correspondence

between transdermal measurements and self-reported alcohol consumption. Abstract accepted for RSA/ISBRA-2020 conference, New Orleans, America. Delayed to 2021 due to COVID-19.

van Egmond, K., Kuntsche, E., Wright, C., Livingston, M. (2020, January 17). Correspondence between transdermal measurements and self-reported alcohol consumption during student orientation over three consecutive days. Oral presentation at the 7th Conference of the Society for Ambulatory Assessment, Melbourne, Australia.

Poster presentations

van Egmond, K., Wright, C., Livingston, M., Kuntsche, E. (2021, June). A parallel test of the SCRAM-CAM ankle monitors ensuring accuracy and reliability. Online poster presentation at the Conference of the Society for Ambulatory Assessment, Zurich, Switzerland.

Acknowledgements

First of all, I am extremely grateful to my supervisors; Emmanuel Kuntsche, Cassandra Wright, Benjamin Riordan and Michael Livingston. You have been such an amazing source of support to me, and without you all I would not have made it (literally). Emmanuel, you always had your door open for me for any advice and was ready to help at any moment of the day and week. Your enthusiasm and hard work have been inspiring and made all challenges we faced so much better. You supported me through all the hard times, and I could not have wished for a better primary supervisor to tackle transdermal alcohol research with. Cassandra, you have a great mentor and friend to me. I am in awe of your achievements and commitment and have always felt super comfortable to ask you anything, thanks so much for all your advice, whether it was life, career, or thesis advice. Benjamin, even though you joined or little transdermal alcohol technology team later, I cannot thank you enough for all the R advice and help you have given. I cannot say learning to deal with TAC data on R was always a smooth journey, but I was glad not to be alone to brainstorm and just stare at the data for hours at a time. Michael, your stats insights were invaluable, and even though you moved universities, you still provided continuous support, feedback, and encouragement throughout. Also, your inputs and thoughts balanced out the conclusions around transdermal monitors, I think we both agreed that they did not work as good as we had hoped.

Even though Amy Pennay, you were not part of my supervision team, as my manager you have been an enormous help to me and this thesis, which is why I wanted to thank you for all your advice, help and support as well. I really enjoyed working together.

To my fellow PhD students and friends, you have made this journey actual fun and I am grateful for all the experiences we have shared. Special thanks to Gabriel Caluzzi, Megan Cook, Hannah Jarman, Thomas Norman, Geoffrey Leggat, Rakhi Vashishtha, Dan Anderson-Luxford and Yvette Mojica Perez. You were not only a fellow PhD student and a friend, but helped me in my research by providing help, advice and even being part of the research yourself.

I would also like to say thanks to all my colleagues at CAPR, who were invaluable in my PhD journey. The team of CAPR had grown so much since I started and the culture makes it a pleasure to work, which is reflected in the quality of work that is produced by the CAPR team. I will miss this so much and wish everyone the best in their research career ahead.

My partner and best friend, Wilson Chen, I cannot thank enough. You have supported me through all the ups and downs and literally did the PhD with me, thank you for being so patient and understanding. COVID and working from home was not easy, but I couldn't have done it without your support. Of course, this thesis would not be complete without giving a thanks to Chunky and Whisky, our guineapigs, that provided us with entertainment and warmth when needed.

Thank you to my family and friends, especially my parents Jos and Peter van Egmond, who had such a hard time with their daughter moving to the other side of the world but did nothing but support and encouraging me throughout it all.

This work was supported by a La Trobe University Postgraduate Research Scholarship and La Trobe Full Fee Research Scholarship. I am very grateful for this financial assistance, without which I would not have been able to undertake my PhD.

I would like to thank all the participants for their time and participation. Without their cooperation in entering the laboratory for three full days, having an alcoholic drink at 10am and letting me strap four monitors on them each time, this work would have been impossible.

Lastly, Dr Campbell Aitken provided professional editing services in accordance with the Institute of Professional Editors' *Guidelines for editing research theses*.

Abbreviations

AFL	Australian Rules Football
AMS	Alcohol Monitoring Systems, Inc.
AUC	Area Under the Curve
AUDIT-C	Alcohol Use Disorders Identification Test – Concise
BAC	Blood Alcohol Concentration
BMI	Body Mass Index
BrAC	Breath Alcohol Concentration
eBAC	estimated BAC
EMA	Ecological Momentary Assessment
ICC	Intra-Class Correlation
K10	Kessler Psychological Distress Scale
SCRAM-CAM	Secure Continuous Remote Alcohol Monitors – Continuous Alcohol Monitoring
TAC	Transdermal Alcohol Concentration
TLFB	Timeline Follow-Back

1. Introduction

Alcohol consumption in Australia

Alcohol is part of everyday life for most Australians (Laslett et al. 2010), with at least 77% of the population aged over 14 years having consumed alcohol in their lifetime (Australian Institute of Health and Welfare 2020). It is associated with many social and cultural activities (Lloyd et al. 2013) and is often present at events such as birthdays, funerals and graduations (Allan et al. 2012). In Australia, per capita consumption has been declining since the 1980s, with an increasing proportion of the population drinking less or abstaining. However, in 2018, 25% of the population aged over 14 years reported drinking at a risky level (more than four Australian standard drinks on a single occasion) at least monthly (Australian Institute of Health and Welfare 2020). Risky alcohol consumption is associated with hospitalisation, lower life expectancy, reduced productivity at work, violence, road accidents, homicide, and suicide (Australian Institute of Health and Welfare 2018; Hurzeler et al. 2021). It has further been identified as the fifth-highest contributor to disease burden in Australia, constituting 4.5% of the total burden and contributing to 4.1% of deaths (Australian Institute of Health and Welfare 2018). Alcohol use is the leading cause of disease (14%) and non-fatal (12%) burden in male adolescents and adults (aged 15–44) (Australian Institute of Health and Welfare 2018). The estimated cost of alcohol-related problems to Australian society in 2019 was $66.8 billion, including $18.2 billion in tangible costs ($4.0 billion from absenteeism/injury, $3.1 billion for crime, $2.8 billion for health care, $2.4 billion for road traffic accidents) and $48.6 billion in intangible costs ($25.9 billion for lost years of life, $20.7 billion from lost quality of life) (National Drug Research Institute 2021).

The considerable impacts of alcohol consumption in Australia make it important to gather accurate information about the phenomenon. Measurement of alcohol consumption is a key part in monitoring population drinking patterns and consequences, enabling understanding of the links between alcohol consumption and harms, mental health, or brain development, and the design of

clinical interventions, policy, and national prevention plans. Researchers have traditionally used retrospective surveys to measure alcohol consumption, but these methods have several limitations that can reduce the accuracy of alcohol reporting (e.g., bias as a result of cognitive impairments, or social desirability) (Callinan 2015; Gmel and Rehm 2004; Greenfield and Kerr 2008; Livingston and Callinan 2015). Advancements in knowledge and technology have led to the development of new methods to measure alcohol consumption and overcome some of the limitations of questionnaire-based methods, including ecological momentary assessment (EMA) surveys (Kuntsche and Labhart 2013b; Voogt et al. 2014; Wright et al. 2018) and breathalysers (Gibb et al. 1984; Gmel and Rehm 2004; Riordan et al. 2017; Riuttanen, Jäntti, and Mattila 2020). However, these methodologies have their own strengths and weaknesses.

Various transdermal alcohol concentration (TAC) measuring devices have been developed over the past two decades, and hold considerable promise as a research tool (Leffingwell et al. 2013). The recent development of small TAC monitoring wristbands that can sync to mobile phones is particularly exciting (Fairbairn and Kang 2019; Lansdorp et al. 2019). Indeed, these devices promise to overcome most of the limitations of self-reports and breathalysers by offering continuous, objective, and non-invasive alcohol monitoring. A reliable wearable alcohol biosensor would be a crucial tool to help researchers gather detailed information on drinking behaviours over longer periods of time.

Test–retest reliability and validity are important concepts for assessing a measuring tool's performance (Cohen and Vinson 1995; Csikszentmihalyi et al. 2014; Heale and Twycross 2015; Riordan et al. 2017). Test–retest reliability has two separate components: reliability and agreement (Berchtold 2016). Reliability is the capacity of the monitor to reproduce a similar measurement pattern across participants when applied under similar conditions (Berchtold 2016; Heale and Twycross 2015; Sobell et al. 1986). High agreement requires a monitor to provide identical values when applied to the same participant under similar conditions over time (Berchtold 2016; Bland and Altman 1986). Thus, devices may demonstrate good reliability, producing similar-shaped alcohol concentration curves across the same drinking event, but low agreement because curves peak at

different levels. A consistent measurement tool has both high agreement and high reliability. Validity, specifically construct or convergent validity, is the extent to which wearable monitors accurately measure the amount of alcohol (Heale and Twycross 2015). Establishing validity requires studying the correspondence between the monitored result and a reference measure of alcohol consumption (e.g., breath alcohol concentration [BrAC] or self-reports) (Csikszentmihalyi et al. 2014; Heale and Twycross 2015; Sakai et al. 2006). Research assessing the test-retest reliability and validity of TAC devices is scarce; the work presented in this thesis represents a critical first step to determine whether these devices are fit for use in research and other settings where detailed measurement of the amount of alcohol consumed in humans is necessary.

In this chapter, I summarise the limitations and strengths of traditional self-report questionnaires, EMAs, and breathalysers for measuring alcohol consumption in real-time, and explain how wearable transdermal devices could overcome these limitations and why they demand research attention.

Alcohol measurement

Traditional self-reported measures

The most popular way to measure alcohol consumption is through self-reported measures, mostly using retrospective surveys (Gmel and Rehm 2004; Lajunen and Summala 2003; May and Foxcroft 1995). In surveys, people are asked about their alcohol consumption and their drinking patterns over a certain period (e.g., the past year, month, or week). This can include general questions on their consumption in the past year (e.g., "how often did you drink more than 6 standard drinks in an occasion?") or more specific questions in a timeline follow-back (TLFB) method, in which (for example) participants might be asked how many drinks they had over how many hours over the past seven days.

Using self-reports to measure alcohol consumption has multiple advantages. Firstly, it is cheap, and can be used to measure the drinking patterns of individuals or populations. Survey data

provides more information than other quantity measures (like alcohol purchasing data or wastewater alcohol concentration data), because participants can be asked for sociodemographic information, about their health status, and any alcohol-related health, social or economic consequences. The ability to collect a range of data in surveys enables researchers to link alcohol consumption to various outcomes on the level of the individual, events and environment, as well as control for individual characteristics. In addition, consecutive surveys allow for tracking of patterns of risky or binge drinking over time (Gmel, Kuntsche, and Rehm 2011). Self-reported methods are known to be reliable; TLFB and online survey methods have shown high test-retest reliability (Khadjesari et al., 2009; Miller et al., 2002; Sobell et al., 1988; Sobell et al., 1986).

Nevertheless, because self-reported surveys rely on people providing data about themselves and alcohol consumption is a somewhat socially sanctioned behaviour and can hence be a sensitive topic to report on, they are prone to bias, such as memory or recall deficits, social desirability and motivation to report (Greenfield and Kerr 2008; Karns-Wright et al. 2018; Shiffman 2009). This often results in underreporting; population surveys generally produce estimates of population-level drinking that are 40–50% lower than objective measurements based on alcohol sales or tax data (Gmel and Rehm 2004; Livingston and Callinan 2015). One of the most important considerations here is the reference period for which consumption data are collected – the longer the reference period, the higher the probability of underreporting due to memory loss (Kuntsche and Labhart 2012). Conversely, shorter reference periods may not represent the individual's typical consumption through a year and the variations throughout (by seasons and during holidays, for example).

Finally, standard drink sizes and the amount of alcohol poured per drink can vary by country of origin (In Australia one standard drink is defined as containing 10 grams of alcohol, while in the United States one standard drink contains 14 grams of alcohol) (Gilligan et al. 2019; Gmel and Rehm 2004; Kerr et al. 2005). Many people are unfamiliar with the concept of standard drinks, which can introduce further reporting bias in surveys (Callinan 2015; Kerr et al. 2008; Stockwell et al. 2004).

Ecological Momentary Assessments

Ecological momentary assessments are repeated surveys with short recall periods, often completed in the participants' natural environments, and are often used to investigate drinking behaviours (Piasecki 2019; Shiffman 2009; Shiffman, Stone, and Hufford 2008; Wray, Merrill, and Monti 2014). EMA surveys are particularly suitable for small to medium-scale studies that are not intended to collect data representative of a whole population. Rather than requiring participants to recall average alcohol consumption from the previous week or month, EMAs involve frequent repeated measures of drinking behaviour during or shortly after the event (e.g., one EMA per hour during a drinking session asking people to report how many drinks they have consumed, how they feel, and what consequences they have experienced) (Kuntsche, Dietze, and Jenkinson 2014; Kuntsche and Labhart 2013b, 2014; Wright et al. 2018). This allows researchers to investigate drinking behaviour in a natural environment ("ecological"), measure current or recent circumstances ("momentary"), and explore changes in behaviour over time and situations (Shiffman et al. 2008; Wray et al. 2014). Most EMA protocols are based on well-established self-report measures, including intensive longitudinal methods such as diaries (Bolger & Laurenceau, 2013; Schafer & Walls, 2006), experience sampling (Csikszentmihalyi et al. 2014), and ambulatory assessment (Trull and Ebner-Priemer 2014).

EMA surveys offer a range of advantages over retrospective self-report measures. Due to taking higher-frequency measurements "in the moment" rather than asking participants to recall their drinking, EMA can avoid biases due to recall deficits and thus reduce the risk of underreporting (Dulin et al. 2017; Jones et al. 2018; Mun et al. 2021). For example, studies found that participants reported more drinks when using EMA surveys than when asked about their consumption retrospectively (Kuntsche and Labhart 2012; Monk et al. 2015). EMA surveys can be carried out on digital devices, such as smartphones, through automated smartphone applications, making them cost-efficient (Kuntsche and Labhart 2014; Shiffman 2009; Shiffman et al. 2008; Wright et al. 2018). Smartphone-based EMA also gives researchers the option of collecting other types of data, such as photos, location data, audio, video and can even include behavioural and cognitive tasks, for example

collecting information on response times (Labhart et al. 2019; Labhart, Muralidhar, et al. 2021; Labhart, Phan, et al. 2021; Voogt et al. 2014). Further, EMA offers an important advantage over cross-sectional surveys in enabling the study of drinking over time and the influence of changes in other momentary variables such as the context, mood and alcohol-related consequences (Kuntsche and Bruno 2015; Kuntsche and Labhart 2013b; Wray et al. 2014).

Despite their strengths, EMA survey methods have limitations associated with participant reactivity, compliance, and burden (Piasecki 2019; Wray et al. 2014). When collecting more information by increasing the survey interval frequency (e.g., one survey per half hour/hour) or lengthening the duration of the study (e.g., to one month), the burden for the participants increases. While traditional retrospective surveys collect data at a single timepoint, EMA can document in almost-real time the participants' drinking behaviours over an extended period of time (e.g. daily for 30 days) (Kuntsche and Bruno 2015; Kuntsche et al. 2014). This increase in burden can influence the participants' willingness to engage and lead to them missing surveys or dropping out of the study, decreasing the representativeness of the results (Piasecki 2019). Behavioural reactivity is change in someone's behaviour when being monitored, and is found to be dependent on the sample. For example, Collins et al. (1998) found that when studying the effectiveness of EMA in a heavy drinking population who were interested in cutting down their drinking, participants in the control condition showed a tendency to report reduced drinking over time, indicating a reactivity effect. However, Luczak et al. (2015) found this effect was less pronounced in a sample of college students, with 71% indicating that the monitoring did not affect their drinking behaviour. Finally, even though recall bias is reduced in EMA, the method still relies on self-report, and similar limitations relating to social desirability and difficulties in reporting the number of standard drinks consumed apply (Callinan 2015; Hufford et al. 2002).

Breathalysers

To increase the objectivity of alcohol measurement and avoid self-report biases, researchers have sought out biometric solutions, such as measuring alcohol concentration in blood serum, breath or urine. Breathalysers are especially suitable for research or situations in which the aim is to measure

the exact level of alcohol in the blood at a specific moment in time. Breathalysers collect a breath sample and use sensors to measure breath alcohol concentration (BrAC), an estimate of blood alcohol concentration (BAC). Once alcohol reaches the bloodstream, it is carried throughout your body, including the lungs. The alcohol in the lungs evaporates while breathing and gets expelled through the mouth. The alcohol in the breath can be measured by the breathalyser through a chemical reaction, which can then be converted to a BAC measurement (ratio BrAC to BAC is about 2100:1). A wide range of devices is available (personal/police-grade, fuel cell/infrared spectroscopy, handheld/wall mounted), with most of the breathalysers used in research being handheld devices that use a fuel cell sensor. If applied correctly, these devices offer accurate BAC estimates, and are therefore used by police in roadside testing to determine if a motorist is over the drink driving limit, in laboratory studies (with cognitive tasks conducted at different BrACs) (Norman et al. 2020; Riordan et al. 2017), clinics and emergency departments (to detect the presence of alcohol) (Kelly et al. 2002; Riuttanen et al. 2020), in-situ/street intercept studies (where a breath test is paired with a brief survey/interview during a night out) (Durbeej et al. 2017; Kingsland et al. 2013), and to supplement EMA (Lauckner et al. 2019).

Measurement of BAC through breathalysers excludes risk for cognitive bias caused by recall or memory errors, and is thus a more objective way of measuring alcohol consumption than either retrospective self-reports or EMA surveys. Many handheld devices are quite large and expensive, but smaller fuel cell devices have been developed for personal use; these are accurate, cheap and highly portable, and can be connected to a smartphone application informing on the breath readings through Bluetooth (Riordan et al. 2017). These small breathalysers allow research participants to collect breath samples themselves, freeing researchers from having to collect multiple breath samples from multiple participants in person.

Biomarkers such as BrAC and BAC are limited in their window of detection, and detection of the amount of alcohol throughout a drinking event requires multiple breath samples from the participant, which can be burdensome for both the researcher and participant. This makes the use of breathalysers impractical in large samples and over long periods of time. Further, personal

breathalysers have been shown to overreport BAC relative to police-graded devices, particularly at higher BACs (Riordan et al. 2017). Because they require repeated active participation during drinking events, they may also be inconvenient or embarrassing to use in public or social settings, which can again cause a decrease in compliance and motivation and might influence the participants' drinking behaviour. Finally, breathalysers can pick up residual alcohol in the mouth, which often results in higher BrAC readings. Manufacturers recommend a 10–20-minute interval between the last sip of alcohol and the breath test, and some use water to reduce the residual alcohol left in the mouth. To ensure an optimal sample and reduce the risk of bias, researchers either have to administer the test themselves or rely on participants to follow a clear protocol, which is especially challenging when participants are under the influence of alcohol.

The potential solution: Transdermal alcohol monitors

As outlined above, existing methods of alcohol data collection have substantial limitations. These limitations are particularly evident for researchers aiming to study patterns of drinking and intoxication in alcohol consumption events. Surveys and EMAs need to balance recall biases against participant burden, while objective measures from breathalysers are likely to be intrusive and/or inaccurate in real-world settings, especially those involving heavy drinking. Thus, the development of a continuous, non-invasive and objective measurement of alcohol consumption is critical to improve measurement and provide a better understanding of drinking behaviours.

Wearable TAC monitors could overcome the limitations of self-reports and breathalysers by measuring alcohol concentration objectively, passively and in detail. Approximately 1% of alcohol consumed by humans is secreted transdermally, both actively through sweat and passively through the skin (Swift 2003) (Figure 1-1). The idea of measuring TAC dates to the 1930s, when it was found that alcohol concentration could be measured in perspiration (Swift 2003). Since then, various devices have been developed to try to measure TAC. Like breathalysers that estimate BAC by measuring the amount of alcohol in expired air, TAC monitors estimate BAC in perspiration, thus eliminating cognitive biases (Table 1-1). However, unlike breathalysers, these monitors have the potential to provide continuous measurement of alcohol during a drinking event and over longer periods without

any action required from the user. Plotting TAC data over time allows researchers to study dynamic drinking data.

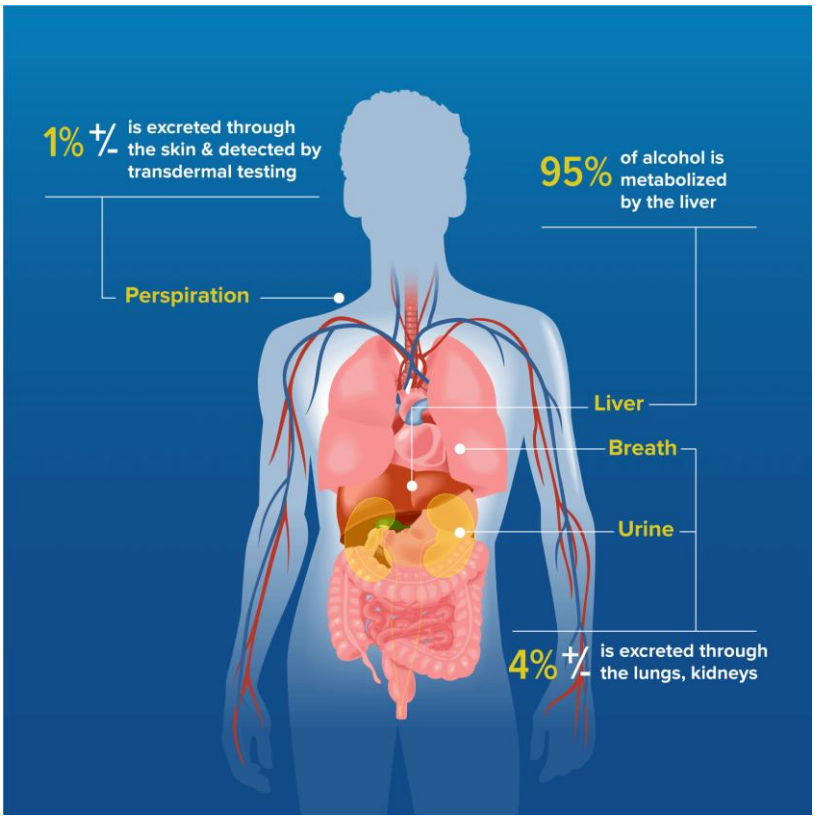

Figure 1-1. Metabolism of alcohol in the body. 1% of the alcohol consumed is secreted through the skin, which is used to measure the amount of alcohol in the body by TAC monitors. (Image from: https://www.scramsystems.com/monitoring/scram-continuous-alcohol-monitoring/transdermal-testing/)

Table 1-1

Features of alcohol measurement methods

	Traditional self-reported measures	EMA	Breathalysers	Transdermal monitors
Features				
Low potential for human error (biases)	Low	Medium	Medium	High
Multiple measurements over time	Low	High	Medium	High
Low participant burden	Medium	Low	Low	High

Note. "High" scores are preferred. The scores are interpreted from published work and limitations reported in previous literature, as discussed in the text.

Available transdermal measurement devices

This section summarises information about the current and most important TAC measurement devices available, but due to rapid developments in transdermal technology, may not be comprehensive.

SCRAM. The secure continuous remote alcohol monitor (SCRAM™), an ankle bracelet developed by Alcohol Monitoring Systems, Inc. (AMS, Littleton), is the most widely used and studied transdermal monitor to date (Figure 1-2A) (Leffingwell et al., 2013). The SCRAM monitor was developed for and is most widely used in the justice system, for monitoring convicted offenders on compliance orders. It measures the ethanol secreted through the skin using fuel cell technology, and is worn tightly against the skin around the ankle. The monitor analyses the transdermal ethanol vapour close to the skin using a platinum-based electrochemical fuel cell at 30-minute intervals, and

determines the grams per decilitre (g/dL) of ethanol in the body through catalytic alcohol oxidation; no action is required from the wearer, who is unable to see the results. TAC data is stored on the device and can be uploaded to a computer via the internet. The data transfer goes through a secure network developed by AMS, which can then be accessed through SCRAMNet, a secure web-based server. According to AMS, the SCRAM monitor can be worn for up to six months without battery change, with all readings date and time-stamped (Alcohol Monitoring Services, 2022). Because the monitor was developed for use in the justice system, it contains circumvention sensors that measure skin temperature (to check whether something is blocking the sensor) and contact with the skin (to check device removal) to detect tampering (Alcohol Monitoring Services, 2022). The fuel cell used in the monitor requires air to be pumped across the sensor to pick up the alcohol vapour, which makes the devices bulky (approximately the size of a deck of cards, weighing about 230 grams). Some studies have found that the clamping system and the bulkiness of the monitor can make it uncomfortable and embarrassing to wear the device, restricting its use for research purposes (Caluzzi et al. 2019; Marques and McKnight 2007).

A. **B.** **C**. **D.**

Figure 1-2. **A**. SCRAM-CAM ankle bracelet, **B.** WrisTAS, **C.** ION RAP, and **D**. BACtrack Skyn

WrisTAS. The WrisTAS (Figure 1-2B) was developed by Giner, Inc. (Newton, Massachusetts) to monitor alcohol abstinence rather than provide BAC estimates (Swette, Griffith, and La Conti 1997). Like the SCRAM, it measures skin temperature and contact to skin through circumvention

detection sensors. Alcohol measurements can be taken at any interval. A sensor placed on the skin measures the constant electrochemical oxidation of ethanol vapour to acetic acid, which is interpreted as a direct measure of the ethanol concentration in the blood. This battery-operated device stores data offline for days or weeks, which can then be downloaded to a computer. The biggest difference between the SCRAM and WrisTAS monitors is that the latter is worn around the wrist, is smaller and has the appearance of a wristwatch. The WrisTAS is not commercially available, but has been used in several validation studies (Bond et al. 2014; Leffingwell et al. 2013; Simons et al. 2015).

ION RAP. ION Wearable, a start-up company located in California, describes its device as the first wearable option for self-monitoring alcohol consumption, and its wristbands are available on the commercial market. Unlike the fuel cell technology used in the SCRAM-CAM, the ION RAP wristband employs an enzymatic detection pathway, with TAC being measured via a raw current (Lansdorp et al., 2019) (Figure 1-2C). Use of the enzyme alcohol oxidase is said to detect not just heavy drinking events, but low to moderate drinking (Lansdorp et al., 2019). The ION RAP is a notably compact device, smaller than either the SCRAM or WrisTAS, and is worn around the wrist, with a rechargeable battery and Bluetooth communication to a smartphone application. It uses a disposable cartridge system that must be replaced with a new cartridge after 24 hours. Using this cartridge system, the ION RAP overcomes the problem of sensor fouling and degradation that occurs with fuel cell sensors, which should increase the reliability of the TAC data (Allan et al., 2017; Campbell et al., 2018; Lansdorp et al., 2019).

Skyn BACtrack. BACtrack Inc, based in San Francisco, has been developing breathalysers for BAC testing since 2001. Its Skyn transdermal wrist monitor was released for research purposes in 2019 (Figure 1-2D) (Anon 2022). This monitor, like the SCRAM, uses fuel cell technology; however, the Skyn relies solely on a passive flow of air containing ethanol across the device, rather than the air pump used in the SCRAM. The Skyn is thus much smaller than the SCRAM, and offers a higher sampling frequency (about every 20 seconds). The Skyn is about the same size as the ION RAP and features smartphone integration via Bluetooth connection, allowing individuals to examine their real-

time intoxication levels using a smartphone application. To be closer to the veins, transporting the alcohol, the monitor is worn on the inside of the wrist to maximise sensitivity and decrease the delay of alcohol measurement (Fairbairn et al. 2019).

Research rationale and objectives

A valid and reliable TAC monitor will allow researchers to gather high-quality information on alcohol consumption. Such monitors could also help individuals make more informed decisions about their alcohol consumption, and thus contribute to better public health outcomes by reducing alcohol-related harms.

There is scant literature offering evidence about the reliability and validity of TAC devices. Leffingwell and colleagues (2013) published a systematic review almost a decade ago, but only SCRAM and WrisTAS monitors were available at the time. With the rapid technological developments in transdermal technology, and newer-generation wristbands being developed and released, it is important to systematically reassess the current state of knowledge about TAC technology. Further, even though several studies have evaluated the performance of the SCRAM monitors, few researchers have assessed their test–retest reliability. As mentioned earlier, test–retest reliability includes both reliability and agreement. Reliability is normally measured using correlation analysis, with agreement referring to the ability to produce highly similar measures, normally evaluated using Bland-Altman plots (Berchtold 2016). Validity is the extent to which the TAC results accurately measure the amount of alcohol, measured with reference to other measures of alcohol consumption that have been proven to give reliable measure of BAC such as breathalysers (Hammersley 1987; Heale and Twycross 2015; Sakai et al. 2006). Additionally, several TAC monitors have been validated in naturalistic settings by comparing them against retrospective self-reported measures. However, to date few studies have tested their performance against EMA surveys (Susan E. Luczak et al. 2015; Simons et al. 2015).

Transdermal alcohol monitoring technology is developing rapidly; alcohol researchers need to understand the devices' accuracy and validity. In order to expand knowledge about the test–retest reliability and validity of available TAC monitors, this research had the following aims:

1) systematically review the literature on test–retest reliability and validity of TAC monitors and identify research gaps;

2) study the test–retest reliability and validity of the SCRAM-CAM ankle bracelet and conduct the first controlled laboratory study of a prototype of the ION RAP wristband; and

3) investigate the validity of the SCRAM-CAM under naturalistic settings by comparing the TAC data against ecological momentary assessments.

Outline of the thesis

Chapter 2 outlines the details about the methods used in the laboratory alcohol administration studies described in Chapters 4–6. The methods section below provides an overview of. Here the recruitment process, participants and the laboratory alcohol administration session protocol used to collect the data analysed in Chapters 3 to 6 are described.

Chapters 3-7 of this thesis consist of either published or submitted articles. At the start of these Chapters, it is stated which article the Chapter relates to.

Chapter 3 provides a comprehensive overview of the evidence on the test–retest reliability and validity of TAC monitors published from 2013 onwards, extending the most recent review substantially (Leffingwell et al. 2013). This chapter includes a review of literature on TAC detection, recent advances toward the calculation of estimated BrAC (eBrAC) and number of standard drinks from TAC data, and gaps in our knowledge and future directions for research to improve the robustness of TAC data collected. I summarise current knowledge about the utility of the TAC monitors in both clinical and research settings, particularly recent advances toward using TAC data to obtain estimates of peak alcohol consumption (peak BrAC) and number of standard drinks.

Chapters 4 to 6 provide new empirical evidence on the test–retest reliability and validity of the SCRAM-CAM ankle bracelet and a prototype of a new-generation alcohol monitor, the ION RAP, in the laboratory. Chapter 4 describes the first assessment of the reliability of the SCRAM-CAM TAC monitor by testing two monitors in parallel, and my examination of the effects of individual differences, such and gender and body mass index (BMI), on the reliability measures. This is important because high reliability will increase researchers' confidence in the use of SCRAM-CAM to measure levels of alcohol consumption.

The work presented in Chapter 5 extends that in Chapter 4 by assessing the agreement as an assessment of the test–retest reliability of the SCRAM-CAM and presenting the first empirical evidence of the test–retest reliability of the ION RAP. I present the agreement between the TAC data from two simultaneously worn SCRAM-CAMs, and the reliability and agreement of TAC data generated from two simultaneously worn ION RAP wristbands. I also investigated the effect of food consumption and drinking rate on the reliability and agreement of the SCRAM-CAM TAC measurements.

In Chapter 6, I outline the validity of the TAC data as measured by the SCRAM-CAM and ION RAP by comparing the TAC data to BrAC data. More precisely, I tested whether TAC, measured through an enzymatic detection pathway as utilised by the ION RAP, shows significantly decreased delays in alcohol detection and improved correlations to BrAC as compared to the SCRAM-CAM fuel cell technology.

Chapter 7 describes my tests of the SCRAM-CAM in a naturalistic setting and comparison of its TAC data to the number of drinks consumed as reported in EMA surveys. Specifically, this chapter reports on the feasibility of using EMA surveys and SCRAM-CAMs to measure alcohol consumption over the course of a day for spectators at Australian Rules football (AFL) matches. This enabled me to study the validity of SCRAM-CAM in uncontrolled, naturalistic settings during periods of potential high consumption – heavy alcohol consumption has been recorded among spectators of various sports (Lloyd et al. 2013; Miller et al. 2013; Wolfe, Martinez, and Scott 1998).

Finally, Chapter 8 contains a summary and discussion of the findings presented in this thesis, and integrates the findings into the broader literature. This chapter focuses mainly on the test–retest reliability and validity of TAC monitors. The limitations and strengths of the research, recommendations and implications for future research are outlined.

References

Alcohol Monitoring Services. 2022. "Alcohol Monitoring Services." (https://www.scramsystems.com/our-company/about-us/).

Allan, Julaine, Anton Clifford, Patrick Ball, Margaret Alston, and Peter Meister. 2012. "'You're Less Complete If You Haven't Got a Can in Your Hand': Alcohol Consumption and Related Harmful Effects in Rural Australia: The Role and Influence of Cultural Capital." *Alcohol and Alcoholism* 47(5):624–29. doi: 10.1093/ALCALC/AGS074.

BACtrack Skyn. 2022. "BACtrack Skyn™." (https://skyn.bactrack.com/).

Australian Bureau of Statistics. 2010. "Spectator Attendance at Sporting Events, Australia (Cat. No. 4174.0)." *Canberra: ABS*.

Australian Institute of Health and Welfare. 2020. "National Drug Strategy Household Survey 2019." *Canberra: AIHW* 90.

Australian Institute of Health and Welfare. 2018. *Australian Burden of Disease Study: Impact and Causes of Illness and Death in Australia 2018*.

Barnett, Nancy P., E. B. Meade, and T. R. Glynn. 2014. "Predictors of Detection of Alcohol Use Episodes Using a Transdermal Alcohol Sensor." *Experimental and Clinical Psychopharmacology* 22(1):86–96. doi: 10.1037/a0034821.

Barnett, Nancy P., Jennifer Tidey, James G. Murphy, Robert Swift, and Suzanne M. Colby. 2011. "Contingency Management for Alcohol Use Reduction: A Pilot Study Using a Transdermal Alcohol Sensor." *Drug and Alcohol Dependence* 118(2–3):391–99.

Berchtold, André. 2016. "Test–Retest: Agreement or Reliability?:" *Http://Dx.Doi.Org/10.1177/2059799116672875* 9. doi: 10.1177/2059799116672875.

Blair, A. L., A. L. Chiaf, E. K. Crockett, T. K. Teague, and J. M. Croff. 2021. "Validation of Hair Ethyl Glucuronide Using Transdermal Monitoring and Self-Reported Alcohol Use in Women of Childbearing Potential." *Neuropsychopharmacology Reports* 41(2):144–51. doi: https://dx.doi.org/10.1002/npr2.12151.

Bland, Martin J., and Douglas G. Altman. 1986. "Statistical Methods for Assessing Agreement between Two Methods of Clinical Measurement." *The Lancet* 327(8476):307–10. doi: 10.1016/S0140-6736(86)90837-8.

Bond, Jason C., Thomas K. Greenfield, Deidre Patterson, and William C. Kerr. 2014. "Adjustments for Drink Size and Ethanol Content: New Results from a Self-Report Diary and Transdermal Sensor Validation Study." *Alcoholism: Clinical and Experimental Research* 38(12):3060–67. doi: http://dx.doi.org/10.1111/acer.12589.

Bühler, Mira, Sabine Vollstädt-Klein, Andrea Kobiella, Henning Budde, Laurence J. Reed, Dieter F. Braus, Christian Büchel, and Michael N. Smolka. 2010. "Nicotine Dependence Is Characterized by Disordered Reward Processing in a Network Driving Motivation." *Biological Psychiatry*

67(8):745–52. doi: 10.1016/j.biopsych.2009.10.029.

Callinan, Sarah. 2015. “How Big Is a Self-Poured Glass of Wine for Australian Drinkers?” *Drug and Alcohol Review* 34(2):207–10. doi: 10.1111/DAR.12154.

Caluzzi, Gabriel, Amy Pennay, Megan Cook, Cassandra Wright, Thomas Norman, and Emmanuel Kuntsche. 2019. “Transdermal Monitors to Assess Alcohol Consumption in Real-Time and Real-Life–a Qualitative Study on User-Experience.” *Addiction Research & Theory* 27(4):354–61.

Cicchetti, Domenic V. 1994. “Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology.” *Psychological Assessment* 6(4):284.

Cohen, B. Beth, and Daniel C. Vinson. 1995. “Retrospective Self-Report of Alcohol Consumption: Test-Retest Reliability by Telephone.” *Alcoholism: Clinical and Experimental Research* 19(5):1156–61.

Cohen, Jacob. 2013. *Statistical Power Analysis for the Behavioral Sciences*. Academic press.

Collins, Allan C., Toni N. Yeager, M. E. Lebsack, and S. Scott Panter. 1975. “Variations in Alcohol Metabolism: Influence of Sex and Age.” *Pharmacology Biochemistry and Behavior* 3(6):973–78.

Collins, R. Lorraine, Elizabeth T. Morsheimer, Saul Shiffman, Jean A. Paty, Maryann Gnys, and George D. Papandonatos. 1998. “Ecological Momentary Assessment in a Behavioral Drinking Moderation Training Program.” *Experimental and Clinical Psychopharmacology* 6(3):306–15. doi: 10.1037/1064-1297.6.3.306.

Croff, J. M., M. L. Hartwell, A. L. Chiaf, E. K. Crockett, and I. J. Washburn. 2021. “Feasibility and Reliability of Continuously Monitoring Alcohol Use among Female Adolescents and Young Adults.” *Drug & Alcohol Review* 40(7):1143–54. doi: https://dx.doi.org/10.1111/dar.13045.

Csikszentmihalyi, Mihaly, Reed Larson, M Csikszentmihalyi, and R Larson. 2014. “Validity and Reliability of the Experience-Sampling Method.” *Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi* 35–54. doi: 10.1007/978-94-017-9088-8_3.

Davis, Christopher G., Jennifer Thake, and Natalie Vilhena. 2010. “Social Desirability Biases in Self-Reported Alcohol Consumption and Harms.” *Addictive Behaviors* 35(4):302–11.

Dawson, Deborah A. 1998. “Measuring Alcohol Consumption: Limitations and Prospects for Improvement.” *Addiction* 93(7):965.

Degenhardt, Louisa, Wayne Hall, and Michael Lynskey. 2001. “Alcohol, Cannabis and Tobacco Use among Australians: A Comparison of Their Associations with Other Drug Use and Use Disorders, Affective and Anxiety Disorders, and Psychosis.” *Addiction (Abingdon, England)* 96(11):1603–14. doi: 10.1046/J.1360-0443.2001.961116037.X.

Dodd, Lori E., and Margaret S. Pepe. 2003. “Partial AUC Estimation and Regression.” *Biometrics* 59(3):614–23.

Dougherty, Donald M., Nathalie Hill-Kapturczak, Yuanyuan Liang, Tara E. Karns, Sarah L. Lake, Sharon E. Cates, and John D. Roache. 2015. “The Potential Clinical Utility of Transdermal Alcohol Monitoring Data to Estimate the Number of Alcoholic Drinks Consumed.” *Addictive Disorders & Their Treatment* 14(3):124–30. doi: 10.1097/ADT.0000000000000060.

Dulin, Patrick L., Corene E. Alvarado, James M. Fitterling, and Vivian M. Gonzalez. 2017. “Comparisons of Alcohol Consumption by Timeline Follow Back vs. Smartphone-Based Daily Interviews.” *Addiction Research & Theory* 25(3):195–200.

Durbeej, Natalie, Tobias H. Elgán, Camilla Jalling, and Johanna Gripenberg. 2017. “Alcohol

Intoxication at Swedish Football Matches: A Study Using Biological Sampling to Assess Blood Alcohol Concentration Levels among Spectators." *PLoS One* 12(11):e0188284.

van Egmond, Kelly, Cassandra J. C. Wright, Michael Livingston, and Emmanuel Kuntsche. 2020. "Wearable Transdermal Alcohol Monitors: A Systematic Review of Detection Validity, and Relationship Between Transdermal and Breath Alcohol Concentration and Influencing Factors." *Alcoholism: Clinical and Experimental Research* 44(10):1918–32.

Estrada, Felipe, and Kalle Tryggvesson. 2001. "A Part of the Game–Alcohol, Football Fans and Male Comradeship." *Nordic Studies on Alcohol and Drugs* 18(3):245–60.

Fairbairn, C. E., and D. Kang. 2019. "Temporal Dynamics of Transdermal Alcohol Concentration Measured via New-Generation Wrist-Worn Biosensor." *Alcoholism: Clinical and Experimental Research* 43(10):2060–69. doi: 10.1111/acer.14172.

Fairbairn, C. E., D. Kang, and N. Bosch. 2020. "Using Machine Learning for Real-Time BAC Estimation from a New-Generation Transdermal Biosensor in the Laboratory." *Drug & Alcohol Dependence* 216:108205. doi: https://dx.doi.org/10.1016/j.drugalcdep.2020.108205.

Fairbairn, C. E., I. G. Rosen, S. E. Luczak, and W. J. Venerable. 2019. "Estimating the Quantity and Time Course of Alcohol Consumption from Transdermal Alcohol Sensor Data: A Combined Laboratory-Ambulatory Study." *Alcohol* 01:1. doi: https://dx.doi.org/10.1016/j.alcohol.2018.08.015.

Flango, Victor E., and Fred L. Cheesman. 2009. "Effectiveness of the SCRAM Alcohol Monitoring Device: A Preliminary Test." *Drug Court Review* 6(2):109–34.

Gibb, Kenneth A., Albert S. Yee, Charles C. Johnston, Steven D. Martin, and Richard M. Nowak. 1984. "Accuracy and Usefulness of a Breath Alcohol Analyzer." *Annals of Emergency Medicine* 13(7):516–20. doi: 10.1016/S0196-0644(84)80517-X.

Gilligan, Conor, Kristen G. Anderson, Benjamin O. Ladd, Yun Ming Yong, and Michael David. 2019. "Inaccuracies in Survey Reporting of Alcohol Consumption." *BMC Public Health* 19(1):1–11. doi: 10.1186/S12889-019-7987-3/TABLES/4.

Glassman, Tavis, Chudley E. Werch, Edessa Jobli, and Hui Bian. 2007. "Alcohol-Related Fan Behavior on College Football Game Day." *Journal of American College Health* 56(3):255–60.

Gmel, Gerhard, Emmanuel Kuntsche, and Jürgen Rehm. 2011. "Risky Single-Occasion Drinking: Bingeing Is Not Bingeing." *Addiction* 106(6):1037–45. doi: 10.1111/J.1360-0443.2010.03167.X.

Gmel, Gerhard, and Jürgen Rehm. 2004. "Measuring Alcohol Consumption:" *Contemporary Drug Problems* 31(3):467–540. doi: 10.1177/009145090403100304.

Greenfield, Thomas K., and William C. Kerr. 2008. "Alcohol Measurement Methodology in Epidemiology: Recent Advances and Opportunities." *Addiction (Abingdon, England)* 103(7):1082. doi: 10.1111/J.1360-0443.2008.02197.X.

Gunn, R. L., J. A. Steingrimsson, J. E. Merrill, T. Souza, and N. Barnett. 2021. "Characterising Patterns of Alcohol Use among Heavy Drinkers: A Cluster Analysis Utilising Alcohol Biosensor Data." *Drug & Alcohol Review* 40(7):1155–64. doi: https://dx.doi.org/10.1111/dar.13306.

Hammersley, Martyn. 1987. "Some Notes on the Terms 'Validity' and 'Reliability.'" *British Educational Research Journal* 13(1):73–82. doi: 10.1080/0141192870130107.

Heale, Roberta, and Alison Twycross. 2015. "Validity and Reliability in Quantitative Studies." *Evidence-Based Nursing* 18(3):66–67. doi: 10.1136/EB-2015-102129.

Hill-Kapturczak, Nathalie, Sarah L. Lake, John D. Roache, Sharon E. Cates, Yuanyuan Liang, and Donald M. Dougherty. 2014. "Do Variable Rates of Alcohol Drinking Alter the Ability to Use

Transdermal Alcohol Monitors to Estimate Peak Breath Alcohol and Total Number of Drinks?" *Alcoholism: Clinical and Experimental Research* 38(10):2517–22. doi: http://dx.doi.org/10.1111/acer.12528.

Hill-Kapturczak, Nathalie, John D. Roache, Yuanyuan Liang, Tara E. Karns, Sharon E. Cates, and Donald M. Dougherty. 2015. "Accounting for Sex-Related Differences in the Estimation of Breath Alcohol Concentrations Using Transdermal Alcohol Monitoring." *Psychopharmacology* 232(1):115–23. doi: 10.1007/s00213-014-3644-9.

Hufford, Michael R., Alan L. Shields, Saul Shiffman, Jean Paty, and Mark Balabanis. 2002. "Reactivity to Ecological Momentary Assessment: An Example Using Undergraduate Problem Drinkers." *Psychology of Addictive Behaviors* 16(3):205–11. doi: 10.1037/0893-164X.16.3.205.

Hurzeler, Tristan, Nicholas A. Buckley, Firouzeh Noghrehchi, Peter Malouf, Andrew Page, Jennifer L. Schumann, and Kate M. Chitty. 2021. "Alcohol-Related Suicide across Australia: A Geospatial Analysis." *Australian and New Zealand Journal of Public Health* 45(4):394–99. doi: 10.1111/1753-6405.13122.

Institute, National Drug Research. n.d. "Social and Economic Costs of Substance Use." *2021*.

Johnson, Richard D., Michael Morowitz, Anne F. Maddox, David J. Shearman. 1991. "Cigarette Smoking and Rate of Gastric Emptying: Effect on Alcohol Absorption." *BMJ : British Medical Journal* 302(6767):20. doi: 10.1136/BMJ.302.6767.20.

Jones, Andrew, Danielle Remmerswaal, Ilse Verveer, Eric Robinson, Ingmar H. A. Franken, Cheng K. Fred Wen, and Matt Field. 2018. "Compliance with Ecological Momentary Assessment Protocols in Substance Users: A Meta-Analysis." doi: 10.1111/add.14503.

Kalist, David E., and Daniel Y. Lee. 2016. "The National Football League: Does Crime Increase on Game Day?" *Journal of Sports Economics* 17(8):863–82.

Karns-Wright, Tara E., Donald M. Dougherty, Nathalie Hill-Kapturczak, Charles W. Mathias, and John D. Roache. 2018. "The Correspondence between Transdermal Alcohol Monitoring and Daily Self-Reported Alcohol Consumption." *Addictive Behaviors* 85:147–52. doi: 10.1016/j.addbeh.2018.06.006.

Karns-Wright, Tara E., John D. Roache, Nathalie Hill-Kapturczak, Yuanyuan Liang, Jillian Mullen, and Donald M. Dougherty. 2017. "Time Delays in Transdermal Alcohol Concentrations Relative to Breath Alcohol Concentrations." *Alcohol and Alcoholism* 52(1):35–41. doi: http://dx.doi.org/10.1093/alcalc/agw058.

Kelly, Thomas M., John E. Donovan, Janet M. Kinnane, and David M. C. D. Taylor. 2002. "A Comparison of Alcohol Screening Instruments Among Under-Aged Drinkers Treated in Emergency Departments." *Alcohol and Alcoholism* 37(5):444–50. doi: 10.1093/ALCALC/37.5.444.

Kerr, William C., Thomas K. Greenfield, Jennifer Tujague, and Stephan E. Brown. 2005. "A Drink Is a Drink? Variation in the Amount of Alcohol Contained in Beer, Wine and Spirits Drinks in a US Methodological Sample." *Alcoholism: Clinical and Experimental Research* 29(11):2015–21. doi: 10.1097/01.ALC.0000187596.92804.BD.

Kerr, William C., Deidre Patterson, Mary Albert Koenen, and Thomas K. Greenfield. 2008. "Alcohol Content Variation of Bar and Restaurant Drinks in Northern California." *Alcoholism, Clinical and Experimental Research* 32(9):1623–29. doi: 10.1111/J.1530-0277.2008.00741.X.

Khadjesari, Zarnie, Elizabeth Murray, Eleftheria Kalaitzaki, Ian R. White, Jim McCambridge, Christine Godfrey, and Paul Wallace. 2009. "Test–Retest Reliability of an Online Measure of Past Week Alcohol Consumption (the TOT-AL), and Comparison with Face-to-Face Interview." *Addictive Behaviors* 34(4):337–42.

Kingsland, Melanie, Luke Wolfenden, Bosco C. Rowland, Karen E. Gillham, Vanessa J. Kennedy, Robyn L. Ramsden, Richard W. Colbran, Sarah Weir, and John H. Wiggers. 2013. "Alcohol Consumption and Sport: A Cross-Sectional Study of Alcohol Management Practices Associated with at-Risk Alcohol Consumption at Community Football Clubs." *BMC Public Health* 13(1):1–9.

Kuntsche, E., and F. Labhart. 2014. "The Future Is Now--Using Personal Cellphones to Gather Data on Substance Use and Related Factors." *Addiction* 109(7):1052–53. doi: 10.1111/ADD.12540.

Kuntsche, Emmanuel, and Raimondo Bruno. 2015. "Moody Booze: Introducing the Special Section on Affect Regulation and Substance Use." *Drug and Alcohol Review* 34(6):569–70. doi: 10.1111/DAR.12349.

Kuntsche, Emmanuel, Paul Dietze, and Rebecca Jenkinson. 2014. "Understanding Alcohol and Other Drug Use during the Event." *Drug and Alcohol Review* 33(4):335–37. doi: 10.1111/DAR.12171.

Kuntsche, Emmanuel, and Florian Labhart. 2012. "Investigating the Drinking Patterns of Young People over the Course of the Evening at Weekends." *Drug and Alcohol Dependence* 124(3):319–24. doi: 10.1016/J.DRUGALCDEP.2012.02.001.

Kuntsche, Emmanuel, and Florian Labhart. 2013a. "Using Personal Cell Phones for Ecological Momentary Assessment: An Overview of Current Developments." *European Psychologist* 18(1):3–11. doi: 10.1027/1016-9040/A000127.

Kuntsche, Emmanuel, and Florian Labhart. 2013b. "Using Personal Cell Phones for Ecological Momentary Assessment." *European Psychologist*.

Labhart, Florian, Skanda Muralidhar, Benoit Massé, Lakmal Meegahapola, Emmanuel Kuntsche, and Daniel Gatica-Perez. 2021. "Ten Seconds of My Nights: Exploring Methods to Measure Brightness, Loudness and Attendance and Their Associations with Alcohol Use from Video Clips." *PLOS ONE* 16(4):e0250443. doi: 10.1371/JOURNAL.PONE.0250443.

Labhart, Florian, Thanh Trung Phan, Daniel Gatica-Perez, and Emmanuel Kuntsche. 2021. "Shooting Shots: Estimating Alcoholic Drink Sizes in Real Life Using Event-Level Reports and Annotations of Close-up Pictures." *Drug and Alcohol Review* 40(7):1228–38. doi: 10.1111/DAR.13212.

Labhart, Florian, Darshan Santani, Jasmine Truong, Flavio Tarsetti, Olivier Bornet, Sara Landolt, Daniel Gatica-Perez, and Emmanuel Kuntsche. 2017. "Development of the Geographical Proportional-to-Size Street-Intercept Sampling (GPSIS) Method for Recruiting Urban Nightlife-Goers in an Entire City." *Https://Doi.Org/10.1080/13645579.2017.1293928* 20(6):721–36. doi: 10.1080/13645579.2017.1293928.

Labhart, Florian, Flavio Tarsetti, Olivier Bornet, Darshan Santani, Jasmine Truong, Sara Landolt, Daniel Gatica-Perez, and Emmanuel Kuntsche. 2019. "Capturing Drinking and Nightlife Behaviours and Their Social and Physical Context with a Smartphone Application – Investigation of Users' Experience and Reactivity." *Https://Doi.Org/10.1080/16066359.2019.1584292* 28(1):62–75. doi: 10.1080/16066359.2019.1584292.

Lajunen, Timo, and Heikki Summala. 2003. "Can We Trust Self-Reports of Driving? Effects of Impression Management on Driver Behaviour Questionnaire Responses." *Transportation Research Part F: Traffic Psychology and Behaviour* 6(2):97–107.

Lansdorp, Bob, William Ramsay, Rashad Hamid, and Evan Strenk. 2019. "Wearable Enzymatic Alcohol Biosensor." *Sensors* 19(10):2380.

Laslett, Anne-Marie, Paul Catalano, Tanya Chikritzhs, Caroline Dale, Christopher Doran, Jason Ferris, Thameemul Jainullabudeen, Michael Livingston, Sharon Matthews, Janette Mugavin,

Robin Room, Morgan Schlotterlein, and Claire Wilkinson. 2010. “The Range and Magnitude of Alcohol’s Harm to Others.” *Victoria: AER Centre for Alcohol Policy Research, Turning Point Alcohol and Drug Centre, Eastern Health*.

Lauckner, Carolyn, Erica Taylor, Darshti Patel, and Alexis Whitmire. 2019. “The Feasibility of Using Smartphones and Mobile Breathalyzers to Monitor Alcohol Consumption among People Living with HIV/AIDS.” *Addiction Science and Clinical Practice* 14(1):1–11. doi: 10.1186/S13722-019-0174-0/TABLES/3.

Lawson, B., K. Aguir, T. Fiorido, V. Martini-Laithier, R. Bouchakour, S. Burtey, C. Reynard-Carette, and M. Bendahan. 2019. “Skin Alcohol Perspiration Measurements Using MOX Sensors.” *Sensors and Actuators, B: Chemical* 280:306–12. doi: 10.1016/j.snb.2018.09.082.

Leffingwell, Thad R., Nathaniel J. Cooney, James G. Murphy, Susan Luczak, Gary Rosen, Donald M. Dougherty, and Nancy P. J. Barnett. 2013. “Continuous Objective Monitoring of Alcohol Use: Twenty-first Century Measurement Using Transdermal Sensors.” *Alcoholism: Clinical and Experimental Research* 37(1):16–22.

Li, B., R. S. Downen, Q. Dong, N. Tran, M. LeSaux, A. C. Meltzer, and Z. Li. 2021. “A Discreet Wearable IoT Sensor for Continuous Transdermal Alcohol Monitoring - Challenges and Opportunities.” *IEEE Sensors Journal* 21(4):5322–30. doi: https://dx.doi.org/10.1109/jsen.2020.3030254.

Lin, K. C., D. Kinnamon, D. Sankhala, S. Muthukumar, and S. Prasad. 2019. “AWARE: A Wearable Awareness with Real-Time Exposure, for Monitoring Alcohol Consumption Impact through Ethyl Glucuronide Detection.” *Alcohol* 81:93–99. doi: https://dx.doi.org/10.1016/j.alcohol.2018.10.006.

Livingston, Michael, and Sarah Callinan. 2015. “Underreporting in Alcohol Surveys: Whose Drinking Is Underestimated?” *Journal of Studies on Alcohol and Drugs* 76(1):158–64.

Lloyd, Belinda, Sharon Matthews, Michael Livingston, Harindra Jayasekara, and Karen Smith. 2013. “Alcohol Intoxication in the Context of Major Public Holidays, Sporting and Social Events: A Time–Series Analysis in M Elbourne, A Ustralia, 2000–2009.” *Addiction* 108(4):701–9.

Luczak, S. E., and I. G. Rosen. 2014. “Estimating BrAC from Transdermal Alcohol Concentration Data Using the BrAC Estimator Software Program.” *Alcoholism: Clinical and Experimental Research* 38(8):2243–52. doi: 10.1111/acer.12478.

Luczak, S E, I. G. Rosen, and T. L. Wall. 2015. “Development of a Real-Time Repeated-Measures Assessment Protocol to Capture Change over the Course of a Drinking Episode.” *Alcohol and Alcoholism* 50(2):180–87. doi: 10.1093/alcalc/agu100.

Luczak, Susan E., I. Gary Rosen, and Tamara L. Wall. 2015. “Development of a Real-Time Repeated-Measures Assessment Protocol to Capture Change over the Course of a Drinking Episode.” *Alcohol and Alcoholism* 50(2):180–87. doi: 10.1093/ALCALC/AGU100.

Maher, Jessica Middlemis, Jonathan C. Markey, and Diane Ebert-May. 2013. “The Other Half of the Story: Effect Size Analysis in Quantitative Research.” *CBE—Life Sciences Education* 12(3):345–51.

Marques, Paul R., and A. Scott McKnight. 2007. *Evaluating Transdermal Alcohol Measuring Devices*. United States. National Highway Traffic Safety Administration.

Marques, Paul R., and A. Scott McKnight. 2009. “Field and Laboratory Alcohol Detection with 2 Types of Transdermal Devices.” *Alcoholism: Clinical and Experimental Research* 33(4):703–11.

May, Carl, and David Foxcroft. 1995. “Minimizing Bias in Self-Reports of Health Beliefs and Behaviours.” *Health Education Research* 10(1):107–12.

Meier, Patrick, and Helmut K. Seitz. 2008. “Age, Alcohol Metabolism and Liver Disease.” *Current Opinion in Clinical Nutrition & Metabolic Care* 11(1):21–26.

Miller, Elizabeth T., Dan J. Neal, Lisa J. Roberts, John S. Boer, Sally O. Cresslr, Jane Metrik, and G. Alan Marlatt. 2002. “Test-Retest Reliability of Alcohol Measures: Is There a Difference between Internet-Based Assessment and Traditional Methods?” *Psychology of Addictive Behaviors* 16(1):56–63.

Miller, Peter, Louise M c Donald, Stephen McKenzie, Kerry O’Brien, and Petra Staiger. 2013. “When the C Ats Are Away: The Impact of Sporting Events on Assault-and Alcohol-related Emergency Department Attendances.” *Drug and Alcohol Review* 32(1):31–38.

Monk, Rebecca Louise, Derek Heim, Adam Qureshi, and Alan Price. 2015. “‘I Have No Clue What I Drunk Last Night’ Using Smartphone Technology to Compare In-Vivo and Retrospective Self-Reports of Alcohol Consumption.” *PLOS ONE* 10(5):e0126209. doi: 10.1371/JOURNAL.PONE.0126209.

Mun, Eun-Young, Xiaoyin Li, Michael S. Businelle, Emily T. Hébert, Zhengqi Tan, Nancy P. Barnett, and Scott T. Walters. 2021. “Ecological Momentary Assessment of Alcohol Consumption and Its Concordance with Transdermal Alcohol Detection and Timeline Follow-Back Self-report Among Adults Experiencing Homelessness.” *Alcoholism: Clinical and Experimental Research* 45(4):864–76.

Nelson, Toben F., and Henry Wechsler. 2003. “School Spirits: Alcohol and Collegiate Sports Fans.” *Addictive Behaviors* 28(1):1–11.

NHMRC. 2020. “Australian Guidelines to Reduce Health Risks from Drinking Alcohol | NHMRC.” (https://www.nhmrc.gov.au/about-us/publications/australian-guidelines-reduce-health-risks-drinking-alcohol).

Niall Bolger, and Jean-Philippe Laurenceau. 2013. “Intensive Longitudinal Methods: An Introduction to Diary and Experience.” *Guilford, New York*. Retrieved January 15, 2022 (https://books.google.com.au/books?hl=en&lr=&id=5bD4LuAFq0oC&oi=fnd&pg=PP1&ots=jnjkZ0GU92&sig=yBv78VkSIjd4yYvE_AHdT8siCJA&redir_esc=y#v=onepage&q&f=false).

Norman, Thomas, Amy Peacock, Stuart G. Ferguson, Emmanuel Kuntsche, and Raimondo Bruno. 2020. “Combining Transdermal and Breath Alcohol Assessments, Real-Time Drink Logs and Retrospective Self-Reports to Measure Alcohol Consumption and Intoxication across a Multi-Day Music Festival.” *Drug and Alcohol Review*. doi: 10.1111/DAR.13215.

Palmer, Catherine. 2011. “Key Themes and Research Agendas in the Sport-Alcohol Nexus.” *Journal of Sport and Social Issues* 35(2):168–85.

Palmer, Catherine, and Kirrilly Thompson. 2007. “The Paradoxes of Football Spectatorship: On-Field and Online Expressions of Social Capital among the ‘Grog Squad.’” *Sociology of Sport Journal* 24(2):187–205.

Peacock, Amy, Raimondo Bruno, Frances H. Martin, and Andrea Carr. 2013. “The Impact of Alcohol and Energy Drink Consumption on Intoxication and Risk-Taking Behavior.” *Alcoholism: Clinical and Experimental Research* 37(7):1234–42. doi: 10.1111/ACER.12086.

Piasecki, Thomas M. 2019. “Assessment of Alcohol Use in the Natural Environment.” *Alcoholism: Clinical and Experimental Research* 43(4):564–77.

Riordan, Benjamin C., Damian Scarf, Saleh Moradi, Jayde A. M. Flett, Kate B. Carey, and Tamlin S. Conner. 2017. “The Accuracy and Promise of Personal Breathalysers for Research: Steps toward a Cost-Effective Reliable Measure of Alcohol Intoxication?” *Digital Health* 3:2055207617746752.

Riuttanen, Antti, Saara J. Jäntti, and Ville M. Mattila. 2020. “Alcohol Use in Severely Injured Trauma

Patients." *Scientific Reports 2020 10:1* 10(1):1–8. doi: 10.1038/s41598-020-74753-y.

Roache, J. D., T. E. Karns-Wright, M. Goros, N. Hill-Kapturczak, C. W. Mathias, and D. M. Dougherty. 2019. "Processing Transdermal Alcohol Concentration (TAC) Data to Detect Low-Level Drinking." *Alcohol* 01:1. doi: https://dx.doi.org/10.1016/j.alcohol.2018.08.014.

Roache, John D., Tara E. Karns, Nathalie Hill-Kapturczak, Jillian Mullen, Yuanyuan Liang, Richard J. Lamb, and Donald M. Dougherty. 2015. "Using Transdermal Alcohol Monitoring to Detect Low-Level Drinking." *Alcoholism: Clinical and Experimental Research* 39(7):1120–27. doi: http://dx.doi.org/10.1111/acer.12750.

Sakai, Joseph T., Susan K. Mikulich-Gilbertson, Robert J. Long, and Thomas J. Crowley. 2006. "Validity of Transdermal Alcohol Monitoring: Fixed and Self-regulated Dosing." *Alcoholism: Clinical and Experimental Research* 30(1):26–33.

Saldich, Emily B., Chunming Wang, I. Gary Rosen, Jay Bartroff, and Susan E. Luczak. 2021. "Effects of Stomach Content on the Breath Alcohol Concentration-Transdermal Alcohol Concentration Relationship." *Drug and Alcohol Review*. doi: 10.1111/DAR.13267.

Schafer, Theodore, and A. Walls Joseph L. 2006. *Models for Intensive Longitudinal Data* . Oxford University Press.

Shiffman, Saul. 2009. "Ecological Momentary Assessment (EMA) in Studies of Substance Use." *Psychological Assessment* 21(4):486.

Shiffman, Saul, Arthur A. Stone, and Michael R. Hufford. 2008. "Ecological Momentary Assessment." *Annu. Rev. Clin. Psychol.* 4:1–32.

Simons, Jeffrey S., Thomas A. Wills, Noah N. Emery, and Russell M. Marks. 2015. "Quantifying Alcohol Consumption: Self-Report, Transdermal Assessment, and Prediction of Dependence Symptoms." *Addictive Behaviors* 50:205–12. doi: http://dx.doi.org/10.1016/j.addbeh.2015.06.042.

Sirlanci, M., I. G. Rosen, T. L. Wall, and S. E. Luczak. 2018. "Applying a Novel Population-Based Model Approach to Estimating Breath Alcohol Concentration (BrAC) from Transdermal Alcohol Concentration (TAC) Biosensor Data." *Alcohol* 20:20. doi: https://dx.doi.org/10.1016/j.alcohol.2018.09.005.

Sobell, Linda C., Mark B. Sobell, Gloria I. Leo, and Anthony Cancilla. 1988. "Reliability of a Timeline Method: Assessing Normal Drinkers' Reports of Recent Drinking and a Comparative Evaluation across Several Populations." *British Journal of Addiction* 83(4):393–402.

Sobell, Mark B., Linda C. Sobell, Felix Klajner, Daniel Pavan, and Ellen Basian. 1986. "The Reliability of a Timeline Method for Assessing Normal Drinker College Students' Recent Drinking History: Utility for Alcohol Research." *Addictive Behaviors* 11(2):149–61.

Stockwell, Tim, Susan Donath, Mark Cooper-Stanbury, Tanya Chikritzhs, Paul Catalano, and Cid Mateo. 2004. "Under-Reporting of Alcohol Consumption in Household Surveys: A Comparison of Quantity-Frequency, Graduated-Frequency and Recent Recall." *Addiction (Abingdon, England)* 99(8):1024–33. doi: 10.1111/J.1360-0443.2004.00815.X.

Swette, L., A. Griffith, and A. La Conti. 1997. "Potential and Diffusion Controlled Solid Electrolyte Sensor for Continuous Measurement of Very Low Levels of Transdermal Alcohol." *Giner, Inc., Waltham, MA. Appl. No. 840802*.

Swift, Robert. 2003. "Direct Measurement of Alcohol and Its Metabolites." *Addiction* 98(SUPPL. 2):73–80. doi: 10.1046/J.1359-6357.2003.00605.X.

Swift, Robert M., Christopher S. Martin, Larry Swette, Anthony LaConti, and Nancy Kackley. 1992. "Studies on a Wearable, Electronic, Transdermal Alcohol Sensor." *Alcoholism: Clinical and*

Experimental Research 16(4):721–25. doi: 10.1111/J.1530-0277.1992.TB00668.X.

Trull, Timothy J., and Ulrich Ebner-Priemer. 2014. "The Role of Ambulatory Assessment in Psychological Science." *Current Directions in Psychological Science* 23(6):466–70. doi: 10.1177/0963721414550706.

Vestal, Robert E., Elizabeth A. McGuire, Jordan D. Tobin, Reubin Andres, Arthur H. Norris, and Esteban Mezey. 1977. "Aging and Ethanol Metabolism." *Clinical Pharmacology & Therapeutics* 21(3):343–54. doi: 10.1002/CPT1977213343.

Voas, Robert B., Robert L. DuPont, Stephen K. Talpins, and Corinne L. Shea. 2011. "Towards a National Model for Managing Impaired Driving Offenders." *Addiction* 106(7):1221–27. doi: 10.1111/J.1360-0443.2010.03339.X.

Voogt, Carmen, Emmanuel Kuntsche, Marloes Kleinjan, Evelien Poelen, and Rutger Engels. 2014. "Using Ecological Momentary Assessment to Test the Effectiveness of a Web-Based Brief Alcohol Intervention Over Time Among Heavy-Drinking Students: Randomized Controlled Trial." *J Med Internet Res 2014;16(1):E5 Https://Www.Jmir.Org/2014/1/E5* 16(1):e2817. doi: 10.2196/JMIR.2817.

Wolfe, Jeannette, Ricardo Martinez, and Warren A. Scott. 1998. "Baseball and Beer: An Analysis of Alcohol Consumption Patterns among Male Spectators at Major-League Sporting Events." *Annals of Emergency Medicine* 31(5):629–32.

Wray, Tyler B., Jennifer E. Merrill, and Peter M. Monti. 2014. "Using Ecological Momentary Assessment (EMA) to Assess Situation-Level Predictors of Alcohol Use and Alcohol-Related Consequences." *Alcohol Research: Current Reviews* 36(1):19.

Wright, Cassandra, Paul M. Dietze, Paul A. Agius, Emmanuel Kuntsche, Michael Livingston, Oliver C. Black, Robin Room, Margaret Hellard, and Megan S. C. Lim. 2018. "Mobile Phone-Based Ecological Momentary Intervention to Reduce Young Adults' Alcohol Use in the Event: A Three-Armed Randomized Controlled Trial." *JMIR MHealth and UHealth* 6(7):e149.

Züll, C. 2016. "Open-Ended Questions." *GESIS Survey Guidelines* 3.

2. Methods

Specific methodological information is provided in each of the papers that make up Chapters 3–7. In this chapter, I provide greater detail about the methods used in the laboratory alcohol administration studies described in Chapters 4–6.

Study design

A laboratory alcohol-administration design was used to collect qualitative data to examine both the test–retest reliability and validity of the SCRAM-CAM ankle bracelet and a prototype of the ION RAP.

Participant recruitment

We used a convenience sampling approach involving street intercepts, online advertisements (on Facebook and Instagram) and snowball sampling from those recruits. The recruitment methods are explained in detail below.

Online recruitment

La Trobe University's Digital Marketing and Customer Engagement team was consulted for the design of the Facebook and Instagram advertisement to ensure its attractiveness to the target audience. The advertisement was live in 2019 between the 25th of July and the 18th of August, targeted adults aged 18–35 years living in metropolitan Melbourne, and included a link which directed them to a website with study information and a screening survey hosted on Qualtrics (survey software from Seattle, Washington, and Provo, Utah, USA). Screening consisted of basic demographic questions (age, gender, location, weight, height, ethnicity), an assessment of drinking behaviour using the 3-item Alcohol Use Disorders Identification Test – Concise (AUDIT-C) questionnaire [19], a question about daily smoker status (Yes/No), and a question about pregnancy or breastfeeding (Yes/No). The participants were asked to sign a consent form indicating that they were willing to be contacted for a

more detailed outlining of the study and that they were aware that the information given was to be used for research purposes.

In-person street approaches

My colleagues and I recruited participants from among attendees at four AFL matches held at a major Melbourne stadium over two days in July 2019. Six researchers were located across the grounds outside the stadium, and approached spectators using the street-intercept recruitment method (Labhart et al. 2017), approaching every tenth person to invite them to complete the screening survey on an iPad. The team also handed out flyers with a QR code for those who did not want to complete the survey on the spot. Again, participants had to sign a consent form indicating that they were willing to be contacted for a more detailed outlining of the study and that they were aware that the information given was to be used for research purposes.

Participant screening

We telephoned potential participants who met the inclusion and exclusion criteria listed below to discuss their involvement in the study. Any person who indicated willingness to participate in the validation study underwent further screening to ensure their eligibility.

Inclusion criteria

- Aged 18–35 years.
- Being inside the healthy BMI range (18-30).
- Able to read and understand English.
- Able to provide informed consent.
- Had consumed at least two standard drinks of alcohol during one drinking occasion in the preceding month.
- Willing to participate and to come into the laboratory for three sessions of 8–10 hours.

Exclusion criteria

- Aged under 18 or above age 35: Participants had to be aged at least 18 years to be of legal drinking age in Australia, and 35 or younger to keep the sample reasonably homogeneous for comparison purposes because age-related changes in body composition can result in high BAC levels for a given amount of alcohol consumption (Collins et al. 1975; Meier and Seitz 2008; Vestal et al. 1977).

- Being under- or overweight (BMI range outside 18–30): BMI has been found to affect the BAC levels, like age; this criterion was intended to keep the sample from becoming overly heterogenous (Collins et al. 1975).

- Daily smoking: Smoking has been found to be associated with slower alcohol absorption and lower peak alcohol levels (Bühler et al. 2010; Johnson et al. 1991).

- High-risk drinking: We did not expect the low amount of alcohol administered to our participants to pose a mental health risk for the participants; however, participants scoring higher than 10 on the AUDIT-C were excluded, out of concern about encouraging individuals with a dependence disorder to drink.

- Being pregnant or lactating: Women who were pregnant, likely to be pregnant or breastfeeding were excluded due to alcohol's potentially harmful effects on the foetus (Gmel et al. 2011).

- Current psychological distress: To assess the participant's level of psychological distress, the Kessler Psychological Distress Scale (K10) was used. People scoring higher than 25 on the K10 were excluded. Participation in the study, meant that participants had to consume four standard drinks, due to the harmful relationship between alcohol consumption and psychological distress such as depression and anxiety , participation could have led to potential mental harm (Degenhardt, Hall, and Lynskey 2001; Peacock et al. 2013).

- Use of any medication shown to interact with alcohol. We compiled a list of medications that could interact with alcohol and compared reported medications to this list to determine eligibility.
- Liver or kidney disease.
- Pre-existing medical conditions related to the legs, ankles or feet: All participants who indicated a pre-existing condition that would prevent them from wearing the SCRAM ankle bracelet were excluded.

Eligible participants were sent:

- an information brochure giving study details;
- a description of the test session set-up, with the full list of what was required of the participants on test days;
- a SCRAM-CAM information booklet; and
- consent forms to sign prior to participation.
- Participants were able to nominate preferred test session days over a five-week period.

Study procedure

Data collection occurred in Melbourne's central business district, to ensure easy access via public transport and from the surrounding suburbs. We leased a classroom from an education centre that had enough space for four participants plus the researchers, and a kitchen enabling us to prepare the alcoholic drinks and meals for the participants to consume. The classroom was temperature controlled (22°C) and contained tables and chairs so that participants were able to remain seated during the session to control for any temperature or movement effects on the TAC monitoring. All sessions were scheduled between Monday and Friday and commenced between 9am and 10am; the first session was held on 8th of August and the last on 13th of September 2019. Participants were

instructed to not drink any alcohol for at least 12 hours prior to attending a session, to and fast for at least three hours prior to arrival. At the start of each session, participants were asked to complete a consent form, provide a breath sample using an Australian standard certified breathalyser (see below), and complete a baseline survey asking about their last meal, last alcoholic drink and how much water or other non-alcoholic drinks (e.g., coffee, tea, milk) they had consumed that morning. In the event of a participant supplying a breath sample that showed a BrAC above 0.000, we would have had to reschedule their session. Fortunately, this did not happen in any of the sessions. At least two hours before the start of each session, trained researchers activated the cartridges for the ION RAP transdermal monitor and inserted them in the wristband as instructed by MILO Sensors (see Figure 2-1). Participants were fitted with ION RAPs and SCRAM-CAMs at least an hour prior to alcohol administration to allow for calibration of the devices. For 58 sessions participants had an ION RAP wristband fitted to one wrist, while for seven sessions participants had them fitted to both wrists. SCRAM-CAMs were fitted to both ankles of the 22 participants. Participants then consumed their alcoholic drinks at varying rates over the session, ate meals or snacks, and provided a breath sample in fixed intervals (more details below). To be able to examine the full alcohol elimination curve, participants were required to stay in the laboratory until their TAC returned to 0.00. Because the TAC curve lags behind the BrAC curve by an average of 140 minutes, most participants reached a BrAC of 0.00 around two hours before leaving. Most participants remained in the laboratory between 9.00am and 6.00pm. At the end of each session, the ION RAPs and SCRAM-CAMs were removed. At the end of the last session, the participant's weight, height and bodyfat percentage were measured using an electronic scale.

Measures

Transdermal alcohol concentration monitoring

SCRAM-CAM

Twenty SCRAM-CAMs were assigned randomly to participants and sessions, the monitors used in each session noted to facilitate tracking of any faulty monitors. The monitors were activated

and connected to AMS's online monitoring software, SCRAM Optix™, using the SCRAM Direct Connect™ software. The SCRAM-CAMs are programmed to sample TAC every 30 minutes. To adapt to the participant and ensure accuracy, the monitors have an in-built calibration process during which they sample TAC in 5-minute intervals. To allow for this calibration process, participants were fitted with the bracelets at least an hour prior to drinking. After each session, we cleaned all devices with alcohol-free wipes and replaced the face plate (the device's metal filter that touches the skin) every third or fourth session.

ION RAP

We had eight ION RAP wristbands, all pre-production prototypes. Again, these monitors were fitted randomly to the participants and across sessions, and the wristbands used in each session noted to aid tracking of any faulty monitors. The ION RAP measures TAC in the form of an electrical current (nA), which starts at 1000 nA and slowly returns to baseline (50 nA or lower) when the cartridge is inserted and once skin contact is achieved. To activate the cartridge, the gel provided in the pink tube had to be extracted using the straw and spread onto the electrode surface, avoiding the area outside the electrode (Figure 2-1). The cartridge was then closed and inserted into the wristband. ION Wearables advised us that the ideal baseline of 50 nA or lower should be reached after one hour. To allow for as much time as possible, the cartridges were prepared and inserted at least one hour before the arrival of the participants. Participants were then fitted with the ION RAP at least an hour before drinking to allow for further calibration.

When alcohol is detected, the current starts to increase; for this reason, the first point of detection was operationalised as the first point of increase after alcohol administration. The ION RAP sensor samples TAC every five seconds when connected with the application on a smartphone or tablet through Bluetooth. The ION RAP wristbands were cleaned with alcohol-free wipes after each session. The cartridges were replaced every session, so no deep cleaning was needed. The monitors were charged overnight.

Figure 2-1. Researchers activated ION RAP cartridges manually. This included extracting gel from the pink tube using the blue straw and spreading it onto the black electrode surface of the cartridge

Breath alcohol monitoring

Breath samples were taken using the Andatech Prodigy–S (Andatech Inc., Nunawading). The Andatech Prodigy–S uses the Andatech FXCell3 advanced fuel cell sensor. All breathalysers were calibrated when obtained. Breath samples were taken at 10-minute intervals and recorded online in Qualtrics. To ensure the mouthpiece and monitor were free of alcohol residue, a new disposable mouthpiece was used and a clean air sample had to be taken in the room before every test. Participants were given 20 millilitres of water to rinse their mouths before each breath test to remove residual alcohol.

Alcohol administration

The participants were asked to drink four standard drinks per session. Each drink contained 32 millilitres of vodka (40% alcohol, 10 grams of alcohol, corresponding to one Australian standard drink) mixed with 68 millilitres of soda and sugar-free cordial. To test the effects of drinking rate on TAC detection, the drinking was varied across sessions. At the beginning of sessions one and two, participants were asked to drink all four standard drinks within 16 minutes. In the other session, participants were asked to drink the two standard drinks within 8 minutes and the other two drinks after one hour within 8 minutes. The order of the sessions varying in drinking rate was randomly assigned to each participant.

Food consumption

In order to test whether food consumption affected TAC detection, the number of calories consumed varied over the sessions. The participants ate snacks (~90kcal) every hour after they started drinking their assigned drinks until four hours, when they had a meal (~500 kcal). In one of the three sessions, participants received a meal instead of a snack at the first hour after beginning drinking. The order of the sessions varying in food conditions was random for each participant; some participant's received a meal at the first hour after drinking in the first session and snacks for the others, some participants received the meal at the first hour in the last session. The participants were offered small glasses of water during the session; their water consumption was tracked, because it diluted the alcohol administered.

Reimbursement

In recognition of the time and effort the study required, the participants received $200 (approx. USD288 in September 2019) in the form of vouchers. They received a voucher worth $50 for completing the first two sessions, and $100 for the third to encourage complete participation. This was deemed an appropriate reimbursement given the large amount of time the study demanded.

Ethics

Ethics approval for this laboratory alcohol administration study was obtained on the 19th of August 2019 from the La Trobe University Human Research Ethics Committee (HEC19249).

Fieldwork and data analytic challenges and experiences

The studies I conducted for my PhD involved complex fieldwork, and the use of methods and devices that were unfamiliar to the team. This is an exciting but challenging aspect of working with emerging technologies in research – our team was the first external research team to have access to the ION RAP monitors. The high cost of the devices also meant that we had a limited time period of four weeks in which we could use them for both my laboratory study and the study on sports spectators. We had expected to have several months of piloting prior to the fieldwork, however delays

in contract negotiation between the university and developer's legal teams pushed the fieldwork out by several months, to the point of almost missing the AFL season for the year. When the time finally came to conduct the fieldwork, the long hours over both weekdays and weekends, high stress of both problem-solving technical issues and managing participant recruitment and contact caused me to experience burnout, which took time to recover from. This experience has taught me the value of substantial time investment in piloting and preparation. While the delays in legal negotiations were not in my control, in retrospect, our team should have negotiated and planned feasible timelines that made room for inevitable technical difficulties associated with emerging technologies. The devices' 'state of readiness' to use in fieldwork did not meet our expectations – we had not realised we would be sent pre-production prototypes rather than market-ready devices. Only four ION RAP wristbands were available for the first week of the study due to delays in production of the devices. Some were faulty, and others experienced battery failure. ION Sensors sent another six devices, for free, in return for feedback that they could use to improve the monitors. The biggest challenge with these monitors was charging the batteries and preparing the cartridges; the chargers had a cord connecting to the wristbands through magnetic strips; however, the cords broke sometimes, and the strips often did not stick resulting in the wristband not being charged. Further, we had to prepare the cartridges on the morning of the study by extracting the gel provided in the pink tube using the straw and spread it onto the electrode surface, avoiding the area outside the electrode (Figure 2-1). This was a very precise task and the tiniest bit of gel outside the electrode area could result in a faulty sensor, which would not be known until the data were collected.

While our team has extensive experience in using technologies for research, my PhD was our first attempt at analysing TAC data. With SCRAM-CAM monitors having been used in research for more than a decade, we were confident in our abilities to work with it and hoped that we could grow a program of research using TAC monitors from this work. Due to our limited experience with TAC data, we planned to collaborate with an international expert and received a grant to bring her to Melbourne to seek advice and input on the fieldwork and data analysis. Unfortunately, the COVID pandemic meant that travel was not possible and had other implications for our expert's availability.

Learning to work with TAC data was a challenging process, especially given the lack of studies using ION RAP monitors, which produce approximately 10,000 data points per participant per session due to the five second sampling interval. Despite the challenges involved in cleaning, processing and analysing the data, I am grateful for the opportunity that I have had to contribute to my field in understanding the value and utility of transdermal monitors.

References

Alcohol Monitoring Services. 2022. "Alcohol Monitoring Services." (https://www.scramsystems.com/our-company/about-us/).

Allan, Julaine, Anton Clifford, Patrick Ball, Margaret Alston, and Peter Meister. 2012. "'You're Less Complete If You Haven't Got a Can in Your Hand': Alcohol Consumption and Related Harmful Effects in Rural Australia: The Role and Influence of Cultural Capital." *Alcohol and Alcoholism* 47(5):624–29. doi: 10.1093/ALCALC/AGS074.

BACtrack Skyn. 2022. "BACtrack Skyn™." (https://skyn.bactrack.com/).

Australian Bureau of Statistics. 2010. "Spectator Attendance at Sporting Events, Australia (Cat. No. 4174.0)." *Canberra: ABS*.

Australian Institute of Health and Welfare. 2020. "National Drug Strategy Household Survey 2019." *Canberra: AIHW* 90.

Australian Institute of Health and Welfare. 2018. *Australian Burden of Disease Study: Impact and Causes of Illness and Death in Australia 2018*.

Barnett, Nancy P., E. B. Meade, and T. R. Glynn. 2014. "Predictors of Detection of Alcohol Use Episodes Using a Transdermal Alcohol Sensor." *Experimental and Clinical Psychopharmacology* 22(1):86–96. doi: 10.1037/a0034821.

Barnett, Nancy P., Jennifer Tidey, James G. Murphy, Robert Swift, and Suzanne M. Colby. 2011. "Contingency Management for Alcohol Use Reduction: A Pilot Study Using a Transdermal Alcohol Sensor." *Drug and Alcohol Dependence* 118(2–3):391–99.

Berchtold, André. 2016. "Test–Retest: Agreement or Reliability?:" *Http://Dx.Doi.Org/10.1177/2059799116672875* 9. doi: 10.1177/2059799116672875.

Blair, A. L., A. L. Chiaf, E. K. Crockett, T. K. Teague, and J. M. Croff. 2021. "Validation of Hair Ethyl Glucuronide Using Transdermal Monitoring and Self-Reported Alcohol Use in Women of Childbearing Potential." *Neuropsychopharmacology Reports* 41(2):144–51. doi: https://dx.doi.org/10.1002/npr2.12151.

Bland, Martin J., and Douglas G. Altman. 1986. "Statistical Methods for Assessing Agreement between Two Methods of Clinical Measurement." *The Lancet* 327(8476):307–10. doi: 10.1016/S0140-6736(86)90837-8.

Bond, Jason C., Thomas K. Greenfield, Deidre Patterson, and William C. Kerr. 2014. "Adjustments for Drink Size and Ethanol Content: New Results from a Self-Report Diary and Transdermal Sensor Validation Study." *Alcoholism: Clinical and Experimental Research* 38(12):3060–67. doi: http://dx.doi.org/10.1111/acer.12589.

Bühler, Mira, Sabine Vollstädt-Klein, Andrea Kobiella, Henning Budde, Laurence J. Reed, Dieter F.

Braus, Christian Büchel, and Michael N. Smolka. 2010. “Nicotine Dependence Is Characterized by Disordered Reward Processing in a Network Driving Motivation.” *Biological Psychiatry* 67(8):745–52. doi: 10.1016/j.biopsych.2009.10.029.

Callinan, Sarah. 2015. “How Big Is a Self-Poured Glass of Wine for Australian Drinkers?” *Drug and Alcohol Review* 34(2):207–10. doi: 10.1111/DAR.12154.

Caluzzi, Gabriel, Amy Pennay, Megan Cook, Cassandra Wright, Thomas Norman, and Emmanuel Kuntsche. 2019. “Transdermal Monitors to Assess Alcohol Consumption in Real-Time and Real-Life–a Qualitative Study on User-Experience.” *Addiction Research & Theory* 27(4):354–61.

Cicchetti, Domenic V. 1994. “Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology.” *Psychological Assessment* 6(4):284.

Cohen, B. Beth, and Daniel C. Vinson. 1995. “Retrospective Self-Report of Alcohol Consumption: Test-Retest Reliability by Telephone.” *Alcoholism: Clinical and Experimental Research* 19(5):1156–61.

Cohen, Jacob. 2013. *Statistical Power Analysis for the Behavioral Sciences*. Academic press.

Collins, Allan C., Toni N. Yeager, M. E. Lebsack, and S. Scott Panter. 1975. “Variations in Alcohol Metabolism: Influence of Sex and Age.” *Pharmacology Biochemistry and Behavior* 3(6):973–78.

Collins, R. Lorraine, Elizabeth T. Morsheimer, Saul Shiffman, Jean A. Paty, Maryann Gnys, and George D. Papandonatos. 1998. “Ecological Momentary Assessment in a Behavioral Drinking Moderation Training Program.” *Experimental and Clinical Psychopharmacology* 6(3):306–15. doi: 10.1037/1064-1297.6.3.306.

Croff, J. M., M. L. Hartwell, A. L. Chiaf, E. K. Crockett, and I. J. Washburn. 2021. “Feasibility and Reliability of Continuously Monitoring Alcohol Use among Female Adolescents and Young Adults.” *Drug & Alcohol Review* 40(7):1143–54. doi: https://dx.doi.org/10.1111/dar.13045.

Csikszentmihalyi, Mihaly, Reed Larson, M Csikszentmihalyi, and R Larson. 2014. “Validity and Reliability of the Experience-Sampling Method.” *Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi* 35–54. doi: 10.1007/978-94-017-9088-8_3.

Davis, Christopher G., Jennifer Thake, and Natalie Vilhena. 2010. “Social Desirability Biases in Self-Reported Alcohol Consumption and Harms.” *Addictive Behaviors* 35(4):302–11.

Dawson, Deborah A. 1998. “Measuring Alcohol Consumption: Limitations and Prospects for Improvement.” *Addiction* 93(7):965.

Degenhardt, Louisa, Wayne Hall, and Michael Lynskey. 2001. “Alcohol, Cannabis and Tobacco Use among Australians: A Comparison of Their Associations with Other Drug Use and Use Disorders, Affective and Anxiety Disorders, and Psychosis.” *Addiction (Abingdon, England)* 96(11):1603–14. doi: 10.1046/J.1360-0443.2001.961116037.X.

Dodd, Lori E., and Margaret S. Pepe. 2003. “Partial AUC Estimation and Regression.” *Biometrics* 59(3):614–23.

Dougherty, Donald M., Nathalie Hill-Kapturczak, Yuanyuan Liang, Tara E. Karns, Sarah L. Lake, Sharon E. Cates, and John D. Roache. 2015. “The Potential Clinical Utility of Transdermal Alcohol Monitoring Data to Estimate the Number of Alcoholic Drinks Consumed.” *Addictive Disorders & Their Treatment* 14(3):124–30. doi: 10.1097/ADT.0000000000000060.

Dulin, Patrick L., Corene E. Alvarado, James M. Fitterling, and Vivian M. Gonzalez. 2017. “Comparisons of Alcohol Consumption by Timeline Follow Back vs. Smartphone-Based Daily

Interviews." *Addiction Research & Theory* 25(3):195–200.

Durbeej, Natalie, Tobias H. Elgán, Camilla Jalling, and Johanna Gripenberg. 2017. "Alcohol Intoxication at Swedish Football Matches: A Study Using Biological Sampling to Assess Blood Alcohol Concentration Levels among Spectators." *PLoS One* 12(11):e0188284.

van Egmond, Kelly, Cassandra J. C. Wright, Michael Livingston, and Emmanuel Kuntsche. 2020. "Wearable Transdermal Alcohol Monitors: A Systematic Review of Detection Validity, and Relationship Between Transdermal and Breath Alcohol Concentration and Influencing Factors." *Alcoholism: Clinical and Experimental Research* 44(10):1918–32.

Estrada, Felipe, and Kalle Tryggvesson. 2001. "A Part of the Game–Alcohol, Football Fans and Male Comradeship." *Nordic Studies on Alcohol and Drugs* 18(3):245–60.

Fairbairn, C. E., and D. Kang. 2019. "Temporal Dynamics of Transdermal Alcohol Concentration Measured via New-Generation Wrist-Worn Biosensor." *Alcoholism: Clinical and Experimental Research* 43(10):2060–69. doi: 10.1111/acer.14172.

Fairbairn, C. E., D. Kang, and N. Bosch. 2020. "Using Machine Learning for Real-Time BAC Estimation from a New-Generation Transdermal Biosensor in the Laboratory." *Drug & Alcohol Dependence* 216:108205. doi: https://dx.doi.org/10.1016/j.drugalcdep.2020.108205.

Fairbairn, C. E., I. G. Rosen, S. E. Luczak, and W. J. Venerable. 2019. "Estimating the Quantity and Time Course of Alcohol Consumption from Transdermal Alcohol Sensor Data: A Combined Laboratory-Ambulatory Study." *Alcohol* 01:1. doi: https://dx.doi.org/10.1016/j.alcohol.2018.08.015.

Flango, Victor E., and Fred L. Cheesman. 2009. "Effectiveness of the SCRAM Alcohol Monitoring Device: A Preliminary Test." *Drug Court Review* 6(2):109–34.

Gibb, Kenneth A., Albert S. Yee, Charles C. Johnston, Steven D. Martin, and Richard M. Nowak. 1984. "Accuracy and Usefulness of a Breath Alcohol Analyzer." *Annals of Emergency Medicine* 13(7):516–20. doi: 10.1016/S0196-0644(84)80517-X.

Gilligan, Conor, Kristen G. Anderson, Benjamin O. Ladd, Yun Ming Yong, and Michael David. 2019. "Inaccuracies in Survey Reporting of Alcohol Consumption." *BMC Public Health* 19(1):1–11. doi: 10.1186/S12889-019-7987-3/TABLES/4.

Glassman, Tavis, Chudley E. Werch, Edessa Jobli, and Hui Bian. 2007. "Alcohol-Related Fan Behavior on College Football Game Day." *Journal of American College Health* 56(3):255–60.

Gmel, Gerhard, Emmanuel Kuntsche, and Jürgen Rehm. 2011. "Risky Single-Occasion Drinking: Bingeing Is Not Bingeing." *Addiction* 106(6):1037–45. doi: 10.1111/J.1360-0443.2010.03167.X.

Gmel, Gerhard, and Jürgen Rehm. 2004. "Measuring Alcohol Consumption:" *Contemporary Drug Problems* 31(3):467–540. doi: 10.1177/009145090403100304.

Greenfield, Thomas K., and William C. Kerr. 2008. "Alcohol Measurement Methodology in Epidemiology: Recent Advances and Opportunities." *Addiction (Abingdon, England)* 103(7):1082. doi: 10.1111/J.1360-0443.2008.02197.X.

Gunn, R. L., J. A. Steingrimsson, J. E. Merrill, T. Souza, and N. Barnett. 2021. "Characterising Patterns of Alcohol Use among Heavy Drinkers: A Cluster Analysis Utilising Alcohol Biosensor Data." *Drug & Alcohol Review* 40(7):1155–64. doi: https://dx.doi.org/10.1111/dar.13306.

Hammersley, Martyn. 1987. "Some Notes on the Terms 'Validity' and 'Reliability.'" *British Educational Research Journal* 13(1):73–82. doi: 10.1080/0141192870130107.

Heale, Roberta, and Alison Twycross. 2015. "Validity and Reliability in Quantitative Studies." *Evidence-Based Nursing* 18(3):66–67. doi: 10.1136/EB-2015-102129.

Hill-Kapturczak, Nathalie, Sarah L. Lake, John D. Roache, Sharon E. Cates, Yuanyuan Liang, and Donald M. Dougherty. 2014. "Do Variable Rates of Alcohol Drinking Alter the Ability to Use Transdermal Alcohol Monitors to Estimate Peak Breath Alcohol and Total Number of Drinks?" *Alcoholism: Clinical and Experimental Research* 38(10):2517–22. doi: http://dx.doi.org/10.1111/acer.12528.

Hill-Kapturczak, Nathalie, John D. Roache, Yuanyuan Liang, Tara E. Karns, Sharon E. Cates, and Donald M. Dougherty. 2015. "Accounting for Sex-Related Differences in the Estimation of Breath Alcohol Concentrations Using Transdermal Alcohol Monitoring." *Psychopharmacology* 232(1):115–23. doi: 10.1007/s00213-014-3644-9.

Hufford, Michael R., Alan L. Shields, Saul Shiffman, Jean Paty, and Mark Balabanis. 2002. "Reactivity to Ecological Momentary Assessment: An Example Using Undergraduate Problem Drinkers." *Psychology of Addictive Behaviors* 16(3):205–11. doi: 10.1037/0893-164X.16.3.205.

Hurzeler, Tristan, Nicholas A. Buckley, Firouzeh Noghrehchi, Peter Malouf, Andrew Page, Jennifer L. Schumann, and Kate M. Chitty. 2021. "Alcohol-Related Suicide across Australia: A Geospatial Analysis." *Australian and New Zealand Journal of Public Health* 45(4):394–99. doi: 10.1111/1753-6405.13122.

Institute, National Drug Research. 2021. "Social and Economic Costs of Substance Use."

Johnson, Richard D., Michael Morowitz, Anne F. Maddox, David J. Shearman, and North Terrace. 1991. "Cigarette Smoking and Rate of Gastric Emptying: Effect on Alcohol Absorption." *BMJ : British Medical Journal* 302(6767):20. doi: 10.1136/BMJ.302.6767.20.

Jones, Andrew, Danielle Remmerswaal, Ilse Verveer, Eric Robinson, Ingmar H. A. Franken, Cheng K. Fred Wen, and Matt Field. 2018. "Compliance with Ecological Momentary Assessment Protocols in Substance Users: A Meta-Analysis." doi: 10.1111/add.14503.

Kalist, David E., and Daniel Y. Lee. 2016. "The National Football League: Does Crime Increase on Game Day?" *Journal of Sports Economics* 17(8):863–82.

Karns-Wright, Tara E., Donald M. Dougherty, Nathalie Hill-Kapturczak, Charles W. Mathias, and John D. Roache. 2018. "The Correspondence between Transdermal Alcohol Monitoring and Daily Self-Reported Alcohol Consumption." *Addictive Behaviors* 85:147–52. doi: 10.1016/j.addbeh.2018.06.006.

Karns-Wright, Tara E., John D. Roache, Nathalie Hill-Kapturczak, Yuanyuan Liang, Jillian Mullen, and Donald M. Dougherty. 2017. "Time Delays in Transdermal Alcohol Concentrations Relative to Breath Alcohol Concentrations." *Alcohol and Alcoholism* 52(1):35–41. doi: http://dx.doi.org/10.1093/alcalc/agw058.

Kelly, Thomas M., John E. Donovan, Janet M. Kinnane, and David M. C. D. Taylor. 2002. "A Comparison of Alcohol Screening Instruments Among Under-Aged Drinkers Treated in Emergency Departments." *Alcohol and Alcoholism* 37(5):444–50. doi: 10.1093/ALCALC/37.5.444.

Kerr, William C., Thomas K. Greenfield, Jennifer Tujague, and Stephan E. Brown. 2005. "A Drink Is a Drink? Variation in the Amount of Alcohol Contained in Beer, Wine and Spirits Drinks in a US Methodological Sample." *Alcoholism: Clinical and Experimental Research* 29(11):2015–21. doi: 10.1097/01.ALC.0000187596.92804.BD.

Kerr, William C., Deidre Patterson, Mary Albert Koenen, and Thomas K. Greenfield. 2008. "Alcohol Content Variation of Bar and Restaurant Drinks in Northern California." *Alcoholism, Clinical and Experimental Research* 32(9):1623–29. doi: 10.1111/J.1530-0277.2008.00741.X.

Khadjesari, Zarnie, Elizabeth Murray, Eleftheria Kalaitzaki, Ian R. White, Jim McCambridge, Christine Godfrey, and Paul Wallace. 2009. "Test–Retest Reliability of an Online Measure of

Past Week Alcohol Consumption (the TOT-AL), and Comparison with Face-to-Face Interview." *Addictive Behaviors* 34(4):337–42.

Kingsland, Melanie, Luke Wolfenden, Bosco C. Rowland, Karen E. Gillham, Vanessa J. Kennedy, Robyn L. Ramsden, Richard W. Colbran, Sarah Weir, and John H. Wiggers. 2013. "Alcohol Consumption and Sport: A Cross-Sectional Study of Alcohol Management Practices Associated with at-Risk Alcohol Consumption at Community Football Clubs." *BMC Public Health* 13(1):1–9.

Kuntsche, E., and F. Labhart. 2014. "The Future Is Now--Using Personal Cellphones to Gather Data on Substance Use and Related Factors." *Addiction* 109(7):1052–53. doi: 10.1111/ADD.12540.

Kuntsche, Emmanuel, and Raimondo Bruno. 2015. "Moody Booze: Introducing the Special Section on Affect Regulation and Substance Use." *Drug and Alcohol Review* 34(6):569–70. doi: 10.1111/DAR.12349.

Kuntsche, Emmanuel, Paul Dietze, and Rebecca Jenkinson. 2014. "Understanding Alcohol and Other Drug Use during the Event." *Drug and Alcohol Review* 33(4):335–37. doi: 10.1111/DAR.12171.

Kuntsche, Emmanuel, and Florian Labhart. 2012. "Investigating the Drinking Patterns of Young People over the Course of the Evening at Weekends." *Drug and Alcohol Dependence* 124(3):319–24. doi: 10.1016/J.DRUGALCDEP.2012.02.001.

Kuntsche, Emmanuel, and Florian Labhart. 2013a. "Using Personal Cell Phones for Ecological Momentary Assessment: An Overview of Current Developments." *European Psychologist* 18(1):3–11. doi: 10.1027/1016-9040/A000127.

Kuntsche, Emmanuel, and Florian Labhart. 2013b. "Using Personal Cell Phones for Ecological Momentary Assessment." *European Psychologist*.

Labhart, Florian, Skanda Muralidhar, Benoit Massé, Lakmal Meegahapola, Emmanuel Kuntsche, and Daniel Gatica-Perez. 2021. "Ten Seconds of My Nights: Exploring Methods to Measure Brightness, Loudness and Attendance and Their Associations with Alcohol Use from Video Clips." *PLOS ONE* 16(4):e0250443. doi: 10.1371/JOURNAL.PONE.0250443.

Labhart, Florian, Thanh Trung Phan, Daniel Gatica-Perez, and Emmanuel Kuntsche. 2021. "Shooting Shots: Estimating Alcoholic Drink Sizes in Real Life Using Event-Level Reports and Annotations of Close-up Pictures." *Drug and Alcohol Review* 40(7):1228–38. doi: 10.1111/DAR.13212.

Labhart, Florian, Darshan Santani, Jasmine Truong, Flavio Tarsetti, Olivier Bornet, Sara Landolt, Daniel Gatica-Perez, and Emmanuel Kuntsche. 2017. "Development of the Geographical Proportional-to-Size Street-Intercept Sampling (GPSIS) Method for Recruiting Urban Nightlife-Goers in an Entire City." *Https://Doi.Org/10.1080/13645579.2017.1293928* 20(6):721–36. doi: 10.1080/13645579.2017.1293928.

Labhart, Florian, Flavio Tarsetti, Olivier Bornet, Darshan Santani, Jasmine Truong, Sara Landolt, Daniel Gatica-Perez, and Emmanuel Kuntsche. 2019. "Capturing Drinking and Nightlife Behaviours and Their Social and Physical Context with a Smartphone Application – Investigation of Users' Experience and Reactivity." *Https://Doi.Org/10.1080/16066359.2019.1584292* 28(1):62–75. doi: 10.1080/16066359.2019.1584292.

Lajunen, Timo, and Heikki Summala. 2003. "Can We Trust Self-Reports of Driving? Effects of Impression Management on Driver Behaviour Questionnaire Responses." *Transportation Research Part F: Traffic Psychology and Behaviour* 6(2):97–107.

Lansdorp, Bob, William Ramsay, Rashad Hamid, and Evan Strenk. 2019. "Wearable Enzymatic Alcohol Biosensor." *Sensors* 19(10):2380.

Laslett, Anne-Marie, Paul Catalano, Tanya Chikritzhs, Caroline Dale, Christopher Doran, Jason Ferris, Thameemul Jainullabudeen, Michael Livingston, Sharon Matthews, Janette Mugavin, Robin Room, Morgan Schlotterlein, and Claire Wilkinson. 2010. "The Range and Magnitude of Alcohol's Harm to Others." *Victoria: AER Centre for Alcohol Policy Research, Turning Point Alcohol and Drug Centre, Eastern Health*.

Lauckner, Carolyn, Erica Taylor, Darshti Patel, and Alexis Whitmire. 2019. "The Feasibility of Using Smartphones and Mobile Breathalyzers to Monitor Alcohol Consumption among People Living with HIV/AIDS." *Addiction Science and Clinical Practice* 14(1):1–11. doi: 10.1186/S13722-019-0174-0/TABLES/3.

Lawson, B., K. Aguir, T. Fiorido, V. Martini-Laithier, R. Bouchakour, S. Burtey, C. Reynard-Carette, and M. Bendahan. 2019. "Skin Alcohol Perspiration Measurements Using MOX Sensors." *Sensors and Actuators, B: Chemical* 280:306–12. doi: 10.1016/j.snb.2018.09.082.

Leffingwell, Thad R., Nathaniel J. Cooney, James G. Murphy, Susan Luczak, Gary Rosen, Donald M. Dougherty, and Nancy P. J. Barnett. 2013. "Continuous Objective Monitoring of Alcohol Use: Twenty-first Century Measurement Using Transdermal Sensors." *Alcoholism: Clinical and Experimental Research* 37(1):16–22.

Li, B., R. S. Downen, Q. Dong, N. Tran, M. LeSaux, A. C. Meltzer, and Z. Li. 2021. "A Discreet Wearable IoT Sensor for Continuous Transdermal Alcohol Monitoring - Challenges and Opportunities." *IEEE Sensors Journal* 21(4):5322–30. doi: https://dx.doi.org/10.1109/jsen.2020.3030254.

Lin, K. C., D. Kinnamon, D. Sankhala, S. Muthukumar, and S. Prasad. 2019. "AWARE: A Wearable Awareness with Real-Time Exposure, for Monitoring Alcohol Consumption Impact through Ethyl Glucuronide Detection." *Alcohol* 81:93–99. doi: https://dx.doi.org/10.1016/j.alcohol.2018.10.006.

Livingston, Michael, and Sarah Callinan. 2015. "Underreporting in Alcohol Surveys: Whose Drinking Is Underestimated?" *Journal of Studies on Alcohol and Drugs* 76(1):158–64.

Lloyd, Belinda, Sharon Matthews, Michael Livingston, Harindra Jayasekara, and Karen Smith. 2013. "Alcohol Intoxication in the Context of Major Public Holidays, Sporting and Social Events: A Time–Series Analysis in M Elbourne, A Ustralia, 2000–2009." *Addiction* 108(4):701–9.

Luczak, S. E., and I. G. Rosen. 2014. "Estimating BrAC from Transdermal Alcohol Concentration Data Using the BrAC Estimator Software Program." *Alcoholism: Clinical and Experimental Research* 38(8):2243–52. doi: 10.1111/acer.12478.

Luczak, S E, I. G. Rosen, and T. L. Wall. 2015. "Development of a Real-Time Repeated-Measures Assessment Protocol to Capture Change over the Course of a Drinking Episode." *Alcohol and Alcoholism* 50(2):180–87. doi: 10.1093/alcalc/agu100.

Luczak, Susan E., I. Gary Rosen, and Tamara L. Wall. 2015. "Development of a Real-Time Repeated-Measures Assessment Protocol to Capture Change over the Course of a Drinking Episode." *Alcohol and Alcoholism* 50(2):180–87. doi: 10.1093/ALCALC/AGU100.

Maher, Jessica Middlemis, Jonathan C. Markey, and Diane Ebert-May. 2013. "The Other Half of the Story: Effect Size Analysis in Quantitative Research." *CBE—Life Sciences Education* 12(3):345–51.

Marques, Paul R., and A. Scott McKnight. 2007. *Evaluating Transdermal Alcohol Measuring Devices*. United States. National Highway Traffic Safety Administration.

Marques, Paul R., and A. Scott McKnight. 2009. "Field and Laboratory Alcohol Detection with 2 Types of Transdermal Devices." *Alcoholism: Clinical and Experimental Research* 33(4):703–11.

May, Carl, and David Foxcroft. 1995. “Minimizing Bias in Self-Reports of Health Beliefs and Behaviours.” *Health Education Research* 10(1):107–12.

Meier, Patrick, and Helmut K. Seitz. 2008. “Age, Alcohol Metabolism and Liver Disease.” *Current Opinion in Clinical Nutrition & Metabolic Care* 11(1):21–26.

Miller, Elizabeth T., Dan J. Neal, Lisa J. Roberts, John S. Boer, Sally O. Cresslr, Jane Metrik, and G. Alan Marlatt. 2002. “Test-Retest Reliability of Alcohol Measures: Is There a Difference between Internet-Based Assessment and Traditional Methods?” *Psychology of Addictive Behaviors* 16(1):56–63.

Miller, Peter, Louise M c Donald, Stephen M c Kenzie, Kerry O’Brien, and Petra Staiger. 2013. “When the C Ats Are Away: The Impact of Sporting Events on Assault-and Alcohol-related Emergency Department Attendances.” *Drug and Alcohol Review* 32(1):31–38.

Monk, Rebecca Louise, Derek Heim, Adam Qureshi, and Alan Price. 2015. “‘I Have No Clue What I Drunk Last Night’ Using Smartphone Technology to Compare In-Vivo and Retrospective Self-Reports of Alcohol Consumption.” *PLOS ONE* 10(5):e0126209. doi: 10.1371/JOURNAL.PONE.0126209.

Mun, Eun-Young, Xiaoyin Li, Michael S. Businelle, Emily T. Hébert, Zhengqi Tan, Nancy P. Barnett, and Scott T. Walters. 2021. “Ecological Momentary Assessment of Alcohol Consumption and Its Concordance with Transdermal Alcohol Detection and Timeline Follow-Back Self-report Among Adults Experiencing Homelessness.” *Alcoholism: Clinical and Experimental Research* 45(4):864–76.

Nelson, Toben F., and Henry Wechsler. 2003. “School Spirits: Alcohol and Collegiate Sports Fans.” *Addictive Behaviors* 28(1):1–11.

NHMRC. 2020. “Australian Guidelines to Reduce Health Risks from Drinking Alcohol | NHMRC.” Retrieved September 10, 2021 (https://www.nhmrc.gov.au/about-us/publications/australian-guidelines-reduce-health-risks-drinking-alcohol).

Niall Bolger, and Jean-Philippe Laurenceau. 2013. “Intensive Longitudinal Methods: An Introduction to Diary and Experience.” *Guilford, New York*.

Norman, Thomas, Amy Peacock, Stuart G. Ferguson, Emmanuel Kuntsche, and Raimondo Bruno. 2020. “Combining Transdermal and Breath Alcohol Assessments, Real-Time Drink Logs and Retrospective Self-Reports to Measure Alcohol Consumption and Intoxication across a Multi-Day Music Festival.” *Drug and Alcohol Review*. doi: 10.1111/DAR.13215.

Palmer, Catherine. 2011. “Key Themes and Research Agendas in the Sport-Alcohol Nexus.” *Journal of Sport and Social Issues* 35(2):168–85.

Palmer, Catherine, and Kirrilly Thompson. 2007. “The Paradoxes of Football Spectatorship: On-Field and Online Expressions of Social Capital among the ‘Grog Squad.’” *Sociology of Sport Journal* 24(2):187–205.

Peacock, Amy, Raimondo Bruno, Frances H. Martin, and Andrea Carr. 2013. “The Impact of Alcohol and Energy Drink Consumption on Intoxication and Risk-Taking Behavior.” *Alcoholism: Clinical and Experimental Research* 37(7):1234–42. doi: 10.1111/ACER.12086.

Piasecki, Thomas M. 2019. “Assessment of Alcohol Use in the Natural Environment.” *Alcoholism: Clinical and Experimental Research* 43(4):564–77.

Riordan, Benjamin C., Damian Scarf, Saleh Moradi, Jayde A. M. Flett, Kate B. Carey, and Tamlin S. Conner. 2017. “The Accuracy and Promise of Personal Breathalysers for Research: Steps toward a Cost-Effective Reliable Measure of Alcohol Intoxication?” *Digital Health* 3:2055207617746752.

Riuttanen, Antti, Saara J. Jäntti, and Ville M. Mattila. 2020. “Alcohol Use in Severely Injured Trauma Patients.” *Scientific Reports 2020 10:1* 10(1):1–8. doi: 10.1038/s41598-020-74753-y.

Roache, J. D., T. E. Karns-Wright, M. Goros, N. Hill-Kapturczak, C. W. Mathias, and D. M. Dougherty. 2019. “Processing Transdermal Alcohol Concentration (TAC) Data to Detect Low-Level Drinking.” *Alcohol* 01:1. doi: https://dx.doi.org/10.1016/j.alcohol.2018.08.014.

Roache, John D., Tara E. Karns, Nathalie Hill-Kapturczak, Jillian Mullen, Yuanyuan Liang, Richard J. Lamb, and Donald M. Dougherty. 2015. “Using Transdermal Alcohol Monitoring to Detect Low-Level Drinking.” *Alcoholism: Clinical and Experimental Research* 39(7):1120–27. doi: http://dx.doi.org/10.1111/acer.12750.

Sakai, Joseph T., Susan K. Mikulich-Gilbertson, Robert J. Long, and Thomas J. Crowley. 2006. “Validity of Transdermal Alcohol Monitoring: Fixed and Self-regulated Dosing.” *Alcoholism: Clinical and Experimental Research* 30(1):26–33.

Saldich, Emily B., Chunming Wang, I. Gary Rosen, Jay Bartroff, and Susan E. Luczak. 2021. “Effects of Stomach Content on the Breath Alcohol Concentration-Transdermal Alcohol Concentration Relationship.” *Drug and Alcohol Review*. doi: 10.1111/DAR.13267.

Schafer, Theodore, and A. Walls Joseph L. 2006. *Models for Intensive Longitudinal Data* . Oxford University Press.

Shiffman, Saul. 2009. “Ecological Momentary Assessment (EMA) in Studies of Substance Use.” *Psychological Assessment* 21(4):486.

Shiffman, Saul, Arthur A. Stone, and Michael R. Hufford. 2008. “Ecological Momentary Assessment.” *Annu. Rev. Clin. Psychol.* 4:1–32.

Simons, Jeffrey S., Thomas A. Wills, Noah N. Emery, and Russell M. Marks. 2015. “Quantifying Alcohol Consumption: Self-Report, Transdermal Assessment, and Prediction of Dependence Symptoms.” *Addictive Behaviors* 50:205–12. doi: http://dx.doi.org/10.1016/j.addbeh.2015.06.042.

Sirlanci, M., I. G. Rosen, T. L. Wall, and S. E. Luczak. 2018. “Applying a Novel Population-Based Model Approach to Estimating Breath Alcohol Concentration (BrAC) from Transdermal Alcohol Concentration (TAC) Biosensor Data.” *Alcohol* 20:20. doi: https://dx.doi.org/10.1016/j.alcohol.2018.09.005.

Sobell, Linda C., Mark B. Sobell, Gloria I. Leo, and Anthony Cancilla. 1988. “Reliability of a Timeline Method: Assessing Normal Drinkers’ Reports of Recent Drinking and a Comparative Evaluation across Several Populations.” *British Journal of Addiction* 83(4):393–402.

Sobell, Mark B., Linda C. Sobell, Felix Klajner, Daniel Pavan, and Ellen Basian. 1986. “The Reliability of a Timeline Method for Assessing Normal Drinker College Students’ Recent Drinking History: Utility for Alcohol Research.” *Addictive Behaviors* 11(2):149–61.

Stockwell, Tim, Susan Donath, Mark Cooper-Stanbury, Tanya Chikritzhs, Paul Catalano, and Cid Mateo. 2004. “Under-Reporting of Alcohol Consumption in Household Surveys: A Comparison of Quantity-Frequency, Graduated-Frequency and Recent Recall.” *Addiction (Abingdon, England)* 99(8):1024–33. doi: 10.1111/J.1360-0443.2004.00815.X.

Swette, L., A. Griffith, and A. La Conti. 1997. “Potential and Diffusion Controlled Solid Electrolyte Sensor for Continuous Measurement of Very Low Levels of Transdermal Alcohol.” *Giner, Inc., Waltham, MA. Appl. No. 840802*.

Swift, Robert. 2003. “Direct Measurement of Alcohol and Its Metabolites.” *Addiction* 98(SUPPL. 2):73–80. doi: 10.1046/J.1359-6357.2003.00605.X.

Swift, Robert M., Christopher S. Martin, Larry Swette, Anthony LaConti, and Nancy Kackley. 1992.

"Studies on a Wearable, Electronic, Transdermal Alcohol Sensor." *Alcoholism: Clinical and Experimental Research* 16(4):721–25. doi: 10.1111/J.1530-0277.1992.TB00668.X.

Trull, Timothy J., and Ulrich Ebner-Priemer. 2014. "The Role of Ambulatory Assessment in Psychological Science." *Current Directions in Psychological Science* 23(6):466–70. doi: 10.1177/0963721414550706.

Vestal, Robert E., Elizabeth A. McGuire, Jordan D. Tobin, Reubin Andres, Arthur H. Norris, and Esteban Mezey. 1977. "Aging and Ethanol Metabolism." *Clinical Pharmacology & Therapeutics* 21(3):343–54. doi: 10.1002/CPT1977213343.

Voas, Robert B., Robert L. DuPont, Stephen K. Talpins, and Corinne L. Shea. 2011. "Towards a National Model for Managing Impaired Driving Offenders." *Addiction* 106(7):1221–27. doi: 10.1111/J.1360-0443.2010.03339.X.

Voogt, Carmen, Emmanuel Kuntsche, Marloes Kleinjan, Evelien Poelen, and Rutger Engels. 2014. "Using Ecological Momentary Assessment to Test the Effectiveness of a Web-Based Brief Alcohol Intervention Over Time Among Heavy-Drinking Students: Randomized Controlled Trial." *J Med Internet Res 2014;16(1):E5 Https://Www.Jmir.Org/2014/1/E5* 16(1):e2817. doi: 10.2196/JMIR.2817.

Wolfe, Jeannette, Ricardo Martinez, and Warren A. Scott. 1998. "Baseball and Beer: An Analysis of Alcohol Consumption Patterns among Male Spectators at Major-League Sporting Events." *Annals of Emergency Medicine* 31(5):629–32.

Wray, Tyler B., Jennifer E. Merrill, and Peter M. Monti. 2014. "Using Ecological Momentary Assessment (EMA) to Assess Situation-Level Predictors of Alcohol Use and Alcohol-Related Consequences." *Alcohol Research: Current Reviews* 36(1):19.

Wright, Cassandra, Paul M. Dietze, Paul A. Agius, Emmanuel Kuntsche, Michael Livingston, Oliver C. Black, Robin Room, Margaret Hellard, and Megan S. C. Lim. 2018. "Mobile Phone-Based Ecological Momentary Intervention to Reduce Young Adults' Alcohol Use in the Event: A Three-Armed Randomized Controlled Trial." *JMIR MHealth and UHealth* 6(7):e149.

Züll, C. 2016. "Open-Ended Questions." *GESIS Survey Guidelines* 3.

3. A review of the literature on the validity of transdermal alcohol monitors

Chapter 3 contains the following publication:

van Egmond, K., Wright, C. J., Livingston, M., & Kuntsche, E. (2020). Wearable transdermal alcohol monitors: a systematic review of detection validity, and relationship between transdermal and breath alcohol concentration and influencing factors. *Alcoholism: Clinical and Experimental Research*, *44*(10), 1918-1932.

Author contributions

KvE collected and analysed the literature and led the manuscript writing. EK, CW and ML provided critical feedback and suggested revisions.

Alcoholism: Clinical and Experimental Research

Vol. 44, No. 10
October 2020

Critical Review

Wearable Transdermal Alcohol Monitors: A Systematic Review of Detection Validity, and Relationship Between Transdermal and Breath Alcohol Concentration and Influencing Factors

Kelly van Egmond, Cassandra J. C. Wright, Michael Livingston, and Emmanuel Kuntsche

Background: Research on alcohol consumption mostly relies on self-reported data, which are subject to recall bias. Wearable transdermal alcohol concentration (TAC) monitors address this limitation by continuously measuring the ethanol excreted via the skin. This systematic review aims to provide an overview of TAC monitors' reliability to detect alcohol consumption and methods to estimate breath alcohol concentration (BrAC) and number of standard drinks consumed in a given time frame.

Methods: The databases MEDLINE, PsycINFO, SCOPUS, Engineering Village, and CINAHL were systematically searched to identify 1,048 empirical research papers published from 2013 onwards, of which 13 were included after full-text screening. The selected studies included 3 TAC monitors: SCRAM™, WristTAS™, and Skyn™.

Results: TAC measures of SCRAM, WrisTAS, and Skyn are found to be positively correlated with BrAC (r = 0.56 to 0.79) and/or self-reports (r = 0.62). Using the AMS criteria for detection results in low sensitivity, adjusted criteria can increase the sensitivity of the SCRAM from 39.9 to 68.5%. The WrisTAS and an early prototype of the Skyn showed high failure rates (17 to 38%). Recent advances toward transforming the TAC data into more clinically relevant measures have led to the development of mathematical models and the *BrAC Estimator Software*. Using TAC data, both approaches produce estimates explaining 70 to 82% of actual BrAC and self-reported drinking or to highly correlate with the actual BrAC measures (β = 0.90 to 0.91).

Conclusions: Transdermal alcohol monitors offer an opportunity to measure alcohol consumption in a valid and continuous way with mathematical models and software estimating BrAC values improving interpretation of TAC data. However, the SCRAM seems unable to detect low-to-moderate drinking levels using the thresholds and criteria set by the manufacturer. Moreover, the WrisTAS and the Skyn prototype show a high failure rate, raising questions about reliability. Future studies will assess the validity of new-generation wristbands, including the next Skyn generations.

Key Words: Transdermal Alcohol Concentration, Alcohol Consumption, Continuous Measurement.

CURRENTLY, RESEARCH ON alcohol consumption is predominately based on self-reported drinking. However, self-report measures are limited as they require active participation resulting in a high-response burden and they are subject to recall bias due to contextual factors, memory capacity, and motivation (Karns-Wright et al., 2018; Maylor et al., 1987; Shiffman, 2009). Studies which collect data during drinking events can reduce some of this bias (Dulin et al., 2017) but still require active participation; this is suspected to influence participants' natural drinking behavior (Fairbairn et al., 2019). Due to these limitations, researchers have sought a biometric solution, such as measuring alcohol in blood, breath, or urine. However, biomarkers such as breath or blood alcohol concentration (BrAC/BAC) are limited in their window of detection and again require repeated active participation during drinking events (Campbell et al., 2018; Fairbairn et al., 2019; Hill-Kapturczak et al., 2014; Karns-Wright et al., 2017; Swift, 2000).

From the Centre for Alcohol Policy and Research (CAPR), (KE, CJCW, ML, EK), La Trobe University, Melbourne, Vic., Australia; Burnet Institute, (CJCW), Melbourne, Vic., Australia; School of Public Health and Preventive Medicine, (CJCW), Monash University, Melbourne, Vic., Australia; School of Clinical Neuroscience, (ML), Karolinska Institute, Stockholm, Sweden; and Institute of Psychology, (EK), Eötvös Loránd University, Budapest, Hungary.

Received for publication February 6, 2020; accepted July 24, 2020.

Reprint requests: Kelly van Egmond, Centre of Alcohol Policy Research, School of Psychology and Public Health, La Trobe University, 3086 Bundoora, Vic., Australia; Tel.: +61 04 98257320; E-mail: k.vanegmond@latrobe.edu.au

DOI: 10.1111/acer.14432

Wearable monitors address these limitations by measuring the transdermal alcohol concentration (TAC) continuously, nonintrusively, and without action required by participants and thus without any response burden (Karns-Wright et al., 2017; Swift, 2000). These monitors have the potential to provide continuous, accurate measurement of alcohol consumption during a drinking event and over an extended period (e.g., several months). By plotting the TAC data over time, the rises and fall of the TAC curves can reflect the absorption and excretion or metabolism of alcohol and thus giving a more in-depth image of an individual's drinking episode as compared to the information either self-report or BrAC/BAC can provide. In this paper, we systematically review the existing literature on the use of transdermal alcohol monitors.

TAC Monitors

The majority of alcohol consumed is processed and excreted via the liver; however, approximately 1% is excreted transdermally, both passively through the skin and actively through sweat glands (Swift, 2000). Most TAC literature to date has examined the secure continuous remote alcohol monitor (SCRAM™), an ankle bracelet from Alcohol Monitoring Systems, Inc. (AMS, Littleton), and the Wrist Transdermal Alcohol Sensor (WrisTAS™; Giner, Inc., Newton). These monitors sample the transdermal ethanol (EtOH) evaporation above the skin and analyze this using a platinum-based electrochemical fuel cell to determine the EtOH content through catalytic alcohol oxidation.

Leffingwell and colleagues (2013) reviewed the available literature on the SCRAM and WrisTAS, which were the only available transdermal monitors at the time. Five studies published between the years of 1992 and 2009 validated the devices by comparing TAC data to self-reported drinking or BrAC. Sakai and colleagues (2006) tested the SCRAM in a laboratory setting and reported a correlation of 0.85 and 0.84 between TAC and BrAC measures for both peak and area under the curve (AUC) measurements, respectively. Additional self-report measures revealed that all self-reported drinking episodes were associated with a positive TAC event as measured with the SCRAM. The WrisTAS has been tested under similar laboratory and ambulatory settings. Peak TAC and AUC measures were highly correlated with the corresponding BrAC levels across individuals (0.61 and 0.91; Swift and Swette, 1992). Test–retest measures showed that peak and AUC measures from 2 WrisTAS devices worn simultaneously were highly correlated ($r = 0.71$, $r = 0.94$; Swift and Swette, 1992). When compared to self-reported drinking, the TAC data from the WrisTAS were highly correlated with self-reported drinks per drinking episode (AUC $r = 0.69$) and agreement between the self-reports and TAC was 83 to 96% (Swift et al., 2004). Overall, the studies included in the review of Leffingwell and colleagues (2013) reported TAC measurements to be highly correlated with BrAC and self-reported drinking.

The studies discussed in the Leffingwell review (Leffingwell et al., 2013) further report on the sensitivity and specificity of the SCRAM and WrisTAS. Sensitivity and specificity are measures of detection accuracy, referring to the ability of the TAC monitors to identify whether a given quantity of alcohol is consumed or not. Due to the SCRAM monitors' use in the justice system, the monitor has been developed to simply identify whether drinking is likely to have occurred with criteria aiming to reduce the risk of false positives. As researchers aim to measure the levels of alcohol consumption in more sophisticated ways, sensitivity and specificity have been a focus in studies testing the TAC monitors (Karns-Wright et al., 2017). Previous studies have used either self-report or BrAC as a reference measure to designate true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The positive or negative readings are identified by either the criteria as developed by AMS (manufacturers of the SCRAM) or adjusted criteria (Roache et al., 2019). Raw TAC data often show unnatural spikes that are likely caused by environmental noise such as alcohol in the carpet or spilling of drinks. When TAC data are not accurately processed or wrong thresholds and criteria are used, this can cause the detection of FPs. Sensitivity, also called the true-positive rate, measures the amount of positive drinking events as indicated by TAC that are truly positive drinking events as compared with self-report or BrAC: sensitivity % = $(TP/(TP + FN)) \times 100$. Specificity, also called the true-negative rate, indicates how many events where alcohol consumption is not detected are actually events without drinking (based on self-report or BrAC): specificity % = $(TN/(TN + FN)) \times 100$. When using AMS' conservative threshold and criteria needing at least 3 TAC data points (also called TAC readings, 3 readings are a total of 90 minutes) above 0.02 g/dl, specificity will remain high and the detection of FNs is prevented, but sensitivity of the TAC drinking detection is lost. The sensitivity of the WrisTAS was reported to reach 84.0% (Swift et al., 2004). Marques and McKnight (2009), as included in the Leffingwell and colleagues (2013) review, validated both the SCRAM and WrisTAS. The SCRAM was reported to show a specificity of 87.7% and sensitivity increased from 65.3 to 86.5% when alcohol levels increased (0.02 to 0.08 g/dl). The WrisTAS showed a specificity of 74.1% and a sensitivity of 84.4 to 82.3% when increasing the BAC levels from 0.02 to 0.08 g/dl. However, failure rate was observed to be much higher for the WrisTAS resulting in a low rate of performance of the WrisTAS, with only 23.6% of the drinking episodes being detected (Marques et al., 2009) and a failure rate of 49% (Swift et al., 2004). This review will discuss the results of studies testing different thresholds and developing new criteria identifying drinking episodes and the effects on sensitivity and specificity.

Leffingwell and colleagues (2013) conclude that the lag in alcohol detection and time to peak is one of the main limitations of the TAC monitors. As alcohol secretion through the skin requires a longer time as compared to breath, this results

in a lag of the TAC curves as compared to BrAC by 1 or more hours (Sakai et al., 2006; Swift, 2003) and peak TAC seems to be lower as compared to peak BrAC (Sakai et al., 2006). Breathalysers allow estimation of BAC based on BrAC, which is reasonably robust across individuals. Conversely, TAC data pose more of a challenge for analysis (Dai et al., 2016). EtOH transportation through the skin is physiologically more complex and appears to be more strongly influenced by individual differences (e.g., skin thickness, gender, body mass index [BMI]) and characteristics of the drinking event and its environment (e.g., drinking rate, food consumption, temperature; Dai et al., 2016; Leffingwell et al., 2013). Leffingwelland colleagues (2013) discussed the development of more sophisticated methods of modeling TAC data into BrAC data. The review discusses an unpublished study by one of the coauthors which studied the validity of a predictive model estimating BrAC values based on TAC data. The estimates of peak BrAC were significantly correlated with peak TAC measures ($r = 0.59$) and estimates of peak BrAC when using Widmark's equation based on self-reported drinking ($r = 0.57$). The Leffingwell review (Leffingwell et al., 2013) further discusses the *BrAC Estimator Software* tool, developed by 2 other coauthors of the review using Matrix Laboratory (MATLAB), a high-performance language for technical computing, producing BrAC estimates using the TAC data (Leffingwell et al., 2013). The software was described to calibrate TAC models to the participant and device by using parameters obtained from a laboratory session and apply these to field data to obtain semi-qualitative estimates of BrAC. In initial testing of the software, consistent models were created across devices, despite a difference in raw TAC data. It was able to compensate for the differences between TAC and BrAC in lag and attenuation. However, as the software was described to be in its initial testing phase no studies had been published at the time the Leffingwell review was published (Leffingwell et al., 2013). The review concluded that at the time, future use of the TAC monitors seemed to be promising, but that further research was needed to ascertain its validity and reliability. Since the publication of this review in 2013, the use of transdermal monitors in research has continued to grow and newer-generation wristbands have been released. It is thus important to systematically reassess the current state of knowledge on the validation of the monitors.

Newer-generation wristbands, such the BACtrack Skyn™ (Fairbairn and Kang, 2019) and the Milo ION™ (Lansdorp et al., 2019), have been developed and tested for validity and accuracy of alcohol detection. The Skyn was released for research purposes in 2019. The IONs have been made available on limited release for studies but are still under development. A notable advance of these wristbands is the wireless connection to smartphone devices; they are also substantially smaller and lighter than previous TAC monitors. Like the SCRAM, the Skyn uses fuel-cell technology to measure the EtOH above the skin. However, as the Skyn does not require a pump to generate the airflow across the sensor, this wristband can be much smaller (Fairbairn and Kang, 2019). The ION Milo sensor uses an enzyme-based sensor that is embedded in a disposable cartridge that needs to be replaced every 24 hours (Lansdorp et al., 2019). There is limited literature available on the Skyn and Milo as the wristbands have only been available to researchers for less than 12 months at the time of this review.

This systematic literature review aims to provide an overview of the recent literature studying the reliability and validity of TAC measurement technology focusing on: (i) TAC detection, (ii) recent advances toward the calculation of eBrAC and number of standard drinks from TAC data, and (iii) gaps in our knowledge and future directions for research to improve the robustness of TAC data collected. We will summarize current knowledge regarding the utility of the TAC monitors and their use in both clinical and research settings, particularly the recent advances toward utilizing the TAC data to obtain estimated levels of peak alcohol consumption (peak BrAC) and the number of standard drinks.

MATERIALS AND METHODS

Identification and Screening Procedure

A systematic search strategy was designed by the main author (KE) after an initial scoping of the available literature in consultation with university library staff who specialize in conducting systematic literature searches and meta-analyses. The search strategy was designed to retrieve articles using the following keywords: (i) "transdermal" or "sweat"; (ii) "alcohol"; (iii) "bracelet"; and (iv) "measure," and synonyms (see Table A1 in the Appendix 1, for details). An initial systematic search was performed in November 2018 in MEDLINE, PsycINFO, SCOPUS, CINAHL, and Engineering Village in accordance with the PRISMA guidelines (McInnes et al., 2018; Moher et al., 2015). The systematic review protocol was registered with PROSPERO database (#116215).

One researcher (KE) screened the records based on title and abstract to ensure that the articles met the inclusion criteria. Eligible records for data extraction included: (i) full-text original articles, published from 2013 onwards (the year the previous review was published [Leffingwell et al., 2013]); (ii) written in English; (iii) studies using wearable transdermal alcohol monitors or data to measure the alcohol consumption, with TAC levels as an outcome measure (studies focussing on a nonwearable transdermal measurement technique (i.e., tattoo sensors) were excluded); (iv) peer-reviewed; (v) studies conducted with adult participants who consumed alcohol; and (vi) studies using a reference measure (BrAC or self-reported drinking) to assess validity and/or effectiveness.

Following the initial search, a total of 1,048 records were retrieved of which 514 duplicates were removed (Fig. 1). Following abstract screening, further records were excluded due to being published before 2013 ($n = 301$), not being a full-text original research article (conference proceedings [$n = 18$]), or not focusing on measuring TAC ($n = 155$). Uncertainties were resolved by consensus with all other authors. Of the remaining 62 records, full-text copies of the articles were screened by 1 author (KE) who discussed any doubts with all other authors. Records that did not focus on the validation of TAC monitors were excluded ($n = 20$) as were records that did not focus on the validation of wearable technology ($n = 14$).

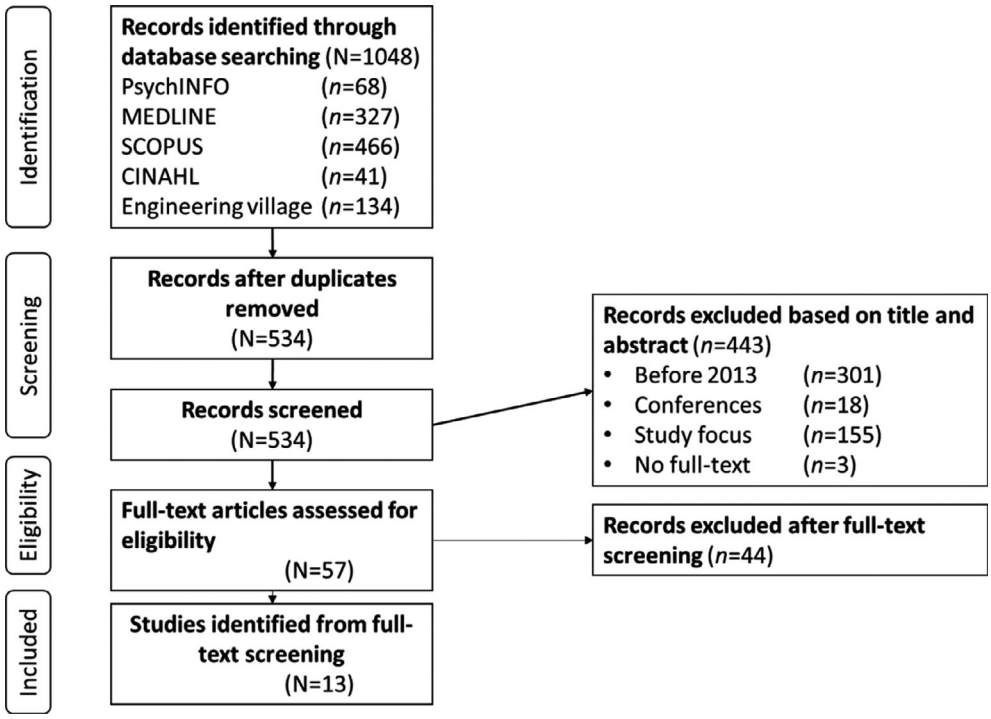

Fig. 1. Flow diagram of selected record screening.

An updated search was performed in November 2019 due to relevant new studies being published, including 3 new articles testing new-generation wristbands. This included the BACtrack Skyn™ (Fairbairn and Kang, 2019), the Milo ION™ (Lansdorp et al., 2019), and the Metal OXide (MOX) sensor (Lawson et al., 2019). The studies with the ION and MOX sensors did not validate TAC measures obtained from the monitors, and no correlation or detection rate measures were given. This resulted in the exclusion of these studies; the final sample of studies therefore focuses on the SCRAM, WrisTAS, and Skyn.

A modified version of the Critical Appraisals Skills Programme (CASP) checklist for diagnostic and randomized controlled studies (Zeng et al., 2015) was used by the main author (KE) to assess risk of bias within each included study. All records considered for data extraction were considered low risk of bias and of sufficient quality to be included in the review.

Data Extraction Procedure

The following information was extracted from the included studies:

1. Study design and methods
2. Outcome variable
3. Analytical approach used
4. Sensitivity and specificity
5. Correlations to control/reference variable
6. Significant predictor variables/correlates
7. Models and their reliability

The included studies are summarized using a structured narrative description based on the monitor used, study design, and analysis methods. A meta-analysis was not possible due to methodological heterogeneity.

RESULTS

Description of the Identified Studies

The final sample included 13 publications. Three of the studies tested the SCRAM in an ambulatory setting (i.e., naturalistic environments; comparing the TAC data to self-reported drinking; Table 1). Five studies tested the SCRAM in a laboratory (comparing the TAC data to number of drinks or BrAC; Table 2). One study used the SCRAM in a combined laboratory–ambulatory design using the *BrAC Estimator Software* (see below) to estimate the BrAC values from TAC data in which participants had to come in the laboratory for a calibration session to test the estimation in the field afterward (Table 3). Two studies tested the WrisTAS in an ambulatory setting (Table 4) and one in a combined laboratory–ambulatory design (Table 5). One study evaluated both the SCRAM and Skyn in a laboratory setting (Table 6). The number of participants in the studies ranged from $n = 1$ to 66; most studies included an equal gender balance of participants.

Studies Using SCRAM

Detection Accuracy of Alcohol Consumption Using SCRAM. Detection accuracy of alcohol consumption can be analyzed using detection rate, and measures of detection such as sensitivity and specificity when using a transdermal monitor. The positive or negative readings are identified by either the criteria as developed by AMS (manufacturers of

Table 1. SCRAM Ambulatory Studies

Publication	*N* participants	% Male	TAC monitor used	*N* days assessed	Reference measure	Outcome	Measurements (average)	Primary results			
								Correlation to reference measure	% Sensitivity	% Specificity	Significant predictors
Karns-Wright and colleagues (2018)	30	53	SCRAM	28	Daily self-report	Correlation measures using different detection thresholds	Self-reported drinking: 324 days. TAC detections: 245 days	Moderate + heavy TAC (M = 75.91%, SD = 15.06%) only heavy TAC (M = 73.69%, SD = 15.23%). 65.4% detected of total self-reported drinking	36 (AMS criteria), 65 (low threshold), 62 (moderate threshold), 51 (heavy threshold)	98 (AMS criteria), 91.7 (lower criteria)	Number of StDr
Roache and colleagues (2019)	30	53	SCRAM	28	Daily self-report	Correlation measures comparing 2 detection methods	Self-reported drinking days: 722. TAC detections: 606. Confirmed events: Research rules: 345 AMS: 163	High sensitivity, without decreasing specificity	39.9 (AMS criteria), 68.5 (own rules). Abstinence: 62.1 (AMS criteria)	99.8 (AMS criteria), 90.4 (own rules), and for abstinence 2% (AMS criteria)	NA
Barnett and colleagues (2014)	66	54.5	SCRAMII and SCRAMx	21 to 28	eBAC from daily self-report	Significant predictors of detection	Self-reported drinking episodes: 690. TAC detections: 502. Self-reported drinks: 6.3 eBAC: 0.083 g/dl TAC peak: 0.100 g/dl	B = 0.54, $p < 0.001$	72.8	NA	Gender, number of StDr, BMI, alcohol dependence scores, eBAC, and StDr

Abbreviations within the table are as follows: AMS, alcohol monitoring systems; AUC, area under the curve; BMI, body mass index; eBAC, estimated blood alcohol concentration; (e)StDr, (estimated) standard drinks; *M*, median; NA, not available; SCRAM, secure continuous remote alcohol monitor; SD, standard deviation; TAC, transdermal alcohol concentration.

[a]Some studies use the AMS criteria to compare their own detection criteria. AMS criteria include the following: (i) At least 3 TAC readings over 0.02 g/dl. (ii) Absorption rate for the event should be lower than 0.05 g/dl/h. (iii) When peak TAC is lower than 0.15 g/dl, elimination rate for the event should be lower than 0.025 g/dl/h. (iv) When peak TAC is above 0.15 g/dl, elimination rate should be less than 0.035 g/dl/h.

[b]TAC classifications: low: 3 or more TAC points >0, but no points >0.01 g/dl; moderate: ≥3 TAC points above 0 and ≥1 TAC point above 0.01 g/dl but <2 points above 0.02 g/dl; and heavy: 2 or more readings >0.02 g/dl.

Table 2. SCRAM Laboratory Studies

Publication	*N* participants	% Male	*N* sessions	Alcohol administration (StDr total, StDr in minutes)	Reference	Outcome	Measurement (average)	Primary results			
								Correlation to reference	% Sensitivity	Significant predictors	Model used
Hill-Kapturczak and colleagues (2014)	22	50	5	1.3 to 6.5 in own pace	BrAC and known number of StDr	eBrAC and eStDr	NA	$Peak_{TAC} - Peak_{BrAC}$: $F(1, 73) = 160.03$, $p < 0.001$, eBrAC and BrAC: $R^2 = 0.70$, eStDr and StDr: $R^2 = 0.79$	NA	Number of StDr, gender, drinking rate (only for eStDr)	eBrAC = 0.02158 + 0.3940*Peak TAC + 0.000149* Time-to-Peak TAC − 0.00366* Gender − 0.1887* Peak TAC*Gender eStDr = 0.6990 + 0.006317*Time-to-Peak TAC + 0.09735*AUC − 0.00097*AUC*AUC + 0.08492*AUC*Gender − 0.00223*AUC *AUC*Gender
Dougherty and colleagues (2015)	42	50	5	1.3 to 6.5, 1.3 in 10 every 24 to 30 minutes	Known number of StDr	eStDr	StDr: eStDr: 1.3 1.0 2.6 2.7 3.9 3.9 5.2 5.1 6.5 6.1	Spearman's $r = 0.92$, $p < 0.0001$, $R^2 = 0.80$ (using independent dataset)	NA	Number of StDr, gender, AUC, peak TAC, time-to-peak TAC, and eStDr	eStDr = 0.6990 + 0.006317 * time-to-peak TAC + 0.09735 * AUC − 0.00097 * AUC2 + 0.08492 * AUC * Gender − 0.00223 * AUC^2 * Gender
Roache and colleagues (2015)	61	52.5	5	1.3 to 6.5, 1.3 in: a) 10 every 24 (men) or 30 (women), b) 10 every 24, c) own pace	Known number of StDr	Frequencies of TAC exceeding a set threshold	Exceeding 0.00: 90.8% Exceeding 0.02: 61.8% Exceeding 0.03: 53.6% AMS[a] resolved: 56.9% AMS[a] confirmed: 53.2%	95 to 100% for 1.8 to 2.7 StDr, AMS[a] only 100% detection at 6.5 StDr	Per unit: (i) 62.5 (males), 58.6 (females), 0 (AMS[a]), (ii) 93.8 (males) 96.6 (females), AMS[a]: 25 (males), 58.6 (females), (iii) 100 (overall), AMS[a]: 50 (males), 62.1 (females), (iv) 100 (overall), 80 (AMS[a]), (v) 95 (AMS[a])	Gender, threshold used and amount of drinks	NA
Hill-Kapturczak and colleagues (2015)	21	52	5	1.3 to 6.5, 1.4 in 10 every 24	BrAC	eBrAC	Men: $Peak_{BrAC}$: 0.045 $Peak_{eBrAC}$: 0.044 Women: $Peak_{BrAC}$: 0.064 $Peak_{eBrAC}$: 0.062	$R^2 = 0.76$. Using independent dataset: Spearman's $r = 0.86$, $p < 0.0001$.	1.3 StDr: 61.9% 2.6 to 6.5 StDr: 100%	Amount of drinks, gender, weight, BMI, peak TAC, and time-to-peak TAC	eBrAC = 0.02158 ± 0.3940 * peak TAC ± 0.000149 * time-to-peak TAC − 0.00366 * Gender − 0.1887 * peak TAC * Gender.
Karns-Wright and colleagues (2017)	61	53	5	1.3 to 6.5, 1.4 in: (i) 10 every 24 (men) or 30 (women), (ii) 10 every 24, (iii) own pace	BrAC	Correlation in peak, and time-to-peak BrAC and TAC	NA	$Peak_{BrAC}$-$Peak_{TAC}$: 132 minutes Peak TAC-to-peak BrAC ratio: <1 at 1.3 StDr, approaching unity (i.e., ratio = 1) as StDr increased	62.5 (men), 58.6 (women) at 1.3 StDr	Number of StDr, gender, gender x StDr	NA

Abbreviations within the table are as follows: AMS, alcohol monitoring systems; AUC, area under the curve; BMI, body mass index; (e)BrAC, (estimated) breath alcohol concentration; (e)StDr, (estimated) number of standard drinks; NA, not available; SCRAM, secure continuous remote alcohol monitor; StDr, standard drinks; TAC, transdermal alcohol concentration.

[a]Some studies use the AMS criteria to compare to their own detection criteria. AMS criteria include the following: (i) At least 3 TAC readings over 0.02 g/dl. (ii) Absorption rate for the event should be lower than 0.05 g/dl/h. (iii) When peak TAC is lower than 0.15 g/dl, elimination rate for the event should be lower than 0.025 g/dl/h. (iv) When peak TAC is above 0.15 g/dl, elimination rate should be less than 0.035 g/dl/h.

[b]Hill-Kapturczak and colleagues (2015) tested their developed model in an independent dataset, and correlation measured are given in Spearmen's *r*.

Table 3. SCRAM Mixed Design

Publication	*N* participants	% Male	*N* days followed	Alcohol administration	Reference	Measurements (average)	Outcome	Primary results: Correlation to reference	Primary results: % Sensitivity	Primary results: % Specificity	Primary results: Model used
Fairbairn et al., 2019	48	50	7	Men: 0.82 g/kg, women: 0.74 g/	BrAC and self-report	Peak BrAC: 0.075 (SD = 0.01) Peak TAC: 0.065 (SD = 0.04) Self-report: 129 drinking events	eBrAC and correlations when using (i) *Calibrated Individual Estimates*, (ii) *Population-based Parameters*	(i) ($\beta = 0.90$, $t = 7.98$, $p < 0.0001$), (ii) ($\beta = 0.91$, $t = 7.34$, $p < 0.0001$).	50	90	*BrAC Estimator Software* using MATLAB

Note: Abbreviations within the table are as follows: AMS, alcohol monitoring systems; (e)BrAC, (estimated) breath alcohol concentration; NA, not available; SCRAM, secure continuous remote alcohol monitor; SD, standard deviation; StDr, standard drinks; TAC, transdermal alcohol concentration.

Table 4. WrisTAS Ambulatory Studies

Publication	*N* participants	% Male	*N* days assessed	Reference measure	Outcome	Measures (average)	Primary results: Correlation to reference measure	Primary results: % Sensitivity	Primary results: % Specificity	Primary results: Significant predictors
Simons and colleagues (2015)	12	NA	14	Three self-report methods: (i) in situ assessments (ii) morning reports, (iii) TLFB	Correlation measures	Percent drinking days: (i) 34.6 (ii) 34.6 (iii) 36.2 WrisTAS: 31.5 Self-reported drinks: b) 7.61 c) 8.04	(i) IRR = 1.022, $p < 0.001$; OR = 0.93, $p < 0.001$, (ii) IRR = 1.021, $p < 0.001$; OR = 0.80, $p < 0.001$, (iii) IRR = 1.026, $p < 0.001$; OR = 0.82, $p < 0.001$	72.4	92.9	Alcohol dependence scores
Bond and colleagues (2014)	30	69	28	Self-report	Correlation measurements using AUC	Self-reported drinking: 174 days. TAC detected: 149 days	$r = 0.62$ (unadjusted StDr), $r = 0.73$ (adjusted number of StDr)	85.6	67.5	Adjustment of number of StDr

Note: Abbreviations within the table are as follows: AUC, area under the curve; IRR, incident rate ratio; NA, not available; OR, odds ratio; StDr, standard drinks; TAC, transdermal alcohol concentration; TLFB, timeline follow-back.

Table 5. WrisTAS Mixed Design Studies

Publication	*N* Participants	% Male	*N* days followed	Alcohol administration	Reference	Measurements (average)	Outcome	Correlation to reference	Primary results: Significant predictors	Primary results: Model used
Luczak and Rosen (2014)	1	0	17	Reach 0.05%	BrAC	*Calibration phase:* Peak: 0.034 $Peak_{eBrAC}$: 0.035 Time: 79 $Time_{eBrAC}$: 86 AUC: 0.076 AUC_{eBrAC}: 0.078 *Estimation phase:* Peak: 0.033 $Peak_{eBrAC}$: 0.035 Time: 168.68 $Time_{eBrAC}$: 168.99 AUC: 0.074 AUC_{eBrAC}: 0.070	eBrAC and correlations with actual BrAC	Difference eBrAC and BrAC: *Calibration phase:* Peak: 0.004 Time to peak: 18 AUC: 0.010 *Estimation phase:* Peak: 0.009 Time to peak: 30 AUC: 0.019	Drinking behavior and shape of drinking behavior, episode used and monitor (calibration)	*BrAC Estimator Software* using MATLAB

Note: Measurements given are of one dataset (Dataset 2). Abbreviations within the table are as follows: (e)BrAC, (estimated) breath alcohol concentration; SD, standard deviation; TAC, transdermal alcohol concentration; time, time of peak in hours.

Table 6. BACtrack and SCRAM Laboratory Studies

Publication	*N* Participants	% Male	TAC monitor used	*N* sessions	Alcohol administration (StDr total, StDr in minutes)	Reference	Measurements (average)	Outcome	Primary results: Correlation to reference
Fairbairn and Kang (2019)	30	50	BACtrack Skyn and SCRAM	1	NA, up to 0.08% BrAC	BrAC	Average peak BrAC: 0.08 Latency to first detection (minutes): Skyn: 22 SCRAM: 23 Latency to peak (minutes): BrAC: 77 Skyn: 132 SCRAM: 197	Correlation to BrAC	Peak Skyn: $r = 0.77$, SCRAM: $r = 0.56$. AUC Skyn: $r = 0.79$, SCRAM: $r = 0.60$.

Note. Abbreviations within the table are as follows: AUC, area under the curve; BrAC, breath alcohol concentration; NA, not available; SCRAM, secure continuous remote alcohol monitor; StDr, standard drinks; TAC, transdermal alcohol concentration.

the SCRAM) or criteria as developed by researchers themselves (Roache et al., 2019).

Roache and colleagues (2015) investigated the sensitivity of 2 versions of the SCRAM (SCRAMII and SCRAMx) in detecting low-level drinking when using different thresholds: (i) TAC > 0.00 g/dl, (ii) TAC > 0.02 g/dl, and (iii) TAC > 0.03 g/dl (Table 7). Almost 40% of events where participants consumed 0.9 standard drinks (Australian standard drink, containing 10 g of alcohol per drink) did not exceed the TAC > 0.00 g/dl. They also found that most low-level drinking (<3 beers containing 4.5% alcohol; under the 2.8 standard drinks) went undetected when using a threshold of >0.02 g/dl. When increasing the threshold above 0.03 g/dl, even drinking events with a higher consumption (e.g., 4.6 standard drinks) went undetected. It was concluded that when aiming to detect low-level drinking, TAC thresholds below 0.02 g/dl are necessary. Karns-Wright and colleagues (2018) aimed to test the correlation between self-reported drinking and positive TAC events from the SCRAM using both AMS criteria and adjusted criteria to detect low-to-moderate drinking. Again, only a small percentage of events with low-to-moderate drinking were detected at the TAC threshold of >0.02 g/dl. When adding a threshold to confirm lower-to-moderate drinking levels (with at least 3 TAC readings above zero and 1 above 0.01 but less than 2 above 0.02 g/dl in a drinking episode), the concordance between self-reports and TAC increased without a significant loss in specificity (Table 7). Using a TAC threshold of >0.02 g/dl using both laboratory data and self-reported drinking, Roache and colleagues (2019) similarly confirmed the low sensitivity of the AMS criteria and developed and examined their own TAC detection criteria consisting of 9 rules in total. These criteria included the use of a lower threshold (removing any positive TAC reading proceeded and followed by at least 2 zero TAC readings). Following this threshold, they removed TAC readings that showed unrealistic spikes within a series of TAC readings. After removing individual TAC readings, rules were applied examining whole TAC events to decide whether to remove the TAC event or not. Using these detection criteria, including a less conservative threshold, resulted in an increase in sensitivity (68.5%) in detecting low level of drinking as compared to the sensitivity when using the AMS criteria (39.9%). There was only a slight decrease in specificity as compared to the AMS criteria (99.8 to 90.4%; Roache et al., 2019; Table 7).

Factors Influencing SCRAM TAC Drinking Detection. Barnett and colleagues (2014) investigated factors that might influence alcohol detection by transdermal monitors. At a univariate level, several factors were found to influence the alcohol detection, that is, amount of alcohol consumed, gender, alcohol dependency status of the drinker, and BMI (Barnett et al., 2014). However, when these factors were mutually adjusted, only the number of drinks consumed was significantly associated with the ability of SCRAM to detect self-reported drinking, with low levels of drinking (up to 2.7 standard drinks) less likely to be detected (Barnett et al., 2014).

Correlation of SCRAM TAC to Self-Reported Drinking. The SCRAM was found to detect 72.8% of 690 self-reported drinking episodes (Barnett et al., 2014) and 65.4% of 324 self-reported drinking days (Karns-Wright et al., 2018; Table 1); the device showed a low failure rate with malfunctions (faulty electronics) reported 5.1% of drinking episodes. The AUC has been observed to be highly correlated (r = 0.79–0.94) with the overall number of drinks consumed during a drinking episode (Barnett et al., 2011; Leffingwell et al., 2013) and is reported to increase as the number of standard drinks increases (Dougherty et al., 2015; Hill-Kapturczak et al., 2014).

Correlation of SCRAM TAC With BrAC. Fairbairn and Kang (2019) reported a significant positive correlation between peak TAC and peak BrAC (r = 0.56, n = 30, $p < 0.001$). In addition, there was a significant association between the AUC of BrAC and AUC of TAC (r = 0.60). Fairbairn and Kang (2019) further examined the cross-correlations, that is, the correlation between 2 measures across time points. The average maximal cross-correlation between BrAC and TAC was r = 0.51 (SD = 0.12).

BrAC has been observed to detect alcohol almost immediately after consumption, and SCRAM TAC has shown a small latency of 23 minutes to first detection (Fairbairn and Kang, 2019). BrAC was shown to reach peak levels after about 77 minutes on average, whereas TAC lagged behind, ranging from 69 minutes (Fairbairn and Kang, 2019) to 4.5 hours, with an average lag of 1.5 hours (Karns-Wright et al., 2017; Leffingwell et al., 2013). Lag times are reported to increase with dose consumed (Karns-Wright et al., 2017). The SCRAM was found to again display low failure rates of 2% (Fairbairn and Kang, 2019).

Estimation of BrAC/BAC and Standard Drinks Consumed From Transdermal Data. Of the 13 studies included, 5 reported models to estimate BAC (eBAC; estimated BAC) or BrAC (eBrAC; estimated BrAC) levels from TAC data, which can be used to infer level of consumption more accurately as compared to raw TAC data or the dichotomous measure of alcohol detection.

eBrAC and eBAC values in these articles were determined using 3 methods: (i) using the Widmark formula: eBAC values can be estimated using the self-reported number of drinks with a model including gender, body weight, and time spent drinking (Barnett et al., 2014); (ii) mathematical modeling: eBrAC values are calculated by model derivation using TAC data considering multiple factors of influence, gender and gender-related variables (weight and height; Dougherty et al., 2015; Hill-Kapturczak et al., 2014, 2015); and (iii) using a version of the *BrAC Estimator Software*: eBrAC values can also be calculated with the use of an automated

Table 7. Sensitivity and Specificity Measures When Using Different Thresholds

Reference	Sample size	Setting	Threshold used	Sensitivity (%)	Specificity (%)
Karns-Wright et al., 2018	30	Ambulatory	2 points >0.02 g/dl	51	96
			≥3 points above 0, ≥1 above 0.01, and ≤2 points above 0.02 g/dl	62	94
Roache and colleagues (2019)	30	Ambulatory	≥3 points above >0.02 g/dl (AMS)	39.9	99.8
			"any drinking"	68.5	90.4
Roache and colleagues (2015)	61	Laboratory	≥3 points above >0.02 g/dl (AMS)	53.2	NA
			Any exceeding 0.03 g/dl	53.6	NA
			Any exceeding 0.02 g/dl	61.8	NA
			Any exceeding 0 g/dl	90.8	NA

software program using calibration estimates and TAC data (Fairbairn et al., 2019; Luczak and Rosen, 2014).

The Widmark Formula—Using the Widmark formula, Barnett and colleagues (2014) estimated eBAC levels based on self-reported data. These eBAC levels were then compared to the raw TAC data from SCRAM. They found that women tended to have slightly higher eBAC as compared to men which in turn resulted in higher TAC levels. They reported an average peak eBAC of 0.08 g/dl and TAC of 0.10 g/dl which were significantly correlated ($r = 0.54$, $p < 0.001$).

Mathematical Modeling—Three laboratory studies explored the estimation of peak BrAC levels with TAC data using mathematical modeling (Dougherty et al., 2015; Hill-Kapturczak et al., 2014, 2015). Hill-Kapturczak and colleagues (2015) used mixed-effects models using multiple variables (including gender, weight, and height) to optimize the best fit and to improve eBrAC calculation from TAC data. They concluded that the optimal model included peak TAC, time-to-peak TAC (minutes from the last 0.00 g/dl TAC-to-peak TAC), and gender.

Using the model, the average peak eBrAC was 0.044 for men and 0.062 for women, explaining 76% of the variance in the actual peak BrAC being 0.045 for men and 0.064 for women on average. The model was further validated using the data collected in a previous study (Dougherty et al., 2012) where the eBrAC values were again highly correlated with the actual BrAC values (Spearman's $r_s = 0.86$). The laboratory study by Hill-Kapturczak and colleagues (2014) used the same model; however, participants were instructed to drink at their own pace to test the impact of drinking rate to the ability of the model to calculate peak eBrAC using TAC data. The peak eBrAC, as calculated using these new data, and the previously developed model, explained 70% of the variance in actual peak BrAC. Adding the drinking rate did not improve the goodness of fit of the model, emphasizing the validity of the model across different drinking rates.

Dougherty and colleagues (2015) also estimated the number of standard drinks consumed across a drinking occasion using the TAC data. Using data from 46 participants who consumed 0.9 to 6.5 standard drinks on 5 separate days in the laboratory, they were able to derive a mathematical model to estimate standard drinks from TAC data. Instead of peak TAC, the AUC was used as it was most strongly associated with the number of standard drinks consumed. They added a quadratic AUC variable (AUC^2) and its interaction with gender to improve model fit. The model explained 82% of the variance of actual standard drinks consumed (all $p < 0.03$). This model was again further validated using 2 independent datasets from independent studies (Dougherty et al., 2012; Hill-Kapturczak et al., 2014) and explained 79 to 80% of the variance of the actual standard drinks consumed.

Taken together, both the estimations of BrAC and standard drinks consumed using TAC data provided consistent results under different drinking conditions: (i) when men and women drank the same amounts at similar drinking rates, (ii) when men and women drank the same amounts, but with lower drinking rates for men, and (iii) under significantly varying individual rates of consumption.

BrAC Estimation Software—One study combined a laboratory and ambulatory design using the *BrAC Estimator Software* to produce eBrAC levels using TAC data. The software calibrated the TAC data obtained to the specific individual and the transdermal monitor used by deconvolving (filtering) the TAC data using 4 parameters combined into one impulse response function: (i) rate at which alcohol diffuses through the various layers of the skin, (ii) effective net rate at which alcohol enters and leaves the skin and is processed by the transdermal sensor, and (iii) and (iv) penalty, or regularization, parameters. These parameters vary across individuals based on several factors (e.g., transdermal monitor used) and were obtained during a calibration session (a laboratory alcohol administration session, during which individuals wear a transdermal monitor to measure TAC while taking breath samples at the same time).

Fairbairn and colleagues (2019) recently investigated the accuracy and validity of the *BrAC Estimator Software* using the SCRAM monitors. A calibration session was used to obtain *Individual Calibration Estimates* as parameters in the model. However, to bypass the individual calibration

sessions, they also used *Population Parameter Estimates* as developed in a previous study (Barnett et al., 2015) to translate TAC to BrAC for individuals in a new sample. During the collection of *Individual Calibration Estimates*, 24 men received 0.82 g/kg of body weight and 24 women 0.74 g/kg. Subsequently, the participants were then asked to wear a SCRAM monitor for 7 days in which they had to self-report their drinking. Using the resulting TAC data, the researchers utilized both the *Individual Calibration Estimates* and the *Population Parameter Estimates* to estimate eBrAC values and compare them with self-reported drinking. With a 129 self-reported drinking episodes, results indicated a strong correlation between the self-reported drinking quantity and area under the eBrAC curve using both the *Individual Calibration Estimates* ($\beta = 0.90$, $t = 7.98$ and $p < 0.0001$) and *Population Parameter Estimates* ($\beta = 0.91$, $t = 7.34$ and $p < 0.0001$).

Studies Validating WrisTAS

Correlation of TAC to Self-Reported Drinking. The WrisTAS showed a relatively high failure rate; Bond and colleagues (2014) observed that of all days for which a TAC curve was available, only 77% of these days contained a TAC curve of high enough quality, as indicated by an alcohol signal that rose and fell in a way that is plausible physiologically. On the days with sufficient data, the correlation between the AUC of TAC and self-report was $r = 0.62$. Simons and colleagues (2015) also found that the WrisTAS had a high failure rate with 17.6% of the days missing due to faulty electronics and 9.9% of the days missing due to the wristband being removed by participants. Of the remaining days, the WrisTAS was able to detect 85.7% of the self-reported drinking days.

Estimation of BrAC/BAC From Transdermal Data

BrAC Estimation Software. Luczak and Rosen (2014) conducted a study with one participant to test the *BrAC Estimator Software*. The *BrAC Estimator Software* fitted the model from 2 separate WrisTAS monitors worn simultaneously during a calibration session obtaining *Individual Calibration Estimates* as parameters in the model. As the raw TAC data differed per monitor, there was a need to calibrate the model not just to the individual, but also to the monitors. Using the calibration estimates, the software was able to produce eBrAC from TAC data that matched well with the actual BrAC data; on average, peak of eBrAC differed only with 0.004 from the actual peak BrAC for both monitors outside the laboratory. Thereby, the model was able to improve the time lag as compared to the raw TAC data with an average of 30 minutes. The software was observed to perform less when alcohol was consumed more sporadically over longer periods of time with 2 episodes resulting in a peak difference of 0.018 and a difference of 0.021 for the AUC between eBrAC and actual BrAC. Luczak and Rosen (2014) also reported a variability in estimated BrAC AUC measures across the 2 datasets, with an average difference between eBrAC and actual BrAC of 0.024 for Dataset 1 and 0.014 for Dataset 2.

Studies Validating BACtrack Skyn

Correlation With BrAC. Fairbairn and Kang (2019) recently studied the correlation of BrAC with TAC as measured by both an early prototype of the Skyn and the SCRAM. They reported a high correlation ($r = 0.77$, $n = 30$, $p < 0.001$) between peak BrAC and peak TAC (Skyn) and a significant association between the AUC BrAC and AUC TAC (Skyn; $r = 0.79$). The average maximal cross-correlation between BrAC and TAC (Skyn) was $r = 0.60$ (SD = 0.15) and for TAC (SCRAM) $r = 0.51$ (SD = 0.12), a difference, which was reported to be statistically significant (Mdiff = 0.09 (SD = 0.20), $t(24) = 2.38$, $p = 0.026$).

On average, TAC (Skyn) lagged behind BrAC by 24 minutes, whereas the same study showed that TAC (SCRAM) lagged behind BrAC by 69 minutes; this difference was again significant (Fairbairn and Kang, 2019). The failure rate of the Skyn prototype (including unusable data, losing data files, and battery failure) was reported to be 18 to 38% (Fairbairn and Kang, 2019).

DISCUSSION AND FUTURE DIRECTIONS

This systematic review aimed to provide an overview of the current evidence on the reliability and validity of wearable TAC measurement technology with a focus on both detection rate and the use of TAC data to produce BrAC estimates. We identified 13 studies that investigated the validity of the SCRAM, WrisTAS, and Skyn TAC monitors using an ambulatory, laboratory, or a combined laboratory–ambulatory design that have been published since a previous review by Leffingwell and colleagues (2013).

Processing of TAC Data, Including Criteria and Thresholds Used to Identify Drinking Episodes

A key benefit of TAC monitors is their ability to provide more rich information on alcohol consumption than other alternatives, and as such, alcohol researchers are looking to make use of this detailed data. Using conservative criteria for detecting drinking (such as the AMS criteria) during the processing of TAC data may limit the richness of the information due to exclusion of low-level drinking events. These criteria are necessarily conservative due to their use in the justice system and the legal implications associated with false positives. Further, the TAC data often show unnatural spikes that are likely caused by environmental noise, which require careful and considerate processing to ensure that we are reviewing alcohol consumption and not environmental noise (Roache et al., 2019). This is especially true of the SCRAM which only samples every 30 minutes, making it

difficult to conclude whether a spike is environmental noise, a result of sudden temperature or humidity change, or movement for example. Roache and colleagues (2015) studied the effect of different thresholds (see Table 7). Roache and colleagues (2019) went on to develop new criteria consisting of 9 rules increasing sensitivity to 68.5% and detecting lower levels of alcohol as compared to using the AMS criteria resulting in a sensitivity of 39.9%. These criteria may provide researchers with a more appropriate method of identifying alcohol consumption when using TAC data for research and clinical purposes. Ensuring that these criteria are consistent across studies will increase comparability of future studies, as measures like the AUC will differ depending on these criteria.

Further laboratory studies are also needed to better understand how various factors influence detection of alcohol and the TAC curves and to better measure alcohol consumption and BrAC. Future studies could replicate the study by Barnett and colleagues (2014) and look at BMI, gender, number of drinks, alcohol dependence, and monitor characteristics, but could also include factors that have not been studied as of yet such as temperature and humidity of the laboratory, skin thickness, and food consumption.

Correlation of TAC and Self-Reported Drinking

Both the SCRAM and WrisTAS have been tested on accuracy in naturalistic settings and compared against self-reports. Leffingwell and colleagues (2013) discussed the results of a study comparing the SCRAM with the WrisTAS and reported a higher sensitivity for the WrisTAS; however, the failure rate of the WrisTAS was found to be much higher as compared to the SCRAM. The studies included in our review do not directly compare the performance of WrisTAS and SCRAM. However, the failure rate of the WrisTAS has again been reported to be high (Bond et al., 2014; Simons et al., 2015), which results in the failure to detect drinking episodes/d and can significantly reduce the sensitivity. It is also important to note that the studies using the WrisTAS monitors used drinking days instead of drinking events, which is problematic because TAC readings from the previous day could carry over to the next day resulting in 2 drinking days instead of just 1 event. This possibly increases the correlation between TAC and self-reported drinking and thus biases detection rate and sensitivity measures. Finally, studies using the WrisTAS do not use any thresholds or other criteria to identify or exclude drinking episodes and has not been investigated in respect to thresholds for the detection of different levels of alcohol consumption (low-to-heavy drinking). As discussed, different criteria can significantly change the sensitivity (Roache et al., 2019).

The ambulatory studies have several limitations. One is that self-report is used as the reference measure to compare TAC against. As mentioned previously, self-report is limited as they are subject to bias due to contextual factors, memory capacity, and motivation (Karns-Wright et al., 2018; Maylor et al., 1987; Shiffman, 2009). Therefore, we cannot be certain whether a TAC drinking episode, without a matched self-reported drinking episode, should be considered a false positive or whether the episode is a false self-report (alcohol was consumed but not reported). Second, the TAC monitors have only been tested in a population drinking on average up to 12.4 standard drinks a week. How the monitors perform in more heavily drinking populations is still unknown. Finally, all studies reported in this review were conducted in North America on healthy participants. How these results apply to other populations and clinical samples has yet to be investigated.

Correlation With BrAC

TAC monitors have consistently been found to have a longer delay in alcohol detection compared with BrAC and BAC measures (Leffingwell et al., 2013). This can be explained by the metabolism through the skin which differs from breath and blood metabolism (Dai et al., 2016). In a recent study (Fairbairn and Kang, 2019), the Skyn TAC data showed a significant decrease in the lag to reach peak values with almost an hour less as compared to SCRAM TAC data. Unfortunately, the one available study of an early prototype of the Skyn reported a failure rate of 18 to 38% (Fairbairn and Kang, 2019). Fairbairn and Kang (2019) demonstrated some promising improvements of the Skyn over previous TAC monitors such as the SCRAM and WrisTAS. However, this study used prototypes of the Skyn so further studies are needed to comprehensively test performance of future generations of the Skyn.

In addition to the Skyn monitor, other new-generation wristbands are being developed including ION by Milo Inc. Unfortunately, only one study (Lansdorp et al., 2019) has been published to date and this study did not validate TAC measures as obtained from the monitors and no correlation or detection rate measures were given. Validation studies testing the accuracy of the ION, comparing them against SCRAM and Skyn are needed to determine its potential use for research and/or clinical purposes.

The laboratory studies have similar limitations to the ambulatory studies; participants were exclusively healthy individuals located in the United States, making it difficult to generalize. Although one study reported that the TAC-to-BrAC relationship improves when the alcohol consumption increases (Karns-Wright et al., 2017), the maximum number of drinks administered was only 6.5 standard drinks in a time frame of 47 to 166 minutes and it is unknown whether and how the TAC-to-BrAC relationship will change with higher levels of BrAC consumed in similar or smaller time frames. When aiming to study higher levels of alcohol consumption, researchers face ethical considerations and this may only be possible when observing participants in their natural environment (as opposed to administering alcohol in a laboratory).

BrAC and Standard Drinks Estimation Using TAC Data

Prior to the Leffingwell review (Leffingwell et al., 2013), no published studies had explored the options to estimate BrAC values using TAC data. Unpublished initial testing reported in Leffingwell and colleagues (2013) showed promising results with significant correlations between eBrAC, TAC, and self-report when using a mathematical model and consistent models across TAC devices as created by the software. The current review discusses the results of 5 published studies that tested the validity of mathematical models and the *BrAC Estimator Software*

The mathematical models have demonstrated higher correlations ($R^2 = 0.70$ to 76), consistent across datasets ($r = 0.86$; Dougherty et al., 2015; Hill-Kapturczak et al., 2014, 2015). These models also show promise in estimating the number of standard drinks using raw TAC data. Unfortunately, a limitation of the models is that they can only estimate the peak BrAC levels and do not provide sufficient detail about the rest of the drinking episode. Furthermore, the studies testing the models have all been conducted in similar controlled laboratory settings; the predictive validity of the models will need to be determined in natural environments. Finally, although wide variations in the drinking rate were reported, the controlled environment and the timing of alcohol administration being in the morning may have impacted the drinking rate. Longer or shorter time spans with higher amounts of alcohol usually occur in natural environments.

The *BrAC Estimator Software* has been refined since the Leffingwell review (Leffingwell et al., 2013). Studies using the software showed significant associations with self-reported drinking when using the SCRAM (Fairbairn et al., 2019) and produced consistent measures of BrAC across 2 WrisTAS devices (Luczak and Rosen, 2014). The software is programmed to not just estimate the peak BrAC but the whole curve including peak and AUC measures; it thereby improves the time-to-peak latency as compared to BrAC by over 30 minutes as compared to an average of an hour for the raw TAC data when using the WrisTAS monitors (Luczak and Rosen, 2014). Without the need of individual calibration sessions, the population-based approach might be a promising and less burdensome method to produce eBrAC values from TAC data (Fairbairn et al., 2019). However, the software is still in development and not widely available to use for research yet. There is limited literature available testing the software, with 1 of the 2 studies using one of the authors as their only participant. Also, the predictive value of the software decreased when drinking occurred more sporadically. Finally, the *BrAC Estimator Software* estimated BrAC AUC measures that were significantly different for the 2 WrisTAS monitors worn simultaneously and were less correlated with the actual BrAC measures.

Luczak and Rosen (2014) found that when conducting a test–retest of the WrisTAS monitors (by having 1 participant wear 2 similar WrisTAS devices simultaneously), the raw TAC data were different for each device. However, the *BrAC Estimator Software* created consistent models for both datasets yielding similar eBrAC values. This is the first time a study has tested 2 of the same monitors in parallel, and this will be crucial to replicate in future studies; it is important to know whether this is a consistent and predictable phenomenon across SCRAM, Skyn, or other recently available TAC wristbands and whether it has implications for using TAC monitors. The test–retest procedure and testing the TAC monitors against each other seem to be an important step to provide better evidence for validation of the monitors in the future.

CONCLUSION

This review found that TAC data as measured by the SCRAM, WrisTAS, and Skyn show a high correlation with both BrAC and self-reported drinks. However, limitations have been reported in identifying lower-to-moderate level drinking episodes, primarily related to the SCRAM and likely due to the conservative criteria and thresholds recommended by the manufacturer due to its use in the justice system. No evidence has been reported in this respect for the WrisTAS and Skyn to date. Both the WristTAS and a prototype of the Skyn have been shown to have a comparatively high failure rate, which substantially decreases the pool of quality data. This raises doubts about whether researchers and clinicians can currently exclusively rely on TAC monitors without further development and validation work. Further improvements to the devices are required to enhance the utility of TAC measurements.

Since the previous review (Leffingwell et al., 2013), there has been a considerable improvement observed in the development of mathematical models and software estimating BrAC values using raw TAC data matching well with the actual BrAC measures and self-reported drinking. Also, newer-generation wristbands have been developed and show promising results in the first study comparing an early prototype of the Skyn to the SCRAM. With only limited research available and continuous development of future generations of the Skyn, additional studies will assess its validity. Further testing is required to demonstrate the reliability, sensitivity, and specificity of the Skyn and other emerging TAC monitors.

CONFLICTS OF INTEREST

All authors declare no conflict of interest.

REFERENCES

Barnett NP, Meade EB, Glynn TR (2014) Predictors of detection of alcohol use episodes using a transdermal alcohol sensor. Exp Clin Psychopharmacol 22:86–96.

Barnett N, Souza T, Glynn T, Luczak S, Swift R, Rosen I (2015) The Transdermal Alcohol Sensor Macro (TASMAC): a rapid data processing tool for use with the SCRAMX alcohol sensor. Alcohol Clin Exp Res 39:192A.

Barnett NP, Tidey J, Murphy JG, Swift R, Colby SM (2011) Contingency management for alcohol use reduction: a pilot study using a transdermal alcohol sensor. Drug Alcohol Depend 118:391–399.

Bond JC, Greenfield TK, Patterson D, Kerr WC (2014) Adjustments for drink size and ethanol content: new results from a self-report diary and transdermal sensor validation study. Alcohol Clin Exp Res 38:3060–3067.

Campbell AS, Kim J, Wang J (2018) Wearable electrochemical alcohol biosensors. Curr Opin Electrochem 10:126–135.

Dai Z, Rosen IG, Wang C, Barnett N, Luczak SE (2016) Using drinking data and pharmacokinetic modeling to calibrate transport model and blind deconvolution based data analysis software for transdermal alcohol biosensors. Math Biosci Eng 13:911–934.

Dougherty DM, Charles NE, Acheson A, John S, Furr RM, Hill-Kapturczak N (2012) Comparing the detection of transdermal and breath alcohol concentrations during periods of alcohol consumption ranging from moderate drinking to binge drinking. Exp Clin Psychopharmacol 20:373–381.

Dougherty DM, Hill-Kapturczak N, Yuanyuan L, Karns TE, Lake SL, Cates SE, Roache JD (2015) The potential clinical utility of transdermal alcohol monitoring data to estimate the number of alcoholic drinks consumed. Addict Disord Their Treat 14:124–130.

Dulin PL, Alvarado CE, Fitterling JM, Gonzalez VM (2017) Comparisons of alcohol consumption by timeline follow back vs. smartphone-based daily interviews. Addict Res Theory 25:195–200.

Fairbairn CE, Kang D (2019) Temporal dynamics of transdermal alcohol concentration measured via new-generation wrist-worn biosensor. Alcohol Clin Exp Res 43:2060–2069.

Fairbairn CE, Rosen IG, Luczak SE, Venerable WJ (2019) Estimating the quantity and time course of alcohol consumption from transdermal alcohol sensor data: a combined laboratory-ambulatory study. Alcohol 81:111–116.

Hill-Kapturczak N, Lake SL, Roache JD, Cates SE, Liang Y, Dougherty DM (2014) Do variable rates of alcohol drinking alter the ability to use transdermal alcohol monitors to estimate peak breath alcohol and total number of drinks? Alcohol Clin Exp Res 38:2517–2522.

Hill-Kapturczak N, Roache JD, Liang Y, Karns TE, Cates SE, Dougherty DM (2015) Accounting for sex-related differences in the estimation of breath alcohol concentrations using transdermal alcohol monitoring. Psychopharmacology 232:115–123.

Karns-Wright TE, Dougherty DM, Hill-Kapturczak N, Mathias CW, Roache JD (2018) The correspondence between transdermal alcohol monitoring and daily self-reported alcohol consumption. Addict Behav 85:147–152.

Karns-Wright TE, Roache JD, Hill-Kapturczak N, Liang Y, Mullen J, Dougherty DM (2017) Time delays in transdermal alcohol concentrations relative to breath alcohol concentrations. Alcohol Alcohol 52:35–41.

Lansdorp B, Ramsay W, Hamidand R, Strenk E (2019) Wearable enzymatic alcohol biosensor. Sensors 19(10):2380.

Lawson B, Aguir K, Fiorido T, Martini-Laithier V, Bouchakour R, Burtey S, Reynard-Carette C, Bendahan M (2019) Skin alcohol perspiration measurements using MOX sensors. Sens Actuators B Chem 280:306–312.

Leffingwell TR, Cooney NJ, Murphy JG, Luczak S, Rosen G, Dougherty DM, Barnett NP (2013) Continuous objective monitoring of alcohol use: twenty-first century measurement using transdermal sensors. Alcohol Clin Exp Res 37:16–22.

Luczak SE, Rosen IG (2014) Estimating BrAC from transdermal alcohol concentration data using the BrAC estimator software program. Alcohol Clin Exp Res 38:2243–2252.

Marques PR, McKnight AS (2009) Field and laboratory alcohol detection with 2 types of transdermal devices. Alcohol Clin Exp Res 33:703–711.

Maylor EA, Rabbit P, Kingstone A (1987) Effects of alcohol on word categorization and recognition memory. Br J Psychol 78:233–239.

McInnes MD, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Clifford T, Cohen JF, Deeks JJ, Gatsonis C, Hooft L (2018) Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. JAMA 319:388–396.

Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA (2015) Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 4:1.

Roache JD, Karns TE, Hill-Kapturczak N, Mullen J, Liang Y, Lamb RJ, Dougherty DM (2015) Using transdermal alcohol monitoring to detect low-level drinking. Alcohol Clin Exp Res 39:1120–1127.

Roache JD, Karns-Wright TE, Goros M, Hill-Kapturczak N, Mathias CW, Dougherty DMJA (2019) Processing transdermal alcohol concentration (TAC) data to detect low-level drinking. Alcohol 81:101–110.

Sakai JT, Mikulich-Gilbertson SK, Long RJ, Crowley TJ (2006) Validity of transdermal alcohol monitoring: fixed and self-regulated dosing. Alcohol Clin Exp Res 30:26–33.

Shiffman S (2009) Ecological momentary assessment (EMA) in studies of substance use. Psychol Assess 21:486–497.

Simons JS, Wills TA, Emery NN, Marks RM (2015) Quantifying alcohol consumption: self-report, transdermal assessment, and prediction of dependence symptoms. Addict Behav 50:205–212.

Swift R (2000) Transdermal alcohol measurement for estimation of blood alcohol concentration. Alcohol Clin Exp Res 24:422–423.

Swift RJA (2003) Direct measurement of alcohol and its metabolites. Addiction 98:73–80.

Swift RM, Swette L (1992) Assessment of ethanol consumption with a wearable, electronic ethanol sensor/recorder, in *Measuring Alcohol Consumption* (Litten RZ, Allen JP eds), pp. 189–202. Totowa, NJ: Humana Press.

Swift R, Tempelman L, Brown GR (2004) A wearable transdermal alcohol sensor for measurement of alcohol drinking. Poster presented at the ISBRA 12th World Congress on Biomedical Alcohol Research, Heidelberg/Mannheim, Germany, September 29–October 2. Alcoholism: Clinical & Experimental Research 28(8).

Zeng X, Zhang Y, Kwong JS, Zhang C, Li S, Sun F, Niu Y, Du L (2015) The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: a systematic review. J Evid Based Med 8:2–10.

APPENDIX

Table A1. Search Strategy per Database

Database	Date searched	Exposure	Technique	No. of hits
PsycINFO	04/11/2019	((transdermal or sweat) and (alcohol or EtOH)).ti,ab. OR ((alcohol* or EtOH*) adj3 (bracelet* or wristband* or ankle*)).ti,ab.	(measur* or monitor* or sensor* or assess*).ti,ab.	68
MEDLINE	04/11/2019	((transdermal or sweat) and (alcohol or EtOH)).ti,ab. OR ((alcohol* or EtOH*) adj3 (bracelet* or wristband* or ankle*)).ti,ab.	(measur* or monitor* or sensor* or assess*).ti,ab.	327
SCOPUS	04/11/2019	TITLE-ABS((transdermal or sweat) and (alcohol or EtOH)) OR	TITLE-ABS (measur* or monitor* or sensor* or assess*)	466
CINAHL	04/11/2019	TI((transdermal or sweat) and (EtOH or alcohol)) OR AB ((transdermal or sweat) and (EtOH or alcohol))	TI(measur* or monitor* or sensor* or assess*) OR AB(measur* or monitor* or sensor* or assess*)	41
Engineering Village	04/11/2019	(((transdermal or sweat) and (EtOH or alcohol)) WN TI) OR (((transdermal or sweat) and (EtOH or alcohol)) WN AB)	((measur* or monitor* or sensor* or assess*) WN TI) OR ((measur* or monitor* or sensor* or assess*) WN AB)	134

4. A parallel test of SCRAM-CAM transdermal monitors to assess reliability

Chapter 4 contains the following publication:

van Egmond, K., Wright, C. J., Livingston, M., & Kuntsche, E. (2021). A parallel test of the SCRAM-CAM transdermal monitors ensuring reliability. *Drug and Alcohol Review, 40*(7), 1122-1130.

Author contributions

KvE and CW collected the data; KvE analysed the data and led data interpretation and manuscript writing. EK, CW, and ML provided critical feedback, suggested revisions, and refined the study design and analytical approach.

Drug and Alcohol REVIEW
APSAD
Drug and Alcohol Review (2021)
DOI: 10.1111/dar.13353

A parallel test of the SCRAM-CAM transdermal monitors ensuring reliability

KELLY VAN EGMOND[1], CASSANDRA J. C. WRIGHT[1,2,3,4], MICHAEL LIVINGSTON[1,5] & EMMANUEL KUNTSCHE[1,6]

[1]Centre for Alcohol Policy and Research, Department of Public Health and Psychology, La Trobe University, Melbourne, Australia, [2]Menzies School of Health Research, Charles Darwin University, Darwin, Australia, [3]Burnet Institute, Melbourne, Australia, [4]School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia, [5]National Drug Research Institute, Curtin University, Melbourne, Australia, and [6]Institute of Psychology, Eötvös Loránd University, Budapest, Hungary

Abstract

Introduction. *Previous studies validating the transdermal alcohol concentration (TAC) as measured by the Secure Continuous Remote Alcohol Monitors Continuous Alcohol Monitoring (SCRAM-CAM) have tested the monitor against self-reports or breath alcohol concentration (BrAC). This study aims to provide further evidence of the reliability of the SCRAM-CAM testing two monitors in parallel.* ***Methods.*** *Participants (*$N = 21$*) received four standard drinks in a laboratory session while wearing SCRAM-CAMs simultaneously on their left and right ankles. The SCRAM-CAMs sampled TAC every 30 min and participants were monitored for at least 2–3 h after their BrAC levels reached zero. Weight and height measures were taken to calculate body mass index (BMI).* ***Results.*** *There was a positive correlation between the TAC measurements from the left and right SCRAM-CAM (*$r = 0.718$*), a cross-correlation model revealed that this correlation was not significantly different for sex or BMI. Area under the TAC curve (AUC) and peak TAC values as measured by the left and right SCRAM-CAM also show positive correlations (*$r = 0.554$ *and* $r = 0.579$*, respectively). Cross-correlation models show a significant effect of BMI on the relationship between left and right peak TAC values, which may be due to outlier effects. No further effects were significant for on both peak and AUC values.* ***Discussion and Conclusions.*** *Results show that TAC measured by SCRAM-CAMs worn on the left and right showed a good correlation, with correlations between AUC and peak TAC values considered to be fair. TAC monitors show promise for use in research settings; however, work is needed testing the reliability of TAC as measured by two TAC monitors.* [van Egmond K, Wright CJC, Livingston M, Kuntsche E. A parallel test of the SCRAM-CAM transdermal monitors ensuring reliability. *Drug Alcohol Rev* 2021]

Key words: transdermal alcohol monitoring, SCRAM-CAM, test–retest reliability, validation.

Introduction

Currently, research on alcohol consumption mostly relies on self-reported data; as with other behavioural research, few alternatives have been traditionally available [1,2]. Previous research has shown high test–retest reliability of self-reported alcohol measures, including the timeline follow-back method ($r = 0.79$–0.99) [3,4] and online survey methods ($r = 0.83$–0.99) [5,6]. Self-report data have a range of advantages: collection is relatively affordable and data on long-term patterns of consumption can easily be collected. However, for in-the-moment alcohol use information, self-reports generally lack detail and accuracy and are routinely subject to bias due to memory deficits, contextual factors and motivation to report [7–9]. This has resulted in new developments, including ecological momentary assessments (in-the-moment data collection, often captured on digital devices) [9–11] and wearable technologies [12–14] that offer opportunities to further increase the detail, accuracy and cost-efficiency of alcohol measurements. Measurement of blood and/or breath alcohol concentration (BAC/BrAC) can provide more objective measures of quantity of alcohol consumed and previous research has shown a high test–retest reliability for fuel-cell breathalysers ($r \approx 0.99$) [15]. However, these measures require active participation and cooperation, which can result in a high response burden and

Kelly van Egmond MSc, PHD Candidate, Cassandra J. C. Wright PhD, Research Fellow, Michael Livingston PhD, Research Fellow, Emmanuel Kuntsche PhD, Director. Correspondence to: Ms Kelly van Egmond, Centre of Alcohol Policy Research, School of Psychology and Public Health, La Trobe University, Bundoora, VIC 3086, Australia. Tel: +61 04 98257320; E-mail: k.vanegmond@latrobe.edu.au

Received 17 December 2020; accepted for publication 11 June 2021.

avoidance of assessments and in turn can influence participants' natural drinking behaviour [16]. Breathalyser readings can also be subject to bias due to the presence of residual alcohol in the linings of the mouth, increasing BrAC levels [13,16,17]. Further, breathalysers only capture single timepoints in a drinking event rather than information about total consumption across the event. Transdermal alcohol monitors, such as the Secure Continuous Remote Alcohol Monitors Continuous Alcohol Monitoring (SCRAM-CAM™; Alcohol Monitoring Systems Inc., Highlands Ranch, CO, USA), can address some of the limitations of self-report and BrAC measurement by using passive, continuous monitoring of transdermal alcohol concentration (TAC). Approximately 1% of consumed alcohol is secreted through the skin; the SCRAM-CAM ankle bracelet samples and analyses the transdermal ethanol vapor close to the skin using a platinum-based electrochemical fuel-cell, which determines the ethanol content through catalytic alcohol oxidation. This enables TAC monitors to provide continuous measurement of alcohol consumption during a drinking event and over an extended period (i.e. months at a time).

The SCRAM-CAM is currently predominantly used in criminal justice systems internationally to monitor court-ordered abstinence for people charged with drink driving, domestic violence and alcohol-related offences [18,19]. TAC monitors also provide an opportunity for researchers to obtain detailed information about alcohol consumption and intoxication. However, depending on the research aims, some researchers might require more detailed drinking information than the binary coding system developed for the justice system. For this purpose, validation is necessary to understand their reliability. This has led to several studies to compare TAC data from SCRAM-CAM monitors to either BrAC [17,20,21] or self-reported data [13,22–25]; a recent systematic review was published which summarises validation studies of TAC monitors [12]. One study found a positive correlation between the TAC and BrAC measures for both the peak values ($r = 0.56$) and the area under the TAC curve (AUC) ($r = 0.60$) [21]. Studies have also found a positive correlation between the AUC and self-reported number of drinks ($r = 0.79$–0.94) [13,22]. Because of this high correlation, previous research used the AUC to compare the TAC data to the number of drinks consumed. Luczak and Rosen studied the validity of the TAC data from the Wrist Transdermal Alcohol Sensor from Giner Inc. (WrisTAS™) [26], with participants each wearing two monitors simultaneously. The authors reported observed differences in the peak TAC measurements, but they did not report on the correlation between the two datasets, as this was not the focus of the study. Collecting data from two wearable monitors simultaneously is a novel method that can be used to assess test–retest reliability, as the datasets can be compared directly. Given that previous research has extensively studied the correlation between TAC with other measures like BrAC and self-reports, this study will take this a step further by studying the correlation between TAC from two similar SCRAM-CAMs. To our knowledge, no published studies have assessed the correlation between TAC datasets from two simultaneously worn devices, and thus test–retest reliability is unknown. A high test–retest reliability and accuracy will increase our confidence in the use of the SCRAM-CAM for alcohol measurement in research.

Reliability is influenced by measurement errors that can be caused by individual differences [15,27,28]. Alcohol metabolism differs across individuals [29] and is known to be influenced by factors, such as sex and body mass index (BMI) [29–31]. This is reflected in differences in peak measures of BrAC [30–32]. In addition, females have been reported to have significantly higher peak TAC [33] and AUC [33] than males. The magnitude of these sex differences increased in line with the number of standard drinks consumed. Moreover, BMI was reported as a significant predictor of alcohol detection (odds ratio 0.95, 95% confidence interval 0.92–0.99, $P = 0.006$), with a lower alcohol detection rate among participants with a higher BMI [12,29]. It is therefore important for any validity or reliability study to assess whether these factors influence the TAC measurements, and if they do to make sure that they affect SCRAM-CAM monitors similarly.

The primary goal of this study was to assess the test–retest reliability of the SCRAM-CAM monitors by investigating the relationship between the TAC measures from two monitors worn in parallel (worn on both the left and right ankle by the same individual). According to the guidelines for clinical research measurements, when the reliability coefficient is below 0.40, the level of clinical significance is poor; when it is between 0.40 and 0.59, the level of clinical significance is fair; when it is between 0.60 and 0.74, the level of clinical significance is good; and when it is between 0.75 and 1.00, the level of clinical significance is excellent [29]. Given previous research into the test–retest reliability of self-reports [3–6] and BrAC [15] measurements finding an excellent correlation (>79) between measurements at different times, we expect the TAC measurements measured at the same time by the SCRAM-CAM to show similar or higher correlations. Secondary aims were to investigate the strength of this relationship and to study whether sex and BMI influenced this relationship. We hypothesised that sex and BMI would affect the TAC measurements from the two SCRAM-CAM monitors similarly and would

thus not affect the relationship between these TAC measurements.

Methods

Participants and criteria

A total of 23 healthy males ($n = 12$) and females ($n = 11$) aged between 18 and 35 years were recruited using targeted Facebook advertisements and street intercept approaches. Participants were screened to determine eligibility and obtain consent to be contacted for further screening. Inclusion criteria were: aged between 18 and 35 years old; regular consumption of alcohol (≥2 standard drinks during one single occasion in the preceding month); and able to read and understand English. Exclusion criteria were: current diagnosed mental health condition or high psychological distress (≥ 25 on the Kessler Psychological Distress Scale; K10 [34]); high-risk drinking behaviour (≥10 Alcohol Use Disorders Identification Test – Alcohol Consumption Questions [35]); being underweight or obese (BMI outside healthy range: 18–30); daily smoking; use of medication proven to interact with alcohol; diagnosis of diabetes mellitus, liver disease, kidney disease or sleep disorder; current infection of any kind or a metal allergy. Participants with a current or possible pregnancy or who were breastfeeding were also excluded. Participants were invited to three laboratory days at times convenient to them. Written consent was collected from every participant prior to the start of the first day. The female participants were additionally asked to sign an informed consent to ensure that they were not pregnant or breastfeeding and were familiar with the risks of drinking and pregnancy, as required by our ethics committee. The study was approved by the Human Research Ethics Committee at La Trobe University (HEC19249).

Study procedure

Laboratory sessions involved the administration of alcohol under controlled conditions. Participants were instructed to not drink any alcohol at least 12 h prior to the session and fast for at least 3–4 h prior to arrival. Laboratory sessions commenced between 9.00 and 10.00 am. At the start of each session, participants were required to demonstrate an alcohol-free breath sample. To ensure accuracy, the monitors have an automated, built-in calibration process in which they initially take samples at 5-min intervals to adjust to a particular subject. Participants were fitted with two randomly selected monitors on each ankle at least an hour prior to drinking to ensure sufficient calibration. The participants received four standard drinks per session; each drink contained 32 mL of vodka (40% alcohol, 10 g of alcohol, corresponding to one Australian standard drink) mixed with 68 mL of soda and sugar-free cordial. The laboratory room was temperature controlled and contained tables and chairs so that participants remained seated during the session. Participants remained in the laboratory for 2–3 h after their BrAC levels reached 0.000. At the end of each participants' session, the TAC readings were uploaded to the online software and monitors were deactivated.

Measures

TAC: SCRAM-CAM bracelets automatically sampled TAC every 30 min.

Sex: Participants self-reported being male or female.

BMI: BMI was measured with scales in the lab and calculated using weight and height by the following formula:

$$\mathrm{BMI} = \mathrm{Weight\,(kg)}/[\mathrm{height\,(m)}]^2$$

Data processing and cleaning

A total of 69 data collection days were conducted. However, for 11 (15.9%) days, we later found that the SCRAM-CAMs had not been correctly calibrated, resulting in the exclusion of data from these days. Data processing was followed by a process of matching the data points from the left and right SCRAM transdermal monitor using time and day records. Several further exclusion criteria were used on the processed data. First, the data points from the one side were matched to the closest data point in time from the monitor on the other side as the monitors may start data collection or calibration at different times. This results in either more datapoints in the left or right TAC dataset and matching datapoints but with large time differences. When a matching data point was missing from either the left or the right side, this data point was excluded ($n_{\mathrm{datapoint}} = 88$). Second, the timing of sampling was compared; we excluded data from both the left and right datasets when the time difference between matched data points sampled by the two devices was greater than 10 min ($n_{\mathrm{datapoint}} = 240$). This was done because differences between the two datasets could be due to the metabolism stage and related intoxication level and not because of actual differences in measurement. This resulted in a mean time

difference of 2.9 (SD = 2.2) min. Thirdly, after cleaning the TAC datapoints we found that 10.1% ($n_{days} = 7$) of all days showed TAC readings of zero or close to zero for across the entire day; as we are only interested in comparing the measured alcohol curves to each other, these days were excluded. Two participants had all days removed from the dataset due to one of the above criteria. We therefore analysed data from 51 days (out of 69), including 664 TAC readings from 21 participants with an average of 2.3 sessions per participant.

Data analysis

First, the intraclass correlation coefficient (ICC) was used for the analysis. A two-way random-effect model based on single ratings and absolute agreement assessed the test–retest reliability of the TAC measurements. A fitted line scatter dot plot was used to illustrate this correlation. Interpretation was as follows: below 0.40, poor; between 0.40 and 0.59, fair; between 0.60 and 0.74, good; 0.75 or above, excellent [36]. To assess whether individual characteristics affected the reliability (i.e. the relationship between left and right SCRAM-CAM measures), two-level random slope regression models with cross-level interactions were estimated using Mplus (V8) statistical software [37]. For the baseline model, we regressed '*SCRAM-CAM reading left side*' (dependent variable) on the *'SCRAM-CAM reading right side'* (first-level predictor). In the cross-level interaction model, variations in the strength of the relationship between the left and right TAC measures (S1) were subsequently regressed on the different characteristics of the participants: *'sex'* and *'BMI'* (second-level predictors). The within-level coefficients included in the cross-level interaction model were adjusted for time difference.

Further, AUC and peak TAC values per participant per day were calculated for each device, given the frequent use of these measures in previous TAC studies [12,13,21,22]. Mean AUC values were calculated using the linear trapezoidal method [38]. Peak TAC values were calculated by taking the highest TAC value. We used similar methods to assess the reliability of peak and AUC measures. A fitted line scatter plot was used to illustrate the relationship between the left and right AUC/peak TAC values. ICC was used to assess test–retest reliability prior to a baseline regression model, in which the '*AUC/Peak values left side'* (as the dependent variable) was regressed on '*AUC/Peak values right side'* (first-level predictor). However, given that the AUC and peak values are summary measures, the AUC and peak values of the right were regressed on the AUC and peak TAC values of the left and on the different characteristics of the participants: *'sex'* and *'BMI'*. To assess whether the relationship between left and right AUC or peak TAC was significantly stronger according to sex or BMI, interaction terms were included. As BMI has continuous values, we z-transformed the interaction variables to make them more comparable. The complex modelling option of Mplus was used to adjust the standard error for the clustering of observations within individuals.

Results

Participant characteristics

Table 1 displays demographics of the 21 participants included in analyses. Just over half were female ($n = 11$), the mean age was 27 and average BMI was 24.4. Males and females did not significantly differ from each other in age and BMI. There were significant sex differences in Alcohol Use Disorders Identification Test – Alcohol Consumption scores with males scoring higher (i.e. showing more hazardous drinking behaviour) compared to females.

Relationship between the left and right TAC values

Figure 1 shows that the left and right SCRAM-CAMs had positively related TAC measurements. The ICC for test–retest reliability was good at 0.72 (0.68–0.74) ($r = 0.717$; 95% confidence interval 0.68, 0.75). See

Table 1. *Participant's demographics*

Characteristics	Males ($n = 10$)	Females ($n = 11$)	Combined ($N = 21$)	Sex differences[a]
Age, years	26.3 (5.3)	27.6 (5.7)	27 (5.4)	$P = 0.584$
Body mass index	25.6 (3.8)	23.2 (3.9)	24.4 (4.0)	$P = 0.169$
AUDIT score	7.7 (1.6)	4.8 (1.9)	6.2 (2.3)	$P = 0.001$

Shown are means with standard deviations in brackets. [a]Sex differences were tested using the independent samples *t*-test. AUDIT, Alcohol Use Disorders Identification Test.

Supporting Information for subgroup correlations (Figure S1).

The baseline model results confirmed the positive relationship between the left and right TAC measures (Table 2).

In the cross-correlation model (Table 2), the S1 intercept parameter shows higher left TAC readings were correlated with higher right TAC readings among the lowest BMI-healthy men. The difference in the strength of this relationship compared to females and those with increasing BMI was small and non-significant (Table 2).

Relationship between the left and right AUC and peak TAC values

Figure 2a shows a positive correlation between the AUC values for the participants per day of the left and right TAC datasets. The ICC for test–retest reliability was considered fair at 0.56 (0.33–0.72). See Supporting Information for subgroup correlations (Figure S2).

Figure 2b shows a positive correlation between the peak TAC values per participant per day of the left and right TAC datasets. The mean difference between

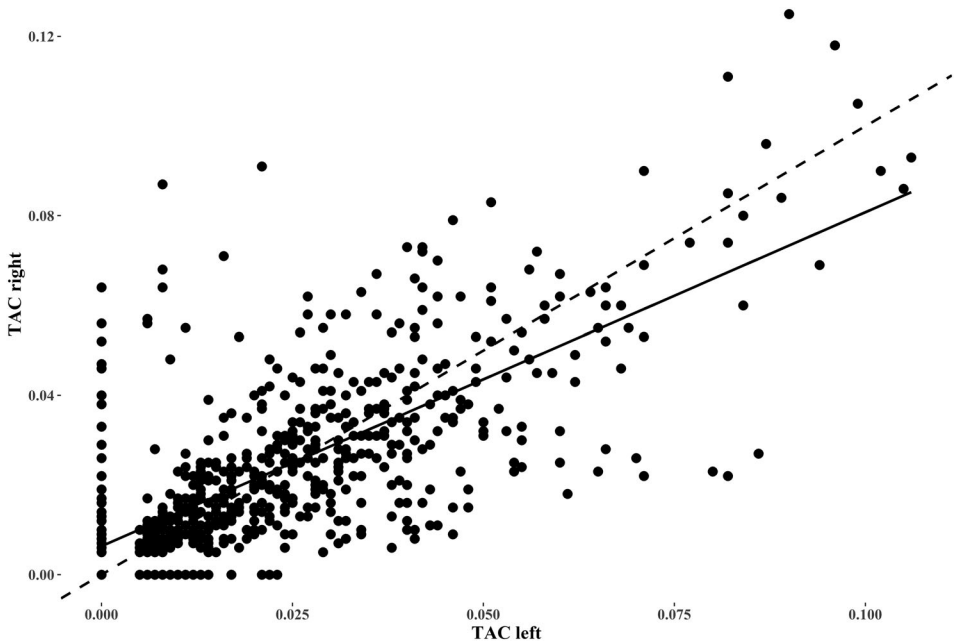

Figure 1. *Scatterplot of TAC measures from left on right. TAC, transdermal alcohol concentration.*

Table 2. *TAC datapoints as measured by the SCRAM-CAM worn on the right regressed on the TAC datapoints as measured by the SCRAM-CAM worn on the left*

	Baseline model		Cross-level interaction model	
Parameter	Coefficient	*P*-value	Coefficient	*P*-value
Intercept	0.007 (0.001)	0.000		
TAC SCRAM right	0.691 (0.026)	0.000		
S1 intercept[a,b]			1.087 (0.287)	0.000
Effect of sex[c] on S1			−0.127 (0.219)	0.536
Effect of BMI[d] on S1			−0.022 (0.026)	0.369

Shown are unstandardised regression coefficients with standard errors in brackets. [a]S1 intercept = Relationship of left TAC with right TAC in the males with the lowest BMI. [b]Within-level coefficients are adjusted for time difference. [c]Effect of sex (coded as 0 = male, 1 = female) = Difference in the strength of the relationship between left with right TAC in females as compared to males. [d]Effect of BMI = Difference in the strength of the relationship between left and right TAC in the participants with a higher BMI. BMI, body mass index; SCRAM-CAM, Secure Continuous Remote Alcohol Monitors Continuous Alcohol Monitoring; TAC, transdermal alcohol concentration.

the peak values of the left SCRAM CAMs and the right SCRAM CAMs was 0.002 (0.2) and the ICC for test–retest reliability was considered fair at 0.58 (0.36–0.74). See Supporting Information for subgroup correlations (Figure S2).

Baseline regression model results confirmed the positive relationships between both AUC and peak TAC values from the left and right SCRAM-CAM (Table 3). The effects of sex on the strength of relationship between the AUC and peak TAC values from the left and right SCRAM-CAM were small and non-significant (Table 3). BMI did not significantly affect the strength of relationship between the AUC TAC values from the left and right SCRAM-CAM. However, there was a significant relationship between reduced peak TAC values from the left and right SCRAM-CAM with increasing BMI (Table 3).

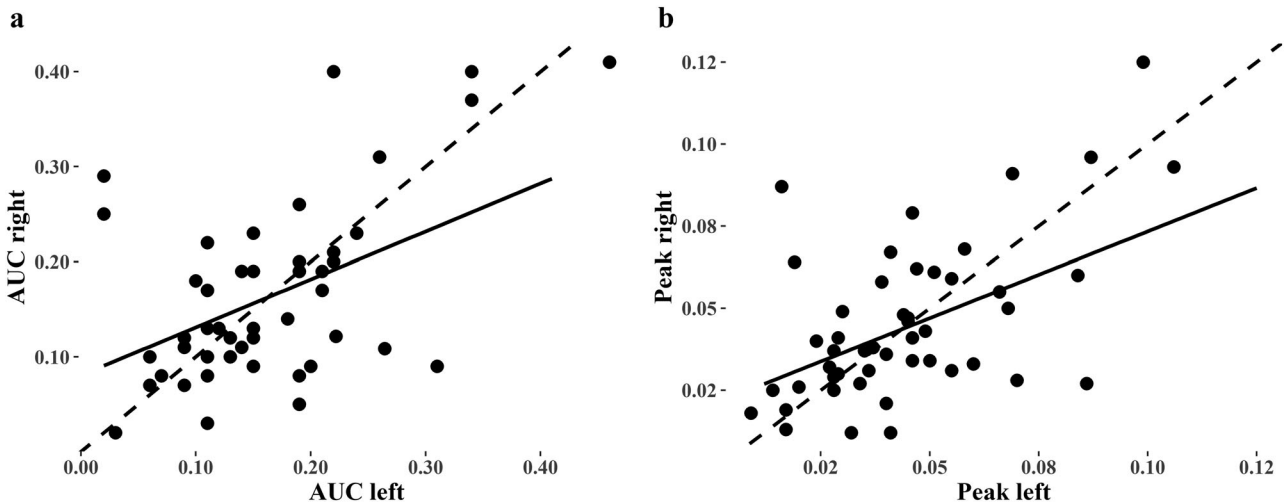

Figure 2. *Scatterplot of (a) area under the TAC curve (AUC) right on left and (b) peak TAC measures of right on left. TAC, transdermal alcohol concentration.*

Table 3. *TAC datapoints as measured by the SCRAM-CAM worn on the right regressed on the TAC datapoints as measured by the SCRAM-CAM worn on the left and sex and BMI*

	AUC		Peak	
Parameter	Coefficient	*P*-value	Coefficient	*P*-value
Baseline model				
Intercept	0.067 (0.024)	0.006	0.017 (0.006)	0.007
TAC SCRAM left	0.611 (0.131)	0.000	0.634 (0.128)	0.000
Cross-correlation interaction model				
Intercept	0.095 (0.148)	0.520	−0.002 (0.036)	0.966
Main effects				
TAC SCRAM left	0.693 (0.364)	0.057	**0.944 (0.353)**	**0.008**[a]
Sex	−0.101 (0.078)	0.199	−0.020 (0.021)	0.364
BMI	0.002 (0.008)	0.770	0.003 (0.002)	0.154
Interaction effects				
TAC left * Sex	0.051 (0.031)	0.098	0.007 (0.008)	0.365
TAC left * BMI	−0.044 (0.032)	0.175	**−0.022 (0.008)**	**0.008**[a]

Intercept = value of the right peak/AUC TAC for lowest BMI males with a peak/AUC TAC reading of zero. Main effects = effect of female on male and higher BNMI on the lowest BMI. Sex coded as 0 = male, 1 = female. Interaction effects = testing the magnitude in difference in the strength of the relationship between the left with right AUC and peak TAC values with increasing BMI and for females, respectively. [a]Significance level = $P < 0.05$. AUC, area under the curve; BMI, body mass index; SCRAM-CAM, Secure Continuous Remote Alcohol Monitors Continuous Alcohol Monitoring; TAC, transdermal alcohol concentration.

Discussion

The aim of this study was to test the reliability of the TAC data as measured by the SCRAM-CAM.

Relationship between TAC data as measured by the left and right SCRAM-CAM

Given previous research into the test–retest reliability of self-reports [3–5] and BrAC [15] measurements finding an excellent correlation (>0.79) between measurements at different times, we had expected the TAC data measured at the same time from the SCRAM-CAM to show similar or even better correlation measures. However, both the ICC and the Baseline Regression Model show coefficients around 0.70; although the correlation is lower than expected, this is still considered to be good for clinical measurements [36]. Further, TAC collects information more frequently as compared to self-reports or BrAC and is much more detailed, which could leave more room for outliers resulting in lower correlations. Future studies replicating the parallel study procedure with TAC monitors will enhance our understanding of the reliability of the TAC data in different settings.

In line with previous research and our second hypothesis, we found no evidence that the strength of the relationship between the left and right values was influenced by sex and BMI. This suggests that small differences between the left and right datasets cannot be explained by the individual characteristics we examined here.

Relationship between the left and right AUC and peak TAC values

Given that the TAC on each side was measured by SCRAM-CAMs, we expected higher correlations than seen in studies testing TAC data against different measures, such as BrAC [21] and self-report measures [13,22]. The ICC for test–retest reliability and the baseline models showed a lower correlation between AUC and peak TAC measures than expected. In line with previous research and our second hypothesis, we found no evidence that the strength of the relationship between the AUC values from the left and right datasets was influenced by either sex or BMI [28]. The results further showed that sex did not affect the relationship between peak TAC values from the left and right SCRAM-CAM, but BMI did. This means that the strength of the relationship was significantly lower in participants with higher BMIs as compared to participants with lower BMIs. As mentioned, previous research found that the detection rate is significantly affected by BMI; the detection of alcohol decreases with an increase in BMI, resulting in a higher chance of undetected drinking events in individuals with a higher BMI [28]. This increases the room for error and differences in the detection of alcohol among the SCRAM-CAM monitors worn left and right. It is possible that these differences were reflected in the peak TAC values from the left and right SCRAM-CAM monitor. Given the small sample size in the peak TAC analysis, the effect of BMI may simply be an outlier effect and should therefore be interpreted with caution. Overall, our results do imply measurement error in the SCRAM-CAM data, and if this cannot be explained by sex or BMI, there may be other factors causing these differences. For example, differences in the tightness of the monitor (the WrisTAS and the SCRAM-CAM are manually strapped on to an individual) or a failure of the monitor itself may lead to lower-than-expected correlations between left and right measures. Future research will need to replicate these findings to provide further evidence of whether the effect of BMI on the reliability of the TAC data of the SCRAM-CAM is consistent as this would affect its use in research settings.

Limitations

A limitation of this study was the small sample size used in the analysis of the AUC and peak TAC. However, the sample size was consistent with previous studies testing the AUC and peak TAC values against BrAC and self-reported drinking [13,21,22]. When interpreting the results and outcomes of this study, one should keep in mind that the process of cleaning and processing the data resulted in the exclusion of a substantial amount of data due to improper calibration and failures in the readings. We did still observe instances of zero TAC values measured by one SCRAM-CAM while the other showed a non-zero TAC value (see Figure 1); this could have resulted in a lower correlation. However, given that these zeros were caused by a lower TAC reading and not a faulty monitor, we chose not to exclude these observations. To investigate the effect that this had our analyses, we trialled excluding these observations and this resulted in a slightly lower ICC of 0.69. Furthermore, while the manufacturer has developed criteria to clean TAC data, this is not meant for research purposes and there are no standardised data cleaning approaches in the research literature [12]. We therefore recommend that researchers invest substantial time in piloting the devices prior to data collection to identify faulty

monitors, refine calibration and ensure consistency in strap tightness to minimise the need to exclude data due to cleaning. Standardised cleaning and processing procedures should be developed and used across different research groups to ensure comparability in future studies. As we randomly selected the SCRAM-CAMs per session per participants and after cleaning the data, we had only 2.3 sessions average per participant; this limited our ability to study between-session effects. Future studies will need to further assess the between-session effects. Given the scope of this article, we focussed solely on the test–retest reliability and did not test the agreement by including the discrepancies between the left and right TAC measurements. Future work will need to examine the magnitude of the differences between the two devices by including agreement tests using the difference score.

Conclusion

The current study was the first to focus on studying TAC monitors in parallel, to provide evidence regarding the reliability of TAC measurements as measured by the SCRAM-CAM. We found a positive relationship between TAC measurements from the two SCRAM-CAMs; however, the correlation was lower than expected, suggesting significant measurement error. The lower-than-expected correlations could be related to individual differences; however, they are more likely due to environmental factors, including the tightness of the monitor strap. Further work is needed to further explore the source of error within and between TAC measurements from SCRAM-CAM devices. The current study highlights a need for further improvement on TAC measurement to improve their application in research settings. TAC monitors are promising with their advantages of providing objective and continuous measures of alcohol consumption. However, improving reliability is important for their use in research.

Acknowledgements

Research reported in this publication was funded using an internal La Trobe University grant. Furthermore, KvE was supported by the La Trobe Full Fee Research Scholarship and La Trobe University Postgraduate Research Scholarship, CJCW was funded by a National Health and Medical Research Council Early Career Fellowship (1161246) and ML was funded by a National Health and Medical Research Council Career Development Fellowship (1123840). We would like to gratefully acknowledge the help and support in the fieldwork from our research colleagues Gabriel Caluzzi and Dan Anderson-Luxford.

Conflict of Interest

The authors have no conflicts of interest.

References

[1] May C, Foxcroft D. Minimizing bias in self-reports of health beliefs and behaviours. Health Educ Res 1995;10:107–12.
[2] Lajunen T, Summala H. Can we trust self-reports of driving? Effects of impression management on driver behaviour questionnaire responses. Transp Res Part F Traffic Psychol Behav 2003;6:97–107.
[3] Sobell MB, Sobell LC, Klajner F, Pavan D, Basian E. The reliability of a timeline method for assessing normal drinker college students' recent drinking history: utility for alcohol research. Addict Behav 1986;11: 149–61.
[4] Sobell LC, Sobell MB, Leo GI, Cancilla A. Reliability of a timeline method: assessing normal drinkers' reports of recent drinking and a comparative evaluation across several populations. Br J Addict 1988;83: 393–402.
[5] Khadjesari Z, Murray E, Kalaitzaki E *et al.* Test–retest reliability of an online measure of past week alcohol consumption (the TOT-AL), and comparison with face-to-face interview. Addict Behav 2009;34:337–42.
[6] Miller ET, Neal DJ, Roberts LJ *et al.* Test-retest reliability of alcohol measures: is there a difference between internet-based assessment and traditional methods? Psychol Addict Behav 2002;16:56–63.
[7] Karns-Wright TE, Dougherty DM, Hill-Kapturczak N, Mathias CW, Roache JD. The correspondence between transdermal alcohol monitoring and daily self-reported alcohol consumption. Addict Behav 2018;85: 147–52.
[8] Maylor EA, Rabbit P, Kingstone A. Effects of alcohol on word categorization and recognition memory. Br J Psychol 1987;78:233–9.
[9] Shiffman S. Ecological momentary assessment (EMA) in studies of substance use. Psychol Assess 2009;21:486–97.
[10] Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1–32.
[11] Wright C, Dietze PM, Agius PA *et al.* Mobile phone-based ecological momentary intervention to reduce young adults' alcohol use in the event: a three-armed randomized controlled trial. JMIR Mhealth Uhealth 2018; 6:e149.
[12] van Egmond K, Wright CJ, Livingston M, Kuntsche E. Wearable transdermal alcohol monitors: a systematic review of detection validity, and relationship between transdermal and breath alcohol concentration and influencing factors. Alcohol Clin Exp Res 2020;44:1918–32.
[13] Leffingwell TR, Cooney NJ, Murphy JG *et al.* Continuous objective monitoring of alcohol use: twenty-first century measurement using transdermal sensors. Alcohol Clin Exp Res 2013;37:16–22.
[14] Lansdorp B, Ramsay W, Hamid R, Strenk E. Wearable enzymatic alcohol biosensor. Sensors 2019;19:2380.
[15] Sorbello JG, Devilly GJ, Allen C, Hughes LR, Brown K. Fuel-cell breathalyser use for field research on alcohol intoxication: an independent psychometric evaluation. PeerJ 2018;6:e4418.
[16] Fairbairn CE, Rosen IG, Luczak SE, Venerable WJ. Estimating the quantity and time course of alcohol consumption from transdermal alcohol sensor data: a combined laboratory-ambulatory study. Alcohol 2018; 81:111–6.
[17] Dougherty DM, Charles NE, Acheson A *et al.* Comparing the detection of transdermal and breath alcohol concentrations during periods of alcohol consumption ranging from moderate drinking to binge drinking. Exp Clin Psychopharmacol 2012;20:373–81.
[18] McKnight AS, Fell JC, Auld-Owens A. Transdermal alcohol monitoring: Case studies. Washington: National Highway Traffic Safety Administration, 2012. (Report no. DOT HS 811 603).

[19] Robertson R, Vanlaar W, Simpson HM. Continuous transdermal alcohol monitoring: a primer for criminal justice professionals. Ottawa: Traffic Injury Research Foundation, 2007.
[20] Hill-Kapturczak N, Lake SL, Roache JD, Cates SE, Liang Y, Dougherty DM. Do variable rates of alcohol drinking alter the ability to use transdermal alcohol monitors to estimate peak breath alcohol and total number of drinks? Alcohol Clin Exp Res 2014;38: 2517–22.
[21] Fairbairn CE, Kang D. Temporal dynamics of transdermal alcohol concentration measured via new-generation wrist-worn biosensor. Alcohol Clin Exp Res 2019;43:2060–9.
[22] Barnett NP, Tidey J, Murphy JG, Swift R, Colby SM. Contingency management for alcohol use reduction: a pilot study using a transdermal alcohol sensor. Drug Alcohol Depend 2011;118:391–9.
[23] Roache JD, Karns TE, Hill-Kapturczak N *et al.* Using transdermal alcohol monitoring to detect low-level drinking. Alcohol Clin Exp Res 2015; 39:1120–7.
[24] Rash CJ, Petry NM, Alessi SM, Barnett NP. Monitoring alcohol use in heavy drinking soup kitchen attendees. Alcohol 2018;81:139–47.
[25] Barnett NP, Celio MA, Tidey JW, Murphy JG, Colby SM, Swift RM. A preliminary randomized controlled trial of contingency management for alcohol use reduction using a transdermal alcohol sensor. Addiction 2017;112:1025–35.
[26] Luczak SE, Rosen IG. Estimating BrAC from transdermal alcohol concentration data using the BrAC estimator software program. Alcohol Clin Exp Res 2014;38:2243–52.
[27] Gullberg RG. Estimating the measurement uncertainty in forensic breath-alcohol analysis. Accredit Qual Assur 2006;11:562–8.
[28] Barnett NP, Meade EB, Glynn TR. Predictors of detection of alcohol use episodes using a transdermal alcohol sensor. Exp Clin Psychopharmacol 2014;22:86–96.
[29] Collins AC, Yeager TN, Lebsack M, Panter SS. Variations in alcohol metabolism: influence of sex and age. Pharmacol Biochem Behav 1975; 3:973–8.
[30] Hill-Kapturczak N, Roache JD, Liang Y, Karns TE, Cates SE, Dougherty DM. Accounting for sex-related differences in the estimation of breath alcohol concentrations using transdermal alcohol monitoring. Psychopharmacology (Berl) 2015;232:115–23.
[31] Dettling A, Fischer F, Böhler S *et al.* Ethanol elimination rates in men and women in consideration of the calculated liver weight. Alcohol 2007; 41:415–20.
[32] Dougherty DM, Hill-Kapturczak N, Yuanyuan L *et al.* The potential clinical utility of transdermal alcohol monitoring data to estimate the number of alcoholic drinks consumed. Addict Disord Their Treat 2015; 14:124–30.
[33] Karns-Wright TE, Roache JD, Hill-Kapturczak N, Liang Y, Mullen J, Dougherty DM. Time delays in transdermal alcohol concentrations relative to breath alcohol concentrations. Alcohol Alcohol 2017;52: 35–41.
[34] Kessler RC, Barker PR, Colpe LJ *et al.* Screening for serious mental illness in the general population. Arch Gen Psychiatry 2003;60:184–9.
[35] Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA, Project ftACQI. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Arch Intern Med 1998;158:1789–95.
[36] Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 1994;6:284–90.
[37] Muthén L. Mplus user's guide. Los Angeles: Muthén & Muthén, 2010: 1998–2020.
[38] Dodd LE, Pepe MS. Partial AUC estimation and regression. Biometrics 2003;59:614–23.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Figure S1. Scatterplot of TAC right on left for (a) sex; male and female, (b) BMI; healthy BMI and overweight (for visual presentation purposes, BMI has been dichotomised based on healthy (18.5–24.9) and overweight (>25).
Figure S2. Scatterplot per subgroup for (a) AUC left on right for males and females, (b) AUC left on right for participants with a healthy BMI and participants with overweight, (c) Peak values left on right for males and females, (e) Peak values left on right for participants with a healthy BMI and participants with overweight (for visual presentation purposes, BMI has been dichotomised based on healthy (18.5–24.9) and overweight (>25).

5. Transdermal monitors – agreement and reliability

Chapter 5 contains the following manuscript:

van Egmond, K., Riordan, B., JC Wright, C., Livingston, M., Kuntsche, E. (submitted for publication). To what degree can we rely transdermal monitors to measure alcohol consumption? A study on agreement and reliability.

Author contributions

KvE and CW collected the data; KvE analysed the data and led data interpretation and manuscript writing. BR provided support for data analysis. EK, CW, BR and ML provided critical feedback, suggested revisions, and refined the study design and analytical approach.

To What Degree Can We Rely Transdermal Monitors to Measure Alcohol Consumption? A Study on Agreement and Reliability.

Kelly van Egmond[1] MSc, Benjamin Riordan[1] PhD, Cassandra J.C Wright[1,2,3] PhD, Michael Livingston[1,5,6] PhD, and Emmanuel Kuntsche[1,4] PhD.

1. Centre for Alcohol Policy and Research (CAPR), NR1 Building, La Trobe University, Melbourne, 3086 VIC, Australia.
2. Menzies School of Health Research, Charles Darwin University, Darwin, Australia
3. Burnet Institute, Melbourne, Australia.
4. Institute of Psychology, Eötvös Loránd University, Kazinczy u. 23-27, 1075 Budapest, Hungary
5. Faculty of Health Sciences, Curtin University, Melbourne, Australia
6. Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

*Corresponding author:

Emmanuel Kuntsche

Centre of Alcohol Policy Research

School of Psychology and Public Health,

La Trobe University

3086 Bundoora, VIC, Australia

e-mail: e.kuntsche@latrobe.edu.au, phone:

Abstract

Background. Wearable transdermal alcohol concentration (TAC) technology is rapidly evolving and there is a need to determine whether the available technology is reliable. In this study we investigate the test-retest reliability of two TAC monitors (SCRAM-CAM and ION RAP).

Methods. Participants received four standard drinks in a laboratory session while wearing SCRAM-CAMs and ION RAP monitors on the left and right ankles and wrists. Participants were assigned to different food (meal or snacks) and drinking rate (fast or slow) conditions. The final samples included 50 sessions from 21 participants (52.4% female, mean age=27.0 (5.4)) for SCRAM-CAM and three sessions from three participants (33.3% female, mean age=24.0 (5.3)) for ION RAP.

Results. There was a mean absolute difference of 0.010 g/dL between the TAC measures of the left and right SCRAM-CAM, with 14% of sessions recording differences higher than 0.020 g/dL. Differences generally increased at higher TAC. The effects of food consumption and drinking rate on this difference were small and non-significant (B=-0.004, p=0.292 and B=-0.005, p=0.143). Left and right ION RAP readings were highly correlated (ICC=0.81) but had an absolute difference of 131 nA with 30% of readings recording differences higher than our predetermined threshold.

Conclusion. Our study showed substantial variation between the same TAC monitors (left vs. right) in the same drinking event, with 14-30% of events producing disparities above a reasonable threshold. Thus, while these monitors show promise for research, there is still a large variability in measuring alcohol content. Depending on the research question, researchers need to determine the acceptable ranges of error in order to rely on the TAC data produced for decisions in clinical, diagnostic, or research settings.

Keywords: *Transdermal alcohol monitoring, Test-retest reliability, Agreement, SCRAM-CAM, ION RAP*.

Introduction

Transdermal alcohol concentration (TAC) monitors offer a way to collect continuous information on alcohol consumption by measuring alcohol that is secreted through the skin. They offer advantages over other methods of alcohol measurement such as breathalysers and self-reports, including reduced participant burden and continuous, non-invasive measurement over long periods of time (Swift, 2003; Karns-Wright *et al.*, 2017; van Egmond *et al.*, 2020).

One of the most widely used and tested TAC monitors is the Secure Continuous Remote Alcohol Monitors Continuous Alcohol Monitoring (SCRAM-CAM™; Alcohol Monitoring Systems Inc., Highlands Ranch, CO) (reviews: Leffingwell et al., 2013; van Egmond et al., 2020). Approximately 1% of consumed alcohol is secreted through the skin as a transdermal ethanol vapor. Using a platinum-based electrochemical fuel-cell, the SCRAM-CAM ankle bracelet samples this vapor in 30-minute intervals and determines the grams per decilitre (g/dL) of ethanol in the body through catalytic alcohol oxidation. To measure the validity and accuracy of the SCRAM-CAM, studies testing SCRAM-CAM TAC measures against breath alcohol concentration (BrAC), and self-reported measures of alcohol consumption found correlation coefficients ranging between 0.56-0.94 (Barnett *et al.*, 2011; Fairbairn and Kang, 2019; van Egmond *et al.*, 2020). Further, due to its low failure rate, the SCRAM-CAM has been found to be the most reliable TAC monitor (Fairbairn and Kang, 2019). However, limited research has focussed on the test-retest reliability of the SCRAM-CAM TAC data. Test-retest reliability is an important concept in the validation process of measurement tools, such as TAC monitors. Test-retest includes two related, but different concepts: reliability and agreement. Reliability is the ability of the monitor to replicate a similar measurement pattern across participants when applied twice. Agreement on the other hand requires the measurement tools to provide strictly identical values when applied twice under similar conditions.

In our recent publication, we investigated the reliability of the SCRAM-CAM TAC measures by comparing the TAC values from a monitor worn on the left to a monitor worn on the right ankle (Egmond *et al.*, 2021). The SCRAM-CAM TAC values were found to show good reliability with intra-class correlation coefficients around 0.70 (Egmond *et al.*, 2021). However, such correlation

measures tell us whether the TAC curves from the left and right SCRAM-CAM follow a generally similar alcohol measurement pattern (absorption and elimination) over time, they may not necessarily depict agreement in terms of the estimated *level* of alcohol. A low agreement between the left and right SCRAM-CAM on the same person and the same time would suggest that the devices lack precision, making it difficult for researchers to trust that the TAC data they produce are accurate. Further, if there are any systematic differences, this would have implications on how we interpret the data and which side of the body we might strap the monitor on. To our knowledge, together with our previous publication (Egmond *et al.*, 2021), these studies are the only studies on the test-retest reliability of the SCRAM-CAM monitor. Alcohol researchers generally aim to measure the amount of alcohol consumed, therefore, an alcohol monitor with low test-retest reliability could obscure the real intoxication levels and lead to inaccurate conclusions that may impact research and clinical outcomes (Bland and Altman, 1986; Berchtold, 2016).

Whereas SCRAM-CAM monitors have been available for years, newer models using different technology have recently been released. New-generation wristbands, such as the ION RAP™ by MILO Sensors™ (Lansdorp *et al.*, 2019), are significantly smaller and have increased sampling frequency as compared to the SCRAM-CAM, improving the user experience (Wright *et al.*, 2021) and the level of detail. In contrast to the fuel-cell technology used by the SCRAM-CAM, the ION RAP uses an enzymatic detection pathway, via alcohol oxidase, to measure TAC in terms of a raw current (nA). The ION RAP samples TAC every five seconds or less, which is markedly more frequent than the SCRAM-CAM's 30-minute sampling intervals. Given that these devices are still so new, studies using the ION RAP are limited, with the only published study conducted by MILO Sensor's own researchers showing that an early prototype of the ION RAP was able to capture in-the-moment alcohol consumption, with only a slight underestimation as compared to a theoretical peak BAC (Lansdorp *et al.*, 2019). To our knowledge, there has been no work assessing the reliability and agreement of the TAC data as measured by the ION RAP.

Reliability and agreement can be affected by measurement errors as a result of individual differences between participants (Gullberg, 2006; Barnett, Meade and Glynn, 2014; Sorbello *et al.*,

2018). Our previous study on SCRAM reliability found no significant effects of gender, and even though the results showed a significant effect of BMI, this was likely due to outliers (Egmond *et al.*, 2021). Food consumption has been observed to slow the absorption and to increase the elimination of alcohol (Holt, 1981; Swift, 2003). Further, peak alcohol levels are lower when a drink is consumed over time as compared to rapidly (Dasgupta, 2017). A first examination did not find any effects of food on the TAC curves (Saldich *et al.*, 2021) and Hill-Kapturczak et al., (2014) found that the drinking rate does not affect the ability of their model to predict BrAC from TAC, however, how food and drinking rate affects TAC reliability and agreement is unknown.

This study had 3 aims: 1) To examine the agreement of the SCRAM-CAM devices, 2) To examine the reliability and agreement of the ION RAP devices, and 3) To assess whether food consumption and drinking rate affect reliability and agreement for the SCRAM-CAM devices. This study investigates the test-retest reliability of two different monitors (SCRAM-CAM and ION RAP), but it does not set out to directly compare them. This is because even though the conditions in which both monitors were tested were the same, the differences in mechanism (fuel-cell technology versus enzymatic detection), output (TAC in g/dL versus raw current (nA)), and state of development (SCRAM-CAM being established versus the ION RAP prototype), make it problematic to strictly compare the monitors' performance to each other.

Methods

Participants. A total of 23 participants, 12 men and 11 women with an average age of 26.4 (SD=5.5) were recruited using targeted Facebook advertisements and street intercept approaches. All were aged between 18-35 years old, regularly consumed alcohol (≥ 2 standard drinks during at least one occasion in the preceding month), were able to read and understand English. Exclusion criteria were: currently diagnosed mental health condition, or high psychological distress (≥ 25 on the Kessler Psychological Distress Scale; K10, (Kessler *et al.*, 2003)), at-risk drinking behaviour (≥ 10 Alcohol Use Disorders Identification Test – Alcohol Consumption Questions; AUDIT-C, (Bush *et al.*, 1998)), being under- or overweight (Body Mass Index (BMI) outside healthy range: 18-30), daily smokers, use of medication proved to interact with alcohol, diagnosis of diabetes mellitus, liver disease, sleep

disorder, current infection of any kind, kidney disease, or a metal allergy, a current or possible pregnancy, or breastfeeding. The study was approved by the Human Research Ethics Committee at La Trobe University (HEC19249).

Study procedure. Participants were invited for three full-day laboratory alcohol sessions for which they were instructed not to drink alcohol 12 hours prior to the session and fast for at least 3-4 hours prior to arrival. Two hours prior to the start of the session, cartridges for the ION RAP were activated and inserted in the wristband by trained researchers (see Figure 1 for cartridge). Upon arrival at the laboratory (between 9-10am), participants had to provide an alcohol-free breath sample and were asked to complete a baseline questionnaire. Participants were then fitted with the SCRAM-CAM and ION RAP monitors at least an hour prior to drinking to allow for calibration. All participants were fitted with a SCRAM-CAM monitor on each ankle simultaneously (two monitors worn simultaneously). When enough devices were available and functioning, participants wore an ION RAP wristband on each wrist simultaneously, however due to a limited number of wristbands available, most participants only wore one monitor. The participants were given four standard drinks each session, containing 32 millilitres of vodka (40% alcohol, 10 grams of alcohol, corresponding to one Australian standard drink) mixed with 68 millilitres of soda and sugar-free cordial. Participants remained in the laboratory for 2-3 hours after their BrAC levels reached 0.000. At the end of each participants' session, the TAC readings were uploaded to the online software and monitors were deactivated.

Measures

SCRAM-CAM TAC. SCRAM-CAM bracelets automatically sampled TAC every 30 minutes. An absolute difference score was calculated (TAC_{left} - TAC_{right}) per datapoint.

ION RAP TAC. ION RAP wristbands were pre-production prototypes and automatically sampled TAC around every five seconds. An absolute difference score was calculated (TAC_{left} - TAC_{right}) for every five minute-level datapoint.

Drinking rate. In two sessions, participants were to drink all four standard drinks within sixteen minutes. In the remaining session, they drank two standard drinks within an eight-minute period and the other two drinks one hour later, again within eight minutes.

Food consumption. Participants received a high-caloric meal (~500kcal) 30 minutes after finishing their drinks during one session and in the other two sessions they received a low-caloric snack (~100kcal) instead. After receiving either a meal or snack, participants were provided with a low-caloric snack in hourly intervals. A meal was always provided four hours after finishing the drinks.

Data processing and cleaning

*SCRAM-CAM***.** A total of 69 data person-days were conducted. However, 11 (15.9%) days were excluded due to incorrect calibration of the SCRAM-CAMs (i.e., not enough time had been allowed for prior to alcohol administration). Data processing and cleaning were completed using the same steps taken in our most recent study (Egmond *et al.*, 2021). Due to differences in calibration, TAC datapoints from the left and right SCRAM-CAM can be uneven in timing, which is why we first matched the data points from the one monitor to the closest data point in time of the other monitor. When a matching datapoint was missing, this datapoint was excluded ($n_{datapoint}$=88, 0.1%). Further, when the time difference between the matching datapoints was greater than 10 minutes, these datapoints were also excluded ($n_{datapoint}$=240, 0.2%). This was done because differences between the two datasets could be due to the metabolism stage and related intoxication level and not because of actual differences in measurement. Data processing and cleaning identified about 10.1% of all days showing TAC readings of zero or close to zero for a full day which were further excluded. This resulted in the inclusion of 50 days (out of 69, 72.5%) including 664 TAC readings from 21 participants with an average of 2.3 days per participant.

ION RAP. Of the 69 data collection days, there were only seven (10.1%) where a participant wore a wristband on both the left and right wrist. This was due to the researchers having only a small number of wristbands and challenges we experienced charging the batteries, meaning that fewer wristbands were available than planned. For reasons unknown, the ION RAP did not detect alcohol or showed a

disturbed alcohol signal (possibly due to a disruption in the Bluetooth signal) for four out of the seven days. Given that this resulted in indiscernible TAC curves, these days were excluded, leaving three days of data for analysis. ION RAP TAC data is in the form of an electrical current (nA) that starts at 1000 nA, this current slowly returns to baseline when the cartridge is inserted and with skin contact. When alcohol is detected, the current will start to increase, for this reason, the first point of detection was operationalized as the first point of increase after alcohol administration. To account for differences in the baseline, the TAC data was centred and standardized by subtracting the start values from all other readings.

To be able to match the TAC data as measured by the left and right ION RAP, the TAC data was aggregated to five minute-level datapoints, by taking the average of all readings in each five-minute interval, for the left and right TAC datasets separately and subsequently matched on this five minute-level. After data cleaning and processing, we were left with data from three days, including 836 matched TAC readings from three participants.

Data analysis

Data analyses for the SCRAM-CAM TAC and ION RAP TAC were completed separately and will thus be described and interpreted separately. Given our previous reliability analysis for the SCRAM-CAM TAC (Egmond *et al.*, 2021), this study will only test the agreement and influences of food consumption and drinking rate for SCRAM-CAMs. Due to limited days of TAC data, it was not possible to investigate the effects of food consumption and drinking rate for the ION RAPs, so we only report on reliability and agreement.

SCRAM-CAM. First an absolute difference score was calculated (TAC_{left} - TAC_{right}) per datapoint (N=544). To test overall agreement on all 544 datapoints, we used a one-sample t-test to study whether the absolute difference was significantly different from zero. Next, we used a modified Bland and Altman (B&A) plot to further describe the difference between the TAC measures from the left and right SCRAM-CAM. Given that we are comparing two SCRAM-CAM devices, the B&A plot was modified to include the overall bias (mean absolute difference), an acceptable limit of agreement

and a trend line to determine whether there was greater bias at higher TAC. A Pearson's correlation coefficient was added to the trend line to test whether there was indeed a correlation between the mean absolute difference and the average TAC. The SCRAM-CAM developers, AMS, have set a threshold of 0.02 g/dL to identify a possible drinking event (Roache *et al.*, 2015), which we also used as the acceptable limit for a good agreement between readings from the left and right monitor.

To test the effect of food consumption and drinking rate on the agreement between left and right and given food consumption and drinking rate were events on the day-level, the data was averaged for all the absolute difference scores within a day (11 on average) for all the 50 days. To investigate the effects of the alcohol administration condition (*meal* or *no meal* and *drinking rate being spread* (hour in between) or *not spread*) on the difference score, we used the complex modelling option of Mplus (version 8) to account for clustering of day-level information (N=21) within individuals (N=23).

ION RAP. First, intraclass correlation coefficients (ICC) were calculated using the two-way random-effect model based on single ratings and absolute agreement assessed the test-retest reliability of the TAC measurements with values below 0.40, between 0.40 and 0.59, between 0.60, and 0.74, and 0.75 or above considered poor, fair, good, and excellent reliability, respectively (Cicchetti, 1994). Next, an absolute difference score was calculated (TAC_{left} - TAC_{right}) per datapoint (N=836). To test agreement, we used a one-sample t-test studying whether the absolute difference between the aggregated five minute-level TAC values from the left and right ION RAP was significantly different from zero. Next, we used a modified B&A plot to further describe the agreement between the TAC measures from the left and right ION RAP. The plot includes the bias (mean absolute difference), an acceptable limit of agreement, and included a trend line to determine whether there was greater bias at higher TAC. A Pearson's correlation coefficient was added to the trend line to test whether there was indeed a correlation between the mean absolute difference and the average TAC. Due to limited research on the ION RAP, no acceptable limit of agreement currently exists. However, all participants were given 40 grams of alcohol, and we decided to use 20% of the mean peak values (~8 grams of alcohol) as the acceptable limit.

Results

Participant characteristics

Table 1 shows the participant demographics the SCRAM-CAM analysis and ION RAP analysis.

Agreement between left and right SCRAM-CAM TAC values

Testing overall agreement between left and right SCRAM-CAM TAC values. The mean absolute difference between TAC measured by the left and right SCRAM-CAM was 0.010 (SD = 0.011) g/dL , and was significantly different from zero, $t(543)=20.8$, $p<0.001$. The B&A plot shows a bias of 0.010 g/dL and the predefined acceptable limit of agreement of 0.02 g/dL (Figure 1). The plot also reveals an increase in the mean absolute difference at higher TAC readings. Furthermore, 14% of the points fell above the limit that we determined was good agreement.

Testing effects of drinking rate and food consumptions on the agreement between left and right SCRAM-CAM TAC values. Table 2 shows that the effects of food consumption and drinking rate on the difference score aggregated over the days were small and non-significant.

Reliability and Agreement of the ION RAP TAC values.

Reliability of the ION RAP TAC values. Test-retest reliability was found to be excellent with an ICC of 0.81 (see Figure 2 for a scatterplot).

Agreement between the left and right TAC values of the ION RAP. The average peak TAC value over the 6 datasets was 793.58 (SD = 223.05). Using these values, we set the acceptable limit of agreement at 159 nA (20% of the mean peak).

The Bland-Altman shows a mean difference of 131 (SD = 94.4) nA (Figure 3), which was significantly different from zero, $t(154)=17.2$, $p<0.0001$. The plot shows an increase in the mean absolute difference at higher average TAC readings (Figure 3). Furthermore, 30% of the points fell above our acceptable limit.

Discussion

This study had 3 aims: 1) to examine the agreement of the SCRAM-CAM devices, 2) to examine the reliability and agreement of the ION RAP devices, and 3) to assess whether food consumption and drinking rate affect reliability and agreement for the SCRAM-CAM devices.

Agreement of the TAC values as measured by the SCRAM-CAM.

We found that on a 30-minute interval and on the day-level, the mean absolute difference between the TAC measures from the two SCRAM-CAM devices was 0.010 g/dL, with 14% of the points falling above the 0.020 g/dL limit that (in the absence of published thresholds) we assumed was good agreement. Roache *et al.*, (2015) found that using this threshold resulted in not capturing the consumption of several standard drinks. When using the monitors in research, a difference of several standard drinks could result in significantly different conclusions. Further, it was observed that the absolute difference increased with higher TAC values and most outliers were observed above an average TAC value of 0.020 g/dL (Figure 1). This suggests that the monitors' agreement becomes less consistent with higher alcohol consumption, which is important for studies using these monitors to track consumption in populations with high levels of consumption.

Effects of food consumption and drinking rate.

In our previous publication (Egmond *et al.*, 2021), we discussed that the observed discrepancy in TAC data from the left and the right SCRAM-CAM could not be explained by participant's weight (BMI) or sex. The current results extend these findings by showing that this variation is also unlikely to be caused by drinking rate and the consumption of a meal. Other factors that could cause TAC measurement differences between the two ankles include muscle and related blood flow, as a result of right or left leg dominance. However, the differences in TAC measurements are most likely caused by random error such as slight differences of the device placement on the ankle, minor contamination on the device faceplate, air flow between the faceplate and the skin (e.g., when participants are moving), or device failures (e.g., calibration errors). For example, the monitors are strapped around the ankle manually by the researchers, which could result in differences in the tightness of the monitor and thus

variation in the distance between the ankle and the monitor. It is important to note that these results are obtained in a controlled laboratory setting, and it is expected that in naturalistic settings where participants would be walking, sleeping, or exercising, this variation in air flow between the faceplate and the skin would be substantially higher and consequently increase the likelihood of measurement error. In a recent study, researchers tried positioning the SCRAM-CAM on the inside of the calf (higher up the leg than the usual ankle placement), to ensure tightness and limit the monitor's movement (Fairbairn and Kang, 2019). Future research should investigate whether this positioning increases the reliability and agreement of the monitors, although it would presumably decrease comfort and limit walking movement in real-world settings.

Reliability and agreement of the TAC values as measured by the ION RAP.

The ION RAP is a new-generation wristband and research on these TAC monitors is limited, with no published studies on the reliability and agreement to our knowledge. We found an excellent test-retest reliability in terms of ICC values (Cicchetti, 1994); however, as mentioned, a high reliability only indicates the degree of association between the left and right ION RAP TAC values and not their equality, and so this may not indicate a good agreement. Indeed, the mean difference between the left and the right ION RAP readings was 130 (SD = 94.4) nA, which was significantly different from zero. Most differences fell under the predefined acceptable limit of 159 nA (20% of the mean peak), however there was still a substantial proportion (30%) of measurements that were above this limit. So even though we observed an excellent reliability, meaning the ION RAPs accurately capture patterns of drinking over time, the level of agreement was relatively poor, meaning estimated levels of consumption were not especially consistent. Together with the SCRAM-CAM results, this suggests that TAC monitors in general might show less agreement with higher levels of alcohol consumption. Given that the participants in this study consumed only four standard drinks and data collected in the participants' natural environment will often exceed this substantially (Dietze *et al.*, 2014; Norman *et al.*, 2020), future research should investigate whether this decline in agreement at higher TAC readings is consistent over a range of levels of alcohol consumption.

Limitations.

We had a limited sample size available for the test-retest assessment of the ION RAP TAC. As mentioned, the ION RAP TAC data only consisted of data from three days and three participants which limited any further investigation into the effects of sex, BMI, drinking rate, or the consumption of a meal. However, given there has been no work published so far testing the agreement and reliability of these new-generation wristbands, these results provide an important first examination of the TAC data produced by the ION RAP. Due to the limited knowledge on how the ION RAP measures relate to BrAC, it was more difficult to interpret what the reported differences correspond to and what an acceptable level of disagreement would be. Future research comparing different monitors will need to develop a common measure (such as estimated BAC) and benchmarks. Given the small sample size remaining after data cleaning, our results should be considered preliminary in terms of the agreement of the ION RAP TAC data and future research using a bigger sample size and different levels of intoxication are necessary.

Recommendations

We found a relatively poor test-retest reliability regarding the TAC monitors, with this study showing substantial disagreement in the TAC data measuring simultaneously on the left and right ankle or wrist, with lower agreement at higher levels of drinking, and our previous publication showing lower than expected reliability (Egmond *et al.*, 2021). This is in addition to previous research reporting lower sensitivity of the SCRAM-CAMs to lower levels of alcohol consumption (Roache *et al.*, 2015, 2019; Karns-Wright *et al.*, 2018). Thus, the TAC monitors might not be sufficiently reliable when the aim is to measure exact quantities of alcohol consumed. Further limitations of TAC monitors are found in the interpretation of TAC and the conversion to BrAC estimates (Luczak and Rosen, 2014; Dougherty *et al.*, 2015), the delays in the measurement (Karns-Wright *et al.*, 2017), the costs, and its comfort (Caluzzi *et al.*, 2019). However, objective event-level measurement of alcohol consumption remains critically important for research. In certain circumstances where heavy drinking takes place, such as festivals, sport events or general heavy drinking populations, breathalysers may be able to measure exact quantities accurately (Gibb *et al.*, 1984) and reliably (Riordan *et al.*, 2017). Unfortunately, these devices tend to be burdensome to both the researcher and the participant,

especially when trying to capture the same amount of drinking information over time as the TAC monitors can provide. Self-reports, such as EMA surveys, are much cheaper and are suitable when measuring larger populations or when interested in additional behaviours or contexts (Kuntsche and Labhart, 2013, 2014; Wright *et al.*, 2018). However, self-reports have been found to show underreporting (Livingston and Callinan, 2015) and tend to miss data during heavy drinking episodes due to bias in memory and compliance (Wray, Merrill and Monti, 2014; Piasecki, 2019). For these situations, the TAC monitors may still be the method of choice because even though the reliability will be lower, the monitors will still be able to provide objective and continuous measures of drinking over longer periods of time. Thus, depending on the research questions, the strengths and weaknesses of these data collection options will need to be weighed against each other to determine the best methodology or a combination of methodologies in order to get the most comprehensive examination on drinking behaviour.

Conclusion

Our results show that, while ION RAP TAC data showed good reliability, both devices produced substantial levels of disagreement when measuring the same drinking events. This problem was exacerbated at higher levels of alcohol consumption. Transdermal monitors offer an exceptionally objective, non-invasive, and continuous measurement of alcohol consumption, however, according to our results, there is still a large variability in measuring alcohol content. Depending on the research question, researchers need to decide on the most suitable methodology and when using TAC monitors it will be important to determine the acceptable ranges of error in order to rely on the TAC data produced for decisions in clinical, diagnostic, or research settings.

References

Barnett, N. P. *et al.* (2011) 'Contingency management for alcohol use reduction: a pilot study using a transdermal alcohol sensor', *Drug and Alcohol Dependence*, 118(2–3), pp. 391–399.

Barnett, N. P., Meade, E. B. and Glynn, T. R. (2014) 'Predictors of detection of alcohol use episodes using a transdermal alcohol sensor', *Experimental and Clinical Psychopharmacology*, 22(1), pp. 86–96. doi: 10.1037/a0034821.

Berchtold, A. (2016) 'Test–retest: Agreement or reliability?':, *http://dx.doi.org/10.1177/2059799116672875*, 9. doi: 10.1177/2059799116672875.

Bland, M. J. and Altman, D. G. (1986) 'Statistical methods for assessing agreement between two methods of clinical measurement', *The Lancet*, 327(8476), pp. 307–310. doi: 10.1016/S0140-6736(86)90837-8.

Bush, K. *et al.* (1998) 'The AUDIT Alcohol Consumption Questions (AUDIT-C): An Effective Brief Screening Test for Problem Drinking', *Archives of Internal Medicine*, 158(16), pp. 1789–1795. doi: 10.1001/archinte.158.16.1789.

Caluzzi, G. *et al.* (2019) 'Transdermal monitors to assess alcohol consumption in real-time and real-life–a qualitative study on user-experience', *Addiction research & theory*, 27(4), pp. 354–361.

Cicchetti, D. V (1994) 'Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology', *Psychological assessment*, 6(4), p. 284.

Dasgupta, A. (2017) 'Alcohol a double-edged sword: Health benefits with moderate consumption but a health hazard with excess alcohol intake', *Alcohol, Drugs, Genes and the Clinical Laboratory*, pp. 1–21. doi: 10.1016/B978-0-12-805455-0.00001-4.

Dietze, P. M. *et al.* (2014) 'The big night out: What happens on the most recent heavy drinking occasion among young Victorian risky drinkers?', *Drug and Alcohol Review*, 33(4), pp. 346–

353. doi: 10.1111/DAR.12117.

Dougherty, D. M. *et al.* (2015) 'The Potential Clinical Utility of Transdermal Alcohol Monitoring Data to Estimate the Number of Alcoholic Drinks Consumed', *Addictive Disorders & Their Treatment*, 14(3), pp. 124–130. doi: 10.1097/ADT.0000000000000060.

van Egmond, K. *et al.* (2020) 'Wearable Transdermal Alcohol Monitors: A Systematic Review of Detection Validity, and Relationship Between Transdermal and Breath Alcohol Concentration and Influencing Factors', *Alcoholism: Clinical and Experimental Research*, 44(10), pp. 1918–1932.

Egmond, K. van *et al.* (2021) 'A parallel test of the SCRAM-CAM transdermal monitors ensuring reliability', *Drug and Alcohol Review*. doi: 10.1111/DAR.13353.

Fairbairn, C. E. and Kang, D. (2019) 'Temporal Dynamics of Transdermal Alcohol Concentration Measured via New-Generation Wrist-Worn Biosensor', *Alcoholism: Clinical and Experimental Research*, 43(10), pp. 2060–2069. doi: 10.1111/acer.14172.

Gibb, K. A. *et al.* (1984) 'Accuracy and usefulness of a breath alcohol analyzer', *Annals of Emergency Medicine*, 13(7), pp. 516–520. doi: 10.1016/S0196-0644(84)80517-X.

Gullberg, R. G. (2006) 'Estimating the measurement uncertainty in forensic breath-alcohol analysis', *Accreditation and quality assurance*, 11(11), pp. 562–568.

Hill-Kapturczak, N. *et al.* (2014) 'Do variable rates of alcohol drinking alter the ability to use transdermal alcohol monitors to estimate peak breath alcohol and total number of drinks?', *Alcoholism: Clinical and Experimental Research*, 38(10), pp. 2517–2522. doi: http://dx.doi.org/10.1111/acer.12528.

Holt, S. (1981) 'Observations on the relation between alcohol absorption and the rate of gastric emptying.', *Canadian Medical Association Journal*, 124(3), p. 267. Available at: /pmc/articles/PMC1705129/?report=abstract (Accessed: 10 November 2021).

Karns-Wright, T. E. *et al.* (2017) 'Time delays in transdermal alcohol concentrations relative to breath alcohol concentrations', *Alcohol and Alcoholism*, 52(1), pp. 35–41. doi: http://dx.doi.org/10.1093/alcalc/agw058.

Karns-Wright, T. E. *et al.* (2018) 'The correspondence between transdermal alcohol monitoring and daily self-reported alcohol consumption', *Addictive Behaviors*, 85, pp. 147–152. doi: 10.1016/j.addbeh.2018.06.006.

Kessler, R. C. *et al.* (2003) 'Screening for serious mental illness in the general population', *Archives of general psychiatry*, 60(2), pp. 184–189. Available at: https://jamanetwork.com/journals/jamapsychiatry/articlepdf/207204/yoa20567.pdf.

Kuntsche, E. and Labhart, F. (2013) 'Using personal cell phones for ecological momentary assessment: An overview of current developments', *European Psychologist*, 18(1), pp. 3–11. doi: 10.1027/1016-9040/A000127.

Kuntsche, E. and Labhart, F. (2014) 'The future is now--using personal cellphones to gather data on substance use and related factors', *Addiction*, 109(7), pp. 1052–1053. doi: 10.1111/ADD.12540.

Lansdorp, B. *et al.* (2019) 'Wearable enzymatic alcohol biosensor', *Sensors*, 19(10), p. 2380.

Leffingwell, T. R. *et al.* (2013) 'Continuous objective monitoring of alcohol use: twenty-first century measurement using transdermal sensors', *Alcoholism: Clinical and Experimental Research*, 37(1), pp. 16–22. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3623856/pdf/nihms446008.pdf.

Livingston, M. and Callinan, S. (2015) 'Underreporting in alcohol surveys: whose drinking is underestimated?', *Journal of studies on alcohol and drugs*, 76(1), pp. 158–164.

Luczak, S. E. and Rosen, I. G. (2014) 'Estimating BrAC from Transdermal Alcohol Concentration Data Using the BrAC Estimator Software Program', *Alcoholism: Clinical and*

Experimental Research, 38(8), pp. 2243–2252. doi: 10.1111/acer.12478.

Norman, T. *et al.* (2020) 'Combining transdermal and breath alcohol assessments, real-time drink logs and retrospective self-reports to measure alcohol consumption and intoxication across a multi-day music festival', *Drug and Alcohol Review*. doi: 10.1111/DAR.13215.

Piasecki, T. M. (2019) 'Assessment of alcohol use in the natural environment', *Alcoholism: Clinical and Experimental Research*, 43(4), pp. 564–577. Available at: https://onlinelibrary.wiley.com/doi/pdfdirect/10.1111/acer.13975?download=true.

Riordan, B. C. *et al.* (2017) 'The accuracy and promise of personal breathalysers for research: Steps toward a cost-effective reliable measure of alcohol intoxication?', *Digital health*, 3, p. 2055207617746752. Available at: https://journals.sagepub.com/doi/pdf/10.1177/2055207617746752.

Roache, J. D. *et al.* (2015) 'Using transdermal alcohol monitoring to detect low-level drinking', *Alcoholism: Clinical and Experimental Research*, 39(7), pp. 1120–1127. doi: http://dx.doi.org/10.1111/acer.12750.

Roache, J. D. *et al.* (2019) 'Processing Transdermal Alcohol Concentration (TAC) data to Detect Low-Level Drinking', *Alcohol*, 01, p. 1. doi: https://dx.doi.org/10.1016/j.alcohol.2018.08.014.

Saldich, E. B. *et al.* (2021) 'Effects of stomach content on the breath alcohol concentration-transdermal alcohol concentration relationship', *Drug and Alcohol Review*. doi: 10.1111/DAR.13267.

Sorbello, J. G. *et al.* (2018) 'Fuel-cell breathalyser use for field research on alcohol intoxication: an independent psychometric evaluation', *PeerJ*, 6, p. e4418. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5857179/pdf/peerj-06-4418.pdf.

Swift, R. (2003) 'Direct measurement of alcohol and its metabolites', *Addiction*, 98(SUPPL. 2), pp. 73–80. doi: 10.1046/J.1359-6357.2003.00605.X.

Wray, T. B., Merrill, J. E. and Monti, P. M. (2014) 'Using ecological momentary assessment (EMA) to assess situation-level predictors of alcohol use and alcohol-related consequences', *Alcohol*

research: current reviews, 36(1), p. 19.

Wright, C. *et al.* (2018) 'Mobile phone-based ecological momentary intervention to reduce young adults' alcohol use in the event: a three-armed randomized controlled trial', *JMIR mHealth and uHealth*, 6(7), p. e149.

Wright, C. J. C. *et al.* (2021) 'Small, slim, sleek, and familiar: user experiences with an ION wearable research alpha prototype transdermal alcohol monitor', *https://doi.org/10.1080/16066359.2021.1982911*. doi: 10.1080/16066359.2021.1982911.

Table 2. Participant demographics

	SCRAM-CAM (N=21)		*ION RAP (N=3)*	
	M (SD)	Range	M (SD)	Range
Sex (% Female)	52.4%		33.3%	
Age (years)	27.0 (5.4)	18 - 35	24.0 (5.3)	18 - 28
% Caucasian	71.4%		100%	
BMI	24.4 (5.4)	18.8-32.3	26.6 (3.2)	23.0 – 29.0
AUDIT Score	6.2 (2.2)	1.0-10.0	6.7 (2.1)	5.0 - 9.0

Note. Shown are means (M) and standard deviations (SD) in brackets. BMI = Body Mass Index, AUDIT = Alcohol Use Disorders Identification Test.

Table 2. Absolute left-right difference in TAC readings regressed on gender and drinking rate.

Parameter	B (SE)	P-Value
Intercept	0.013 (0.003)	0.000
Food consumption	-0.004 (0.003)	0.292
Drinking rate	-0.005 (0.003)	0.143

Note. Shown are unstandardized regression coefficients B with standard errors (SE) in brackets. Food consumption was coded as 0 = no meal, 1 = meal; drinking rate was coded as 0 = four standard drinks within sixteen minutes, 1 = four standard drinks within one hour.

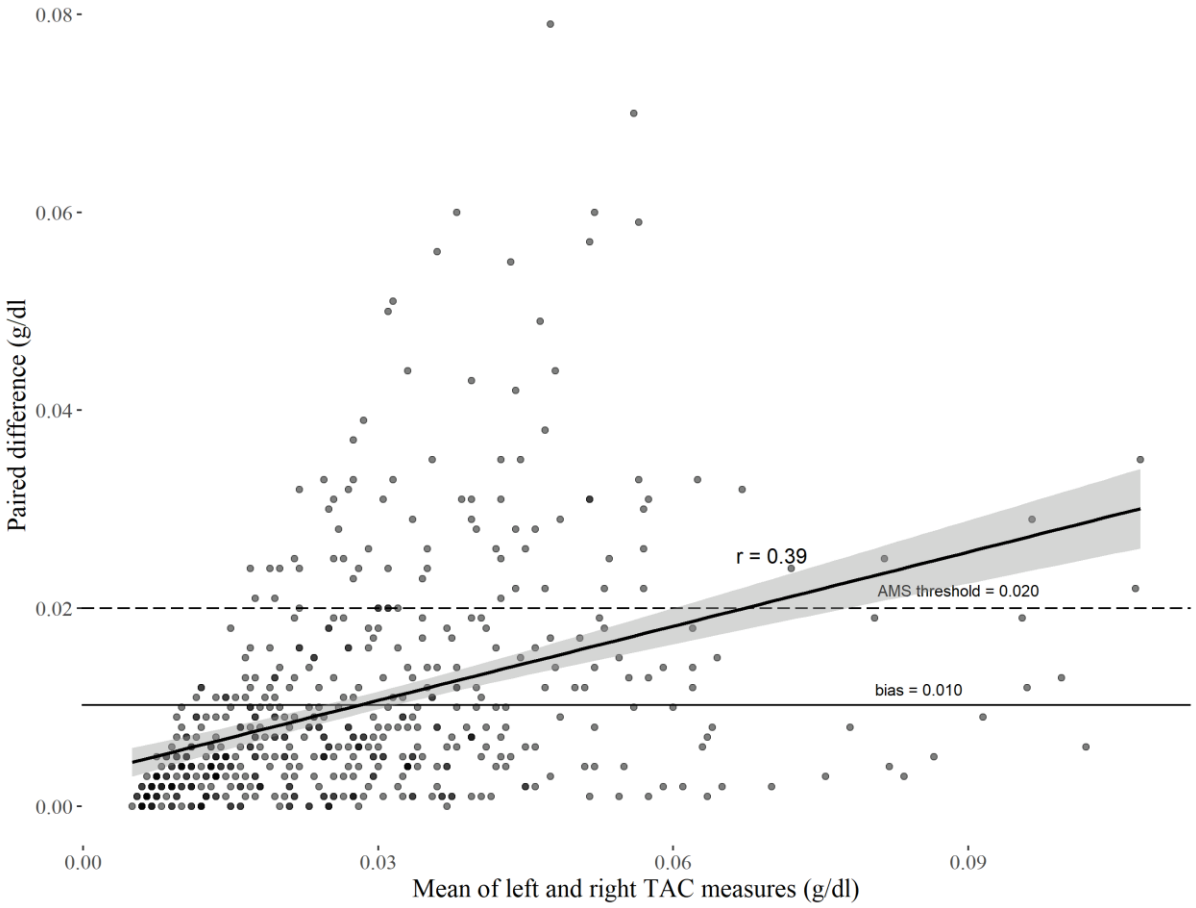

Figure 2. Bland and Altman plot of the paired difference between left and right on the mean values of left and right TAC measures as measured by the SCRAM-CAM.

Note. The plot shows a bias (overall mean absolute difference) of 0.010 g/dL, the predefined acceptable limit of agreement of 0.020 g/dL, a trend line indicating that bias increased as TAC increased (r = .0.39), and the standard error depicted as the shaded region of the line.

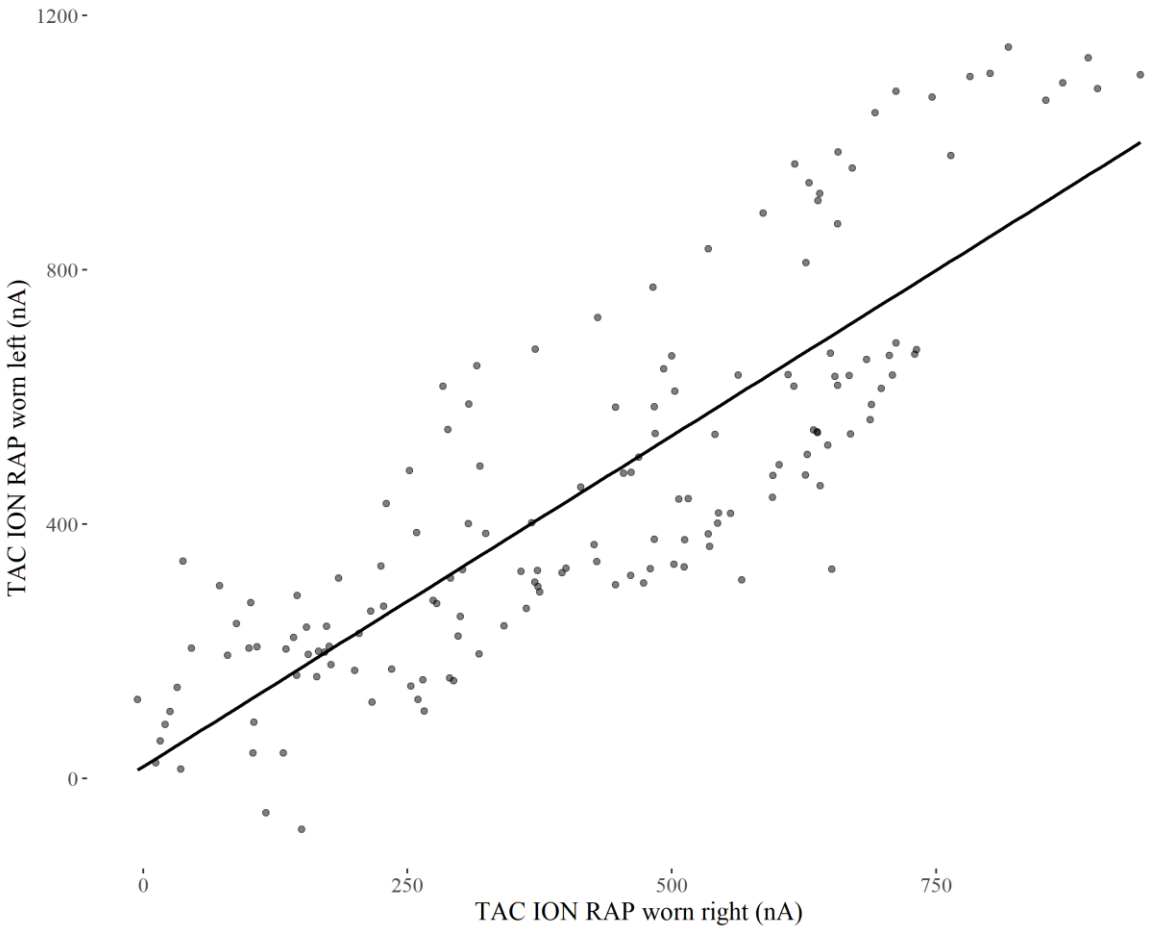

Figure 2. Scatterplot of ION RAP TAC measures from left on right.

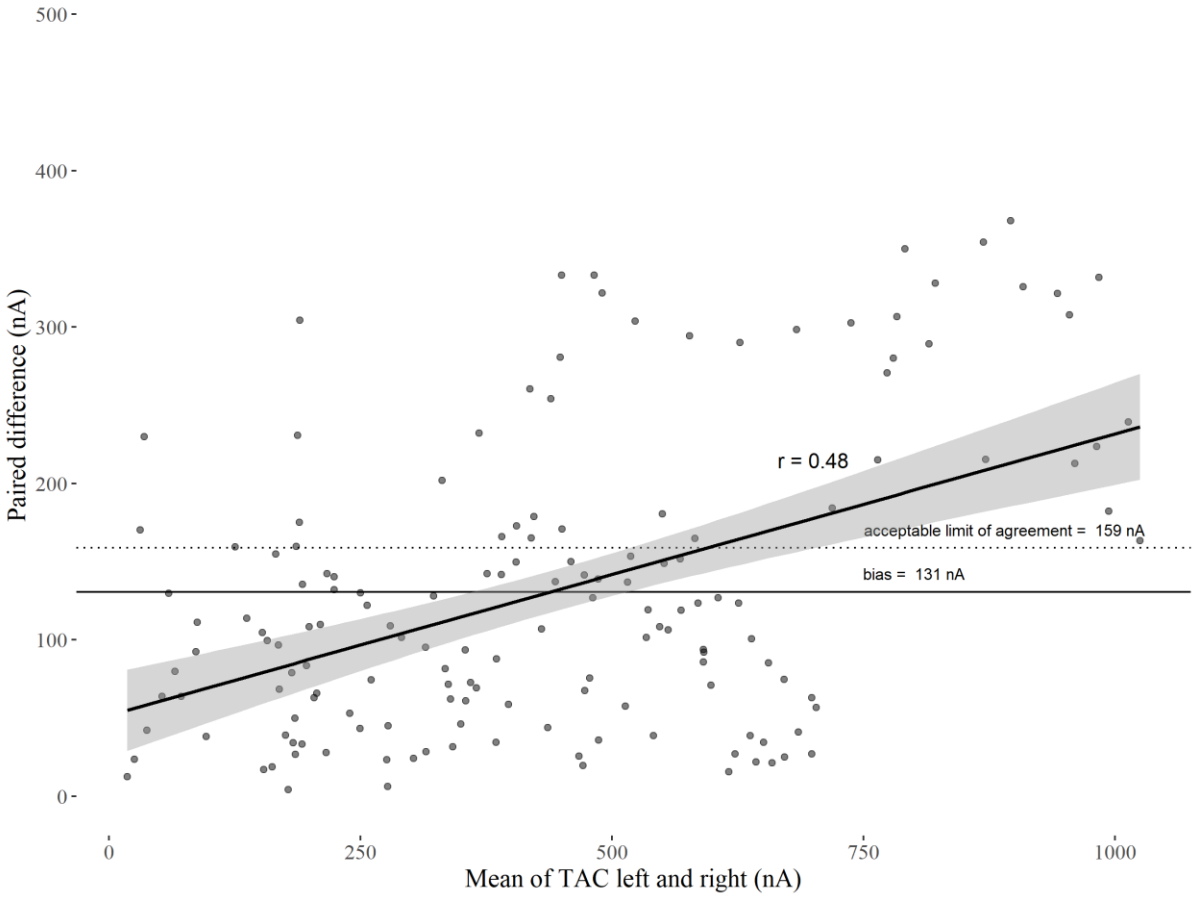

Figure 3. Bland and Altman plot of the paired absolute difference between left and right on the mean values of left and right TAC as measured by the ION RAP.

Note. The plot shows a bias (overall mean absolute difference) of 131 nA, the acceptable limit of agreement (159 nA), and a trend line indicating that bias increased as TAC increased (r = .0.48), and the standard error depicted as the shaded region of the line.

6. Comparing data from the ION RAP and SCRAM-CAM against breath alcohol concentration readings

Chapter 6 contains the following manuscript:

van Egmond, K., Riordan, B., JC Wright, C., Livingston, M., Kuntsche, E. (submitted for publication). Measurement of transdermal alcohol concentration using a wrist worn enzymatic transdermal monitor.

Author contributions

KvE and CW collected the data; KvE analysed the data and led data interpretation and manuscript writing. BR provided support for data analysis. EK, CW, BR and ML provided critical feedback, suggested revisions, and refined the study design and analytical approach.

Measurement of Transdermal Alcohol Concentration using a Wrist Worn Enzymatic Transdermal Monitor.

Kelly van Egmond[1] MSc, Benjamin Riordan[1] PhD, Cassandra J.C Wright[1,2,3] PhD, Michael Livingston[1,5,6] PhD, and Emmanuel Kuntsche[1,4] PhD.

7. Centre for Alcohol Policy and Research (CAPR), NR1 Building, La Trobe University, Melbourne, 3086 VIC, Australia.
8. Menzies School of Health Research, Charles Darwin University, Darwin, Australia
9. Burnet Institute, Melbourne, Australia.
10. Institute of Psychology, Eötvös Loránd University, Kazinczy u. 23-27, 1075 Budapest, Hungary
11. Faculty of Health Sciences, Curtin University, Melbourne, Australia
12. Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

*Corresponding author:

Kelly van Egmond

Centre of Alcohol Policy Research

School of Psychology and Public Health,

La Trobe University

3086 Bundoora, VIC, Australia

e-mail: k.vanegmond@latrobe.edu.au, phone: +61 04 98257320

Declarations of interest: none

Abstract

Background: With improvements in size, comfort, and sampling frequency, new-generation transdermal monitors such as the ION Research Alpha Prototypes (ION RAP) hold promise for real-time alcohol measurement. This paper aims to provide the first comparisons of the wrist worn enzyme-based ION RAP and the fuel-cell based SCRAM-CAM against breath alcohol concentration (BrAC) readings.

Methods: Participants (N=17) completed a total of 25 laboratory alcohol administration sessions, while wearing both a prototype of the ION RAP wristband and a SCRAM-CAM ankle monitor; they also gave breath samples each 10 minutes. Analysis focused on latencies of transdermal alcohol concentration (TAC) after alcohol ingestion, correlations, and cross-correlations between BrAC and TAC measurements.

Results: A high failure rate of the ION RAP was observed (61.5% of the sessions were removed due to the sessions not containing enough valid data). On average, the SCRAM-CAM and ION RAP detected alcohol 43 (SD=21) and 50 (SD=27) minutes after the first drink, with peak values reached after 138 (SD=47) and 154 (SD=56) minutes, respectively. SCRAM-CAM TAC peak (r=0.185p=0.375) and area under the curve (AUC; r=0.320, p=0.118) showed small- and medium-sized correlations to BrAC. ION RAP TAC peak (r=-0.082, p=0.698) and AUC (r=0.040, p=0.852) correlations to BrAC were close to zero.

Conclusions: In this study, the new-generation ION RAP and the traditionally used SCRAM-CAM show similar delays and TAC measurement patterns over time, despite using either enzyme or fuel-cell based technologies, respectively. Due to high failure rates of the ION RAP prototypes and close to zero correlations to BrAC, further developments and improvements of these TAC wristbands are required for reliable and valid use in real-time alcohol measurement.

Keywords: Transdermal alcohol monitors, validation, breath alcohol concentration, biosensor.

Highlights

- First evaluation of ION wristband and SCRAM-CAM against breath alcohol concentration
- The failure rate for the ION RAP was 62% and 0% for the SCRAM-CAM
- The enzyme-based and fuel-cell devices had similar alcohol detection delays
- BrAC correlations higher for SCRAM-CAM (r=0.19-0.32) than for ION RAP (r=-0.09-0.04).
- Maximal cross-correlation coefficients ranged from 0.63 to 0.67.

Introduction

Transdermal monitors, such as the Secure Continuous Remote Alcohol Monitor Continuous Alcohol Monitoring (SCRAM-CAM^{TM}, Alcohol Monitoring Systems, AMS, Littleton, USA), are predominantly used in justice systems for detecting alcohol consumption among people with orders that include conditions of abstinence (Bock, 2003; Alcohol Monitoring Services, 2021). Transdermal monitors are wearable monitors that measure the estimated 1% of alcohol consumed that is diffused transdermally (in passive sweat or actively through sweat glands secreted through the skin). The SCRAM-CAM is worn around the ankle and measures transdermal alcohol concentration (TAC) using a fuel-cell sensor; it can provide detailed information on an individual's drinking patterns by measuring TAC continuously. To evaluate how well a monitor measures the presence of alcohol, two concepts are tested: test-retest reliability and validity. Test-retest reliability is the capacity of the monitor to reproduce a similar measurement and measurement pattern when applied multiple times under similar conditions (Berchtold, 2016; Heale & Twycross, 2015; Sobell et al., 1986) and the validity refers to the accuracy of the measure. Validity is often measured by testing the correlation of the TAC measurement to another way of measuring alcohol consumption (e.g., breath/blood alcohol concentration (BrAC/BAC). Previous research has established that there is a strong correlation between TAC and BrAC/BAC (Leffingwell et al., 2013; van Egmond et al., 2020) and self-reported alcohol use (Dougherty et al., 2012; Fairbairn & Kang, 2019; Hill-Kapturczak et al., 2014). A recent laboratory study including 30 participants, with five controls receiving no alcohol and 25 participants who received an alcohol dose intended to bring them up to a BrAC level of 0.08%, observed a strong positive correlation of r=0.56-0.60, between BrAC (peak BrAC and AUC) and TAC (peak TAC and Area Under the Curve; AUC) (Fairbairn & Kang, 2019).

As the SCRAM-CAMs were originally developed to monitor convicted offenders on compliance orders, robustness and tamper-proofing appear to have been prioritised over user-experience. Previous research reported discomfort, especially within the first few days, and during exercise and sleep (Barnett et al., 2011; Caluzzi et al., 2019). Furthermore, given the stigma

associated with wearing a device known to be associated with convicted offenders, some participants reported public embarrassment (Barnett et al., 2011; Marques & McKnight, 2009).

Using the AMS criteria for detection, the SCRAM-CAM is was found to be limited in its ability to detect low levels of alcohol consumption (Roache et al., 2015). In the justice system, the SCRAM-CAM is primarily used to provide a dichotomous measure of whether alcohol has been consumed or not. To keep the false positives low, the manufacturers use conservative criteria when confirming alcohol consumption, resulting in the detection of heavy drinking events, but missing most low to moderate drinking events (Roache et al., 2015). Further, another important challenge the SCRAM-CAM poses is the latency to first detected alcohol and to peak compared to BrAC measures (Leffingwell et al., 2013). This occurs due to the ethanol transportation through the skin that is physiologically more complex when compared to the more robust measurement of ethanol in either breath or blood. Specifically, while breathalysers detect alcohol almost immediately after consumption, the SCRAM-CAMs detect alcohol in around 23 minutes after consumption (Fairbairn & Kang, 2019). Greater delays are seen in the time to reach peak levels; one study showed BrAC reached peak levels within 77 minutes, and SCRAM-CAM within 3.3 hours (Fairbairn & Kang, 2019) and another study observed a lag of 4.5 hours (Karns-Wright et al., 2017). However, as Fairbairn and Kang (2019) pointed out, these latencies are based on only one point on the BAC curve and do not examine the latencies across the entire drinking sessions, encompassing the full TAC and BrAC curves. When examining the latencies along the full curves SCRAM-CAM TAC was about 69 minutes behind BrAC (Fairbairn & Kang, 2019). Given that researchers are aiming to use transdermal devices for real-time alcohol measurement, it is important to know the exact latencies to be able to map alcohol use over time.

New-generation wristbands have been developed using advances in technology, resulting in a significant reduction in size improving user experience and increasing sampling frequency. These wristbands include the BACtrack Skyn™ (BACtrack Breathalyzers/KHN Solutions Inc., San Francisco, USA) and the ION RAP™ (ION Wearable Inc™ , Santa Barbara, USA). One study showed that Skyn TAC reached peak values over an hour prior to SCRAM-CAM TAC (Fairbairn &

Kang, 2019). Further, when examining the full curves, the Skyn was found to be twice as fast in its alcohol measurement as compared to the SCRAM-CAM (Fairbairn & Kang, 2019). The mechanisms behind the difference in latencies that were observed remain unaddressed. A possible explanation could be related to the body positioning, i.e., the wrist as compared to the ankle. Relative distribution of sweat glands and permeability of the skin might differ across these body parts resulting in differences in alcohol detection latency. Moreover, correlations between peak and AUC values from Skyn TAC and BrAC were strong and significant (Fairbairn & Kang, 2019).

Unlike the fuel-cell technology used in the Skyns and the SCRAM-CAMs, the ION RAP wristbands employ an enzymatic detection pathway, with TAC being measured in terms of a raw current (Lansdorp et al., 2019). Using the enzyme, alcohol oxidase is said to detect not just heavy drinking events, but also low to moderate drinking (Lansdorp et al., 2019). With a higher sensitivity to lower alcohol levels, it should be expected to observe smaller delays in the time to first detected alcohol. The ION RAP is a small device worn around the wrist, with a rechargeable battery and Bluetooth communication to a smartphone application. They use a disposable cartridge system that can be used for up to 24 hours before they need to be replaced with a new cartridge. Using this disposable cartridge system, the ION RAPs are able to overcome the problem of sensor fouling and degradation as is observed for the fuel-cell sensors which should in turn increase the reliability of the TAC data (Allan et al., 2017; Campbell et al., 2018; Lansdorp et al., 2019). ION Wearable researchers studied an early prototype of the ION RAP wristband and concluded that the wristband was able to capture real-world drinking events with time to first detected alcohol ranging between 70-90 minutes (Lansdorp et al., 2019). It was observed that the ION RAP TAC slightly underestimated a theoretical peak BAC as calculated by the authors (Lansdorp et al., 2019). However, research on the ION RAP wristband is limited and no validation study comparing the ION RAP TAC to either SCRAM-CAM TAC and/or BrAC has been published to date.

The aim of the current study is to investigate and compare the SCRAM-CAM TAC and ION RAP TAC to BrAC, when obtained in a laboratory alcohol session under controlled settings. Specifically, this study aims to test whether TAC, measured through an enzymatic detection pathway

(ION RAP), shows significantly decreased latencies and improved correlations to BrAC as compared to the SCRAM-CAM fuel-cell measured TAC.

Materials and Methods

Participants

Participants were recruited using targeted Facebook advertisements and street intercept approaches. Participants were screened to determine basic eligibility and to provide informed consent to be contacted for further screening. The inclusion criteria were: aged between 18-35; consumption of ≥2 standard drinks during one single occasion in the preceding month; and able to read and understand English. Exclusion criteria were: current diagnosed mental health condition, or high psychological distress (≥25 on the Kessler Psychological Distress Scale; K10 (Kessler et al., 2003)); high-risk drinking behaviour (≥10 Alcohol Use Disorders Identification Test–Alcohol Consumption Questions; AUDIT-C, (Bush et al., 1998)); being underweight or obese (BMI outside the healthy range: 18-30); daily smoking; use of medication proven to interact with alcohol; diagnosis of diabetes mellitus, liver disease, kidney disease or sleep disorder; current infection of any kind; or a metal allergy. Participants with a current or possible pregnancy or who were breastfeeding were also excluded. The study was approved by the Human Research Ethics Committee at La Trobe University (HEC19249).

Study procedure

Participants were invited to three laboratory sessions during days convenient to them. Written consent was collected from every participant prior to the start of the first day. Participants were instructed to not drink any alcohol at least 12 hours prior to attending the session and fast for at least 3 hours prior to arrival. Laboratory sessions commenced between 9.00-10.00am. Two hours prior to the start of every session, the cartridges for the ION RAP were activated and inserted in the wristband by trained researchers as instructed by ION Wearable (See

Figure B for cartridge). Instructions for cartridge activation included extracting gel from the pink tube using the straw provided and spreading it onto the electrode surface of the cartridge making sure not to spill any gel on the area outside the electrode. The cartridge was then closed and inserted into the wristband. Participants were fitted on arrival with the ION RAP and SCRAM-CAM monitors

at least an hour prior to drinking to allow for calibration. When enough devices were available and functioning, participants wore an ION RAP wristband on each wrist simultaneously, however due to a limited number of wristbands available, most participants only wore one monitor. One SCRAM-CAM was fitted around both participant's ankles. At the start of each session, participants were breathalysed to ensure they had not consumed any alcohol prior to the study. It was observed that the fixed dosing procedures by Fairbairn and Kang, (2019), resulted in a restricted range of BrAC peak levels. To be able to study the associations between TAC and BrAC over a wide range of peak levels, we asked all participants to consume four standard drinks (total of 40 grams of alcohol), each drink contained 32 millilitres of vodka (40% alcohol) mixed with 68 millilitres of soda and sugar-free cordial. To control for any effects of temperature and movement on the TAC detection, the laboratory room temperature was set to 22 degrees Celsius using a central heating/cooling system and the participants were instructed to remain seated during the session. Following administration, participants remained in the laboratory for 2-3 hours after their BrAC levels reached 0.000; on average, participants stayed in the laboratory between 9.00am-6.00pm, up to 9 hours in total. At the end of each participant's session, both the ION RAP and SCRAM-CAM monitors were removed.

Measures

Breath alcohol monitoring

Breath samples were taken every 10 minutes after finishing the drinks using an Andatech Prodigy–S (Andatech Inc., Nunawading, VIC, AU; an Australian standard certified breathalyser). The Andatech Prodigy – S uses the Andatech FXCell3 advanced fuel-cell sensor reacting specifically to alcohol. The monitors were calibrated when received and had a feature keeping track of the calibration statistics. Before every test, participants were given a controlled amount of water to rinse their mouths, and the breathalyser would take a clean air sample to ensure the mouthpiece and monitor were free of alcohol residues. For every test, a new disposable mouthpiece was used to avoid testing alcohol residuals.

Transdermal alcohol concentration monitoring

SCRAM-CAM

A sample of 20 SCRAM-CAM monitors were used randomly. Each session the researchers would note which monitor was worn on which ankle on the participant. The SCRAM-CAMs sample TAC every 30 minutes and were activated before use, using SCRAM Direct Connect™, connecting the monitor to the online monitoring software: SCRAM Optix™ (See Figure C). After each session, the devices were cleaned with alcohol-free wipes and every three to four sessions the face plate (the metal filter on the device touching the skin) was replaced.

ION RAP

The 8 ION RAP wristbands available were pre-production prototypes (see Figure A) and used randomly, with the researchers recording the wristband used in each session to keep track of possible faulty monitors. ION RAP electrical current data start at 1000 nA, this current slowly returns to baseline when the cartridge is inserted and with skin contact. As per instructions from MILO Sensors, the ideal baseline would be 50 nA or lower and should be reached after one hour of calibration. To allow for as much time as possible, trained researchers developed and inserted the cartridges at least one hour prior to the arrival of the participants. Participants were then fitted with the ION RAP an hour prior to drinking to allow for further calibration, so there were at least 2 hours in total for calibration. Despite this lengthy calibration time, only one session showed a baseline of 50 nA or lower. Due to the limited time we could keep participants in the laboratory, we could not afford to have more time for calibration purposes. When alcohol is detected, the current will start to increase, for this reason, the first point of detection was operationalized as the first point of increase after alcohol administration. The ION RAP sensor samples TAC measures every five seconds or more frequently when connected with the application on a smartphone or tablet through Bluetooth. For hygiene purposes, the ION RAP wristbands were wiped clean after each session, however given that the ION RAP cartridges had to be replaced every day, there were no deep cleaning measures needed. The monitors had to be charged overnight.

Data processing and analysis

Out of the 69 laboratory sessions we conducted with a total of 23 participants, 11 (15.9%) sessions were excluded due to incorrect calibration of the SCRAM-CAMs (i.e., not enough time had been allowed for prior to alcohol administration). Another 33 sessions were excluded because the ION RAP TAC data was of insufficient quality (See 'Results: Failure rates' for details). Due to this data exclusion, our final sample included, including 25 full-data sessions incorporating full data (including BrAC, SCRAM-CAM TAC, and ION RAP TAC) from 17 participants. A visual representation on the data exclusion procedures is shown in the supplementary materials. These 25 sessions did not have any missing or invalid data points for either the SCRAM-CAM or ION RAP TAC. With the aim to directly compare the TAC data from the ION RAP to the SCRAM-CAM and the breathalysers and previous results showing that the TAC data from the SCRAM-CAM monitors worn on opposite ankles can differ (Egmond et al., 2021), we decided to only compare the monitors on the same side of the body. For example, we compared ION RAP TAC data worn on the left wrist with the SCRAM-CAM TAC data worn on the left ankle. Unfortunately, due to a limited number of wristbands available and functioning, some participants ended up wearing only one ION RAP wristband. In these occasions, we matched data based on the side of the body that the ION RAP was worn on. However, there was still a small subsample that wore an ION RAP on each wrist, while also wearing a SCRAM-CAM monitor on each ankle simultaneously. In these occasions we analysed the matched TAC data from both the left and right for that participant and session.

To analyse the latencies of the TAC values (SCRAM-CAM and ION RAP) as compared to the start of drinking and BrAC values, three metrics were used: 1) the time to first detected alcohol, which is the time elapsed between the start of drinking to the time of the first non-zero TAC value, 2) the time to reach peak value, which is the time passed between the start of drinking and the time of peak TAC, and 3) latency to maximal cross-correlation across BrAC and TAC curves, assessing the similarities between the BrAC and TAC as a function of time[1]. For the cross-correlation analysis,

[1] Cross-correlation analyses produce the correlations between the two curves at different time "lags" and can thus be used to estimate the level of displacement between the two curves.

BrAC and SCRAM-CAM TAC values were interpolated, and ION RAP TAC values aggregated to minute-level estimates. Due to disturbances in the ION RAP signal, three sessions were cut short. Cross-correlation analysis were conducted for each of the 25 sessions separately, and the latency to maximal cross-correlation between the BrAC, and TAC curve was recorded. To test the differences in latencies between the BrAC and TAC (SCRAM-CAM and ION RAP) values, we used a paired samples t-test. To test the size of these differences we calculated the Cohen's d with 0.2, 0.5, and 0.8 indicating small, medium, and large effects, respectively (Cohen, 2013).

We further analysed the correlation between BrAC and TAC using: 1) the peak values; the highest BrAC and TAC values after drinking; 2) area under the TAC curve (AUC), calculated using the trapezoidal formula (Yeh & Kwan, 1978); and 3) maximal value of the cross-correlation coefficients between BrAC and TAC. To calculate both peak and AUC values, the ION RAP TAC data was centred and standardized by subtracting the start values from all other readings (Fairbairn & Kang, 2019). Pearson's correlation coefficients were used to analyse the relationships between the BrAC and TAC (SCRAM-CAM and ION RAP) peak and AUC values. Scatterplots were used to visualize the relationship between these measures. A correlation of .10, .30, .50, and .70, was considered to indicate a small, medium, large, and very large effect (Cohen, 2013; Maher et al., 2013). Maximum values of the cross-correlation coefficients between BrAC and TAC for SCRAM-CAM and ION RAP were calculated and paired samples t-test were used to test for significant differences. Again, to test the size of these differences, we calculated the Cohen's d (Cohen, 2013).

There is limited validation literature available on the ION RAP, with only one study validating a new-generation TAC wristband (Fairbairn & Kang, 2019). Given that the TAC data measured by ION RAP poses similar challenges as the TAC data measured by the Skyn, we used similar to methods to the validation study of the Skyn (Fairbairn & Kang, 2019). Alternative analysis methods were suggested by ION Wearable which are discussed and shown in the supplementary materials.

Results

Participants and descriptive data

Table 1 describes our final sample (N=17) and the sessions included (N=25).

Figure 1 shows visualisations of the raw BrAC and TAC data from all sessions included in the analysis. Here it is observed that both TAC and BrAC curves are similar in their alcohol consumption patterns, ascending with alcohol absorption and descending with alcohol elimination.

Failure rates

A total of 65 TAC sessions with the ION RAP wristbands were collected. During 40 sessions, the ION RAP either did not detect any alcohol or the detection was disturbed (possibly due to Bluetooth connection losses), resulting in indiscernible TAC curves. Out of the 40 sessions, three devices were deemed faulty, one due to charging problems and the other two were unable to measure TAC in three sessions (one session for each device, 7.5% of the excluded sessions). Once found to be faulty, the devices were excluded from future sessions. However, the cause of the indiscernible alcohol curve for the remaining 37 sessions is unknown. This resulted in a failure rate of 61.5% for the ION RAP wristband. Besides the incorrect calibration, there were no actual device failures for the SCRAM-CAM. A visual representation on the data exclusion procedures and examples of ION RAP TAC data are shown in the supplementary materials.

Latency in alcohol measurement

SCRAM-CAM was on average about seven minutes faster in detecting alcohol as compared to the ION RAP. However, this difference was relatively small and statistically non-significant (Table 2). Regarding time to peak, BrAC reached peak values after 48.1 (SD=25.6) minutes. The SCRAM-CAM TAC reached peak values after an average of 89.6 minutes after peak BrAC and ION RAP readings after an average of 105.9 minutes. The difference in latencies between peak SCRAM-CAM and peak ION RAP was small and non-significant (Table 2). The ION RAP TAC reached maximal cross-correlation with BrAC about 22 minutes faster than the SCRAM-CAM TAC, this difference was found to be close to medium-sized (Table 2).

Associations between BrAC and TAC measures

Peak

The relationship between peak SCRAM-CAM TAC and BrAC values was positive and small, r=0.185, p=0.375 (See Figure 3A). The relationship between peak ION RAP TAC and BrAC was slightly negative, but also considered small, , r=-0.082, p=0.698 (See Figure 3B).

AUC

Figure 4A shows a medium and positive correlation between the SCRAM-CAM TAC and the BrAC AUC values (r=0.320, p=0.118). Figure 4B shows a small and close to zero correlation between the ION RAP TAC and the BrAC AUC values (r=0.040, p=0.852).

Maximal cross-correlation

The average maximal cross-correlation between BrAC and SCRAM-CAM TAC was found to be large at 0.63 (SD=0.15) and the average maximal cross-correlation between BrAC and ION RAP TAC was also found to be large at 0.67 (SD=0.17). The difference between these correlations were small, M_{diff}=-0.04 (SD=0.14), t(28)=-1.19, d=0.25.

Discussion

The current study is the first to systematically validate the ION RAP; a new-generation, wrist-worn device, utilising an enzymatic detection pathway. An important observation was that the ION RAP had a rather high failure rate of 61.5%. This limited the sample size and, after processing the ION RAP TAC data, only 25 sessions were left. For the failure of the three devices, it remains unclear whether this is a device or human error. As the manual activation of the cartridges is prone to human error resulting in device failure, some of the failures may have occurred when the researchers activated the cartridges (see

Figure B). Since it was difficult to extract the gel with the straw the ION Wearable company provided spilling of the gel, or an insufficient amount of gel placed on the electrode could result in a faulty sensor. Even though this was standard procedure at the time, it is probable errors happened at this stage. Since the time that this work was completed in 2019, the ION Wearable team have released updates to their products which may address some of the shortcomings of the ION RAP that we encountered.

Due to the complex nature of ethanol transportation through the skin, significant delays in the latency to first detected alcohol and to reach peak for TAC as compared to BrAC measures have repeatedly been reported (Fairbairn & Kang, 2019; Lansdorp et al., 2019; Leffingwell et al., 2013; Marques & McKnight, 2009; van Egmond et al., 2020). This study also observed substantial latencies in TAC alcohol detection, with alcohol being detected 40-50 minutes after ingestion and reaching peak values after 2.5-3 hours. However, as Fairbairn and Kang (2019), pointed out, these latencies are based on only one point on the BAC curve. When examining latencies across the full curves, the latency between TAC and BrAC was approximately 48-69 minutes. Even though SCRAM-CAM was observed to be faster in alcohol detection measurements as compared to the ION RAP, these differences were statistically non-significant. Fairbairn and Kang (2019), observed significantly shorter latencies for the Skyn as compared to the SCRAM-CAM and suggested that this might be related to the body positioning of the monitors (wrist vs ankle). However, the current study did not find any significant differences in the latencies between the wrist- and ankle-worn monitors. This observation suggests that independent of the body positioning and alcohol detection pathway used (enzymatic or fuel-cell), transdermal monitors show a similar delay of alcohol measurements that is most likely solely caused by the ethanol transportation through the skin. However, further research will be necessary to make any firmer conclusions. It is, according to our findings, important that this lag is accounted for in the analysis approach when comparing TAC to BrAC or for the interpretation of TAC when aiming to use these monitors in naturalistic settings. Adding to this, the maximal cross-correlation coefficient ranging from 0.63 to 0.67 respectively, suggests that both the SCRAM-CAM and ION RAP TAC curves show very similar patterns over time as compared to the BrAC curves. However, when extracting the peak measures, the TAC data from both the SCRAM-CAM and ION RAP show only small-sized correlations to the BrAC data measures. To note, the correlation between SCRAM-CAM TAC and BrAC AUC was positive and a medium-sized correlation whereas the correlation for the ION RAP TAC was negative and close to zero. This suggests that together with the results on the failure rates, the SCRAM-CAM monitors are the more reliable and valid monitors, a conclusion also drawn by Fairbairn & Kang, (2019), when testing the SCRAM-CAM against the

BACtrack Skyn. It is again important to keep in mind that the ION RAPs in this study were prototypes and the newer models might also show stronger correlations to BrAC.

One notable difference between our study and previous studies was that previous studies found stronger correlations between the SCRAM-CAM TAC data and BrAC. However, this may be due to the fact that previous studies used a non-alcohol receiving control group and restricted weight-based dosing protocols in which participants received a dose intended to bring them up to the US legal driving limit (0.08%) (Fairbairn & Kang, 2019; Sakai et al., 2006). For example, using a non-alcohol receiving control group tends to increase the size of the correlation between TAC and BrAC as both values were zero or close to zero. Further, a limited variation in BrAC produced by these restricted weight-based alcohol administration protocols makes these procedures less suitable to quantify the magnitude of BrAC-TAC correlations. The current study had participants drinking four standard drinks, independent of their gender, weight, and height. This resulted in different peak measures for all participants and can thus result in wider ranges of variance between the peak BrAC and TAC measures. Future studies could try and estimate BrAC using the TAC data and test what factors are influencing both TAC detection and the TAC to BrAC relationship (including gender, BMI and temperature) (See also Figure S1 and Lansdorp et al., 2019).

Limitations

One of the main limitations of this study is that the sample size ended up being smaller than expected, which as explained, is partly related to errors in calibration of the SCRAM-CAM and the high ION RAP failure rate. For this reason, concerning the correlation analysis, we decided to place more importance on the effects sizes rather than the significance levels. Further, this study did not include a control experiment at zero alcohol consumption. Therefore, instead of extrapolating a baseline curve, the current study used the first point of detection (or first point of change) and subtracted this value from all other readings to create a baseline (Fairbairn & Kang, 2019). Note that it may be that variations in the SCRAM-CAM and ION RAP TAC could be caused by factors such as skin thickness, temperature, and humidity, although we did perform the study in a climatically controlled environment. Variation in the SCRAM-CAM TAC data could also have been due to the

ankle-positioning of the monitor. In the recruitment stage, participants were asked to report their weight and height, which was used to calculate their BMI. Anyone with a self-reported BMI outside the healthy range (18-30) was excluded from the study to prevent the study sample from becoming too heterogeneous. However, in the laboratory where researchers measured the actual height and weight, two participants were found to be slightly overweight (BMI>30). Given the small discrepancy (BMI<33) we did not exclude these participants. Recently, researchers have reported positioning the SCRAM-CAM ankle bracelet higher up the leg, against the calf (Fairbairn & Kang, 2019). This could possibly secure the positioning of the SCRAM-CAM better as compared to securing on the outside of the ankle, resulting in more consistent and improved TAC readings in laboratory settings. However, wearing the monitor on the inside of the calf would limit the participants' movement and is prone to falling down, which is why the SCRAM-CAM is worn around the ankles in real-world settings. Given this, having participants wear the monitor around the ankles still gives us the best understanding of its validity. Future research should test the new-generation ION Wearable in larger and more diverse samples of participants and in varied settings, to develop a clearer understanding of the detection and validity of TAC data measured using an enzymatic pathway.

Conclusion

We found similar delays and TAC measurement patterns over time when comparing the enzyme-based ION RAP to the fuel-cell SCRAM-CAM. With the advantages and improvements in the size, comfort, and sampling frequency over the traditionally used SCRAM-CAM TAC monitor, the ION RAP holds promises for future real-time alcohol measurement. However, the high failure rate of the ION RAP and its close to zero correlations to BrAC makes the SCRAM-CAM still the more reliable and valid transdermal monitor. With the rapid developments in transdermal technology, it is important to further validate the devices available and to clarify and add further evidence concerning the TAC to BrAC relationship to better assess the potential of transdermal devices for real-time alcohol measurement.

Acknowledgements

We would like to gratefully acknowledge the help and support in the fieldwork from our research colleagues Gabriel Caluzzi and Dan Anderson-Luxford.

Funding

The research reported in this publication was funded by start-up funding of La Trobe University grant for Professor Emmanuel Kuntsche. Kelly van Egmond was supported by the full fee research scholarship (LTUFFRS) and postgraduate research scholarship (LTUPRS) from La Trobe University, Cassandra J. C. Wright PhD, was funded by a NHMRC Early Career Fellowship (1161246), and Michael Livingston PhD, was funded by NHMRC Career Development Fellowship (1123840). The funding was used to purchase the transdermal monitors from both ION Wearable (ION RAPs) and AMS (SCRAM-CAMs). Both companies were informed on results obtained, however they did not have any influence over the reporting or interpretation of the findings. The content is solely the responsibility of the authors and does not necessarily represent the official views of any funding sources.

References

Alcohol Monitoring Services. (2021) *Alcohol Monitoring Services*. https://www.scramsystems.com/our-company/about-us/

Allan, J. T. S., Rahman, M. R., & Easton, E. B. (2017). The influence of relative humidity on the performance of fuel cell catalyst layers in ethanol sensors. *Sensors and Actuators B: Chemical*, *239*, 120–130. https://doi.org/10.1016/J.SNB.2016.07.156

Barnett, N. P., Tidey, J., Murphy, J. G., Swift, R., & Colby, S. M. (2011). Contingency management for alcohol use reduction: a pilot study using a transdermal alcohol sensor. *Drug and Alcohol Dependence*, *118*(2–3), 391–399.

Berchtold, A. (2016). Test–retest: Agreement or reliability?: *Http://Dx.Doi.Org/10.1177/2059799116672875*, *9*. https://doi.org/10.1177/2059799116672875

Bock, S. (2003). Michigan DOC Runs BETA Test of New Remote Transdermal Alcohol Monitoring System. *The Joumal of Offender Monitoring*.

Bush, K., Kivlahan, D. R., McDonell, M. B., Fihn, S. D., Bradley, K. A., & Project, for the A. C. Q. I. (1998). The AUDIT Alcohol Consumption Questions (AUDIT-C): An Effective Brief Screening Test for Problem Drinking. *Archives of Internal Medicine*, *158*(16), 1789–1795. https://doi.org/10.1001/archinte.158.16.1789

Caluzzi, G., Pennay, A., Cook, M., Wright, C., Norman, T., & Kuntsche, E. (2019). Transdermal monitors to assess alcohol consumption in real-time and real-life–a qualitative study on user-experience. *Addiction Research & Theory*, *27*(4), 354–361.

Campbell, A. S., Kim, J., & Wang, J. (2018). Wearable electrochemical alcohol biosensors. *Current Opinion in Electrochemistry*, *10*, 126–135. https://doi.org/10.1016/j.coelec.2018.05.014

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Academic press.

Dougherty, D. M., Charles, N. E., Acheson, A., John, S., Furr, R. M., & Hill-Kapturczak, N. (2012).

Comparing the detection of transdermal and breath alcohol concentrations during periods of alcohol consumption ranging from moderate drinking to binge drinking. *Experimental and Clinical Psychopharmacology*, *20*(5), 373. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3601483/pdf/nihms443283.pdf

Egmond, K. van, Wright, C. J. C., Livingston, M., & Kuntsche, E. (2021). A parallel test of the SCRAM-CAM transdermal monitors ensuring reliability. *Drug and Alcohol Review*. https://doi.org/10.1111/DAR.13353

Fairbairn, C. E., & Kang, D. (2019). Temporal Dynamics of Transdermal Alcohol Concentration Measured via New-Generation Wrist-Worn Biosensor. *Alcoholism: Clinical and Experimental Research*, *43*(10), 2060–2069. https://doi.org/10.1111/acer.14172

Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-Based Nursing*, *18*(3), 66–67. https://doi.org/10.1136/EB-2015-102129

Hill-Kapturczak, N., Lake, S. L., Roache, J. D., Cates, S. E., Liang, Y., & Dougherty, D. M. (2014). Do variable rates of alcohol drinking alter the ability to use transdermal alcohol monitors to estimate peak breath alcohol and total number of drinks? *Alcoholism: Clinical and Experimental Research*, *38*(10), 2517–2522. https://doi.org/http://dx.doi.org/10.1111/acer.12528

Karns-Wright, T. E., Roache, J. D., Hill-Kapturczak, N., Liang, Y., Mullen, J., & Dougherty, D. M. (2017). Time delays in transdermal alcohol concentrations relative to breath alcohol concentrations. *Alcohol and Alcoholism*, *52*(1), 35–41. https://doi.org/http://dx.doi.org/10.1093/alcalc/agw058

Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Normand, S.-L. T., Manderscheid, R. W., & Walters, E. E. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry*, *60*(2), 184–189. https://jamanetwork.com/journals/jamapsychiatry/articlepdf/207204/yoa20567.pdf

Lansdorp, B., Ramsay, W., Hamid, R., & Strenk, E. (2019). Wearable enzymatic alcohol biosensor.

Sensors, *19*(10), 2380.

Leffingwell, T. R., Cooney, N. J., Murphy, J. G., Luczak, S., Rosen, G., Dougherty, D. M., & Barnett, N. P. J. (2013). Continuous objective monitoring of alcohol use: twenty-first century measurement using transdermal sensors. *Alcoholism: Clinical and Experimental Research*, *37*(1), 16–22. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3623856/pdf/nihms446008.pdf

Maher, J. M., Markey, J. C., & Ebert-May, D. (2013). The other half of the story: effect size analysis in quantitative research. *CBE—Life Sciences Education*, *12*(3), 345–351.

Marques, P. R., & McKnight, A. S. (2009). Field and laboratory alcohol detection with 2 types of transdermal devices. *Alcoholism: Clinical and Experimental Research*, *33*(4), 703–711.

Roache, J. D., Karns, T. E., Hill-Kapturczak, N., Mullen, J., Liang, Y., Lamb, R. J., & Dougherty, D. M. (2015). Using transdermal alcohol monitoring to detect low-level drinking. *Alcoholism: Clinical and Experimental Research*, *39*(7), 1120–1127. https://doi.org/http://dx.doi.org/10.1111/acer.12750

Sakai, J. T., Mikulich-Gilbertson, S. K., Long, R. J., & Crowley, T. J. (2006). Validity of transdermal alcohol monitoring: fixed and self-regulated dosing. *Alcoholism: Clinical and Experimental Research*, *30*(1), 26–33.

Sobell, M. B., Sobell, L. C., Klajner, F., Pavan, D., & Basian, E. (1986). The reliability of a timeline method for assessing normal drinker college students' recent drinking history: Utility for alcohol research. *Addictive Behaviors*, *11*(2), 149–161.

van Egmond, K., Wright, C. J. C., Livingston, M., & Kuntsche, E. (2020). Wearable Transdermal Alcohol Monitors: A Systematic Review of Detection Validity, and Relationship Between Transdermal and Breath Alcohol Concentration and Influencing Factors. *Alcoholism: Clinical and Experimental Research*, *44*(10), 1918–1932.

Yeh, K. C., & Kwan, K. C. (1978). A comparison of numerical integrating algorithms by trapezoidal, Lagrange, and spline approximation. *Journal of Pharmacokinetics and Biopharmaceutics*, *6*(1),

79–98.

Table 3. Characteristics of participants and included data.

Characteristics	M (SD) (N = 17)	Range
Participants		
% Female	35.3	
Age (years)	25.7 (5.5)	18.0 – 35.0
BMI	25.2 (3.8)	20.7 - 32.8
% Caucasian	76.5	
AUDIT Score	6.5 (2.2)	1.0 – 10.0
Included sessions		
Peak BrAC	0.08 (0.2)	0.05 – 0.11
Peak SCRAM TAC	0.05 (0.03)	0.01 - 0.13
Peak ION RAP TAC	87.20 (30.54)	37.0 - 163.0

Notes. Shown are means (M) with standard deviations in brackets. BrAC = and units, ION RAP TAC = nA

Table 2. Latencies for transdermally detected alcohol.

	BrAC	SCRAM-CAM TAC	ION RAP TAC	Paired sample t-test (SCRAM-CAM vs ION RAP)
	Mean (SD)	Mean (SD)	Mean (SD)	
Latency to first detected alcohol (minutes)	0 (0)	42.5 (21.3)	49.5 (26.7)	t(27) = -1.042, p = 0.307, d = -0.20
Latency to peak (minutes)	48.1 (25.6)	137.7 (46.5)	154.0 (56.3)	t(28) = -1.358, p = 0.185, d = -0.25
Max cross-correlation lag*	NA	69.9 (46.1)	48 (53)	t(28) = -1.6092, p = 0.119, d = 0.44

Notes. BrAC = breath alcohol concentration, TAC = transdermal alcohol concentration, SD = standard deviation. Mean values are presented in minutes. All time values are calculated regarding the start of drinking.
*Max cross-correlation lags are calculated as compared to the BrAC curves (from start of drinking to BrAC < 0.000%). 1261 minute-datapoints were excluded, mostly due to Bluetooth disturbances and battery failures of the ION RAP ION wristband, resulting in missing data.

A. **B**. **C**.

Figure 4. **A**. ION RAP wristband, **B.** ION RAP cartridge activation was completed manually by researchers. This included extracting gel from the pink tube using the blue straw and spreading it onto the electrode surface of the cartridge, and **C.** SCRAM-CAM ankle bracelet.

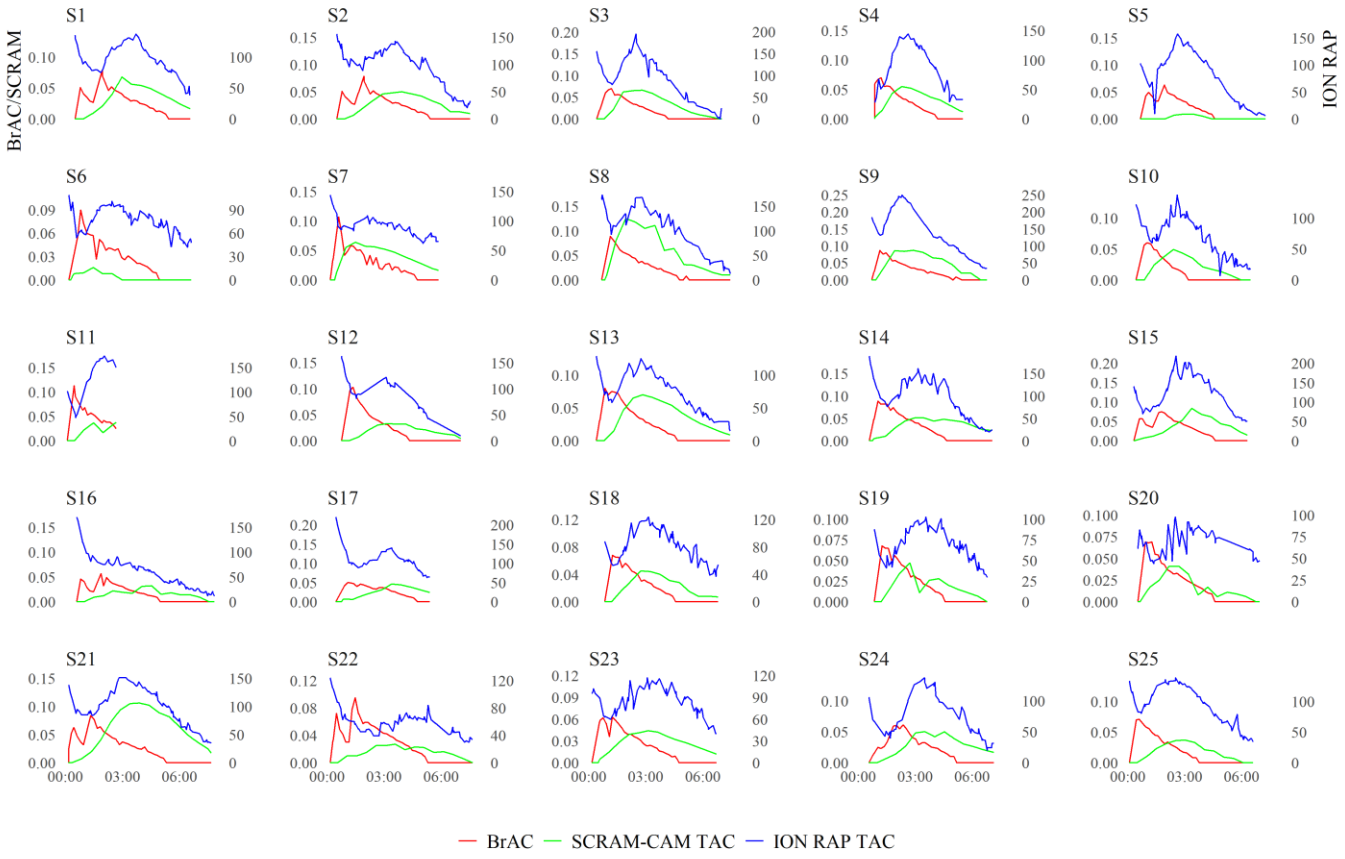

Figure 2. SCRAM-CAM TAC, ION RAP TAC, and BrAC data for each of the 25 sessions, including 17 participants. The data is reflecting the moment of receiving alcohol to the end of the session (about two to three hours after the participants reached 0.000 BrAC.

Notes. S refers to the session. BrAC = breath alcohol concentration, TAC = transdermal alcohol concentration.

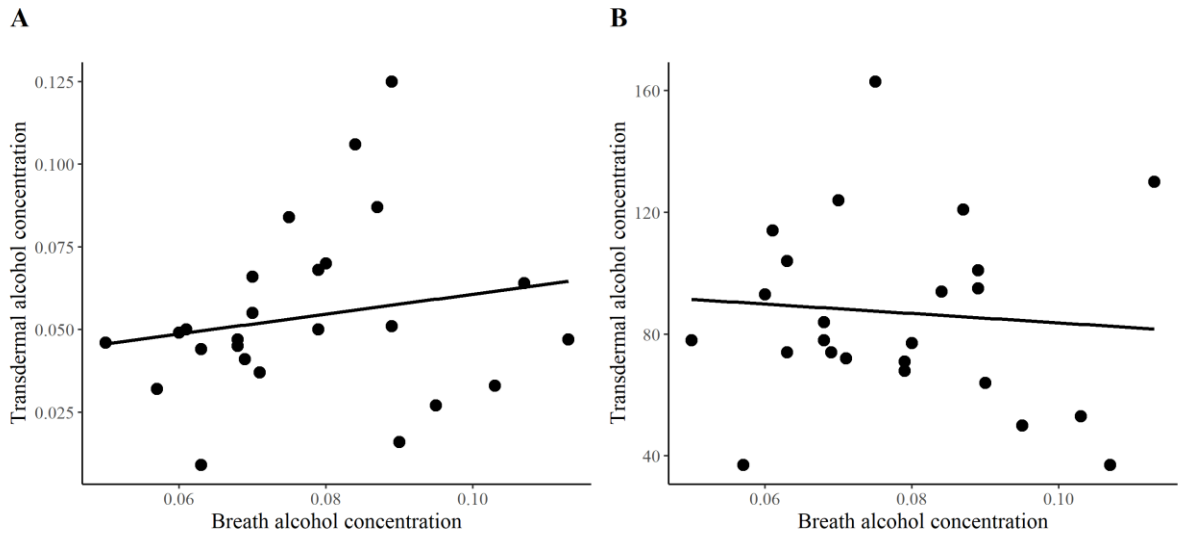

Figure 3. Regressions of the individual participants TAC values on BrAC values. **A.** Regressing the individual SCRAM TAC peak values on the BrAC peak values. **B.** Regressing the individual ION RAP TAC peak values on the BrAC peak values.

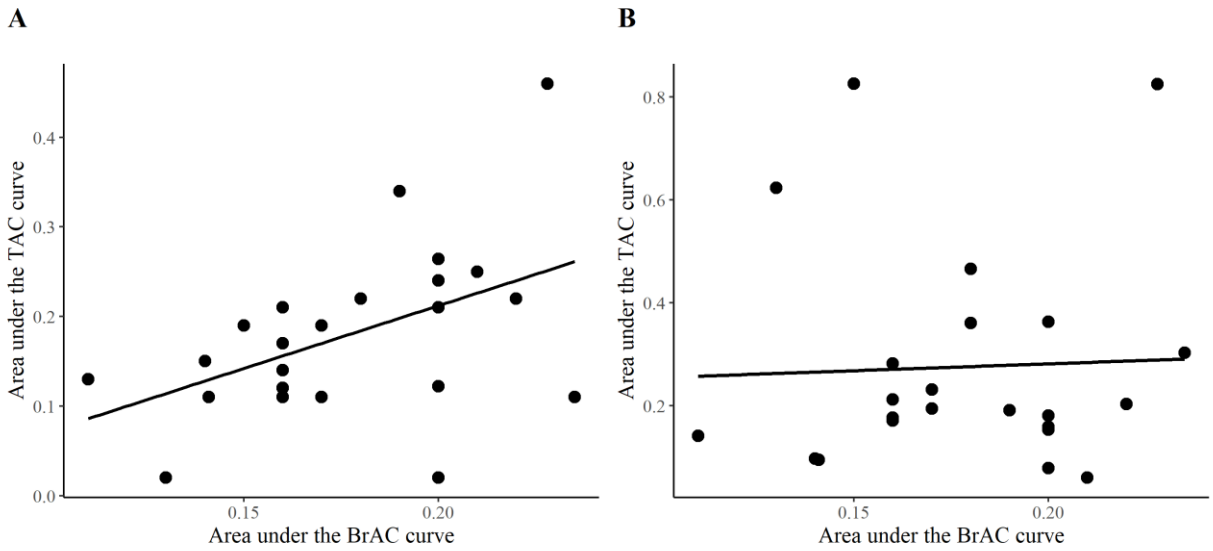

Figure 4. Regressions of the individual participants TAC AUC values on BrAC AUC values. **A.** Regressing the individual SCRAM TAC AUC values on the BrAC AUC values. **B.** Regressing the individual ION RAP AUC values on the BrAC AUC values.

7. Measuring alcohol consumption in sporting spectators using ecological momentary assessments and transdermal alcohol monitors

Chapter 7 contains the following manuscript:

van Egmond, K., Anderson-Luxford, D., Kuntsche, E., JC Wright, C., Caluzzi, G., Pennay, A. (submitted for publication). Measuring alcohol consumption while watching sport events: a feasibility study comparing ecological momentary assessments and transdermal alcohol monitors.

Author contributions

KvE led the data collection with help from DAL, CW, and GC. AP led study design, and the qualitative data analysis, provided critical feedback and suggested revisions. KvE led the quantitative data analysis and led manuscript writing. DAL was in involved in the data analysis and suggested revisions. EK and CW were involved in study design, provided critical feedback and suggested revisions. GC further suggested revisions.

Measuring Alcohol Consumption while Watching Sport Events: A Feasibility Study Comparing Ecological Momentary Assessments and Transdermal Alcohol Monitors

Kelly van Egmond[1], Daniel Anderson-Luxford[1], Emmanuel Kuntsche[1,2], Cassandra Wright[1,3,4], Gabriel Caluzzi[1] and Amy Pennay[1]

[1.] Centre for Alcohol Policy and Research (CAPR), La Trobe University, Melbourne, Australia.

[2.] Institute of Psychology, Eötvös Loránd University, Budapest, Hungary

[3.] Menzies School of Health Research, Charles Darwin University, Darwin, Australia

[4.] Burnet Institute, Melbourne, Australia.

Corresponding author: Kelly van Egmond, Centre of Alcohol Policy Research, School of Psychology and Public Health, La Trobe University, 3086 Bundoora, VIC, Australia. e-mail: k.vanegmond@latrobe.edu.au, phone: +61 04 98257320.

Keywords: alcohol consumption, ecological momentary assessment, transdermal alcohol measurement, sport spectators.

Abstract

Background: Accurate real-time information about the relationship between alcohol consumption and sports spectatorship is necessary to inform meaningful approaches to reducing harmful drinking patterns while watching sport. This study aimed to evaluate and compare ecological momentary assessment (EMA) surveys and transdermal SCRAM-CAM monitors to measure drinking over the course of a day while watching Australian Rules Football (AFL).

Methods: During 29 AFL events, 13 participants wore a SCRAM-CAM monitor while simultaneously completing EMA surveys about their drinking behaviour. Correspondence and correlation between the self-reported drinks and transdermal alcohol concentration (TAC) was measured. An exit survey assessed experiences with EMA and SCRAM-CAM.

Results: Alcohol consumption was self-reported on 24 (83.3%) of the 29 events, with an average of 5.0 standard drinks consumed over 2.3 hours. Correspondence was considered to be good at r=0.62. TAC curves showed large-sized correlations to the number of self-reported drinks (r=0.55-0.67). Participants noted discomfort while wearing the SCRAM-CAM, whilst also noting some annoyance at having to complete EMA surveys during a match, which became harder after drinking more alcohol.

Conclusion: This preliminary study found that it is feasible to monitor alcohol consumption in real-time using both EMA and transdermal monitors in an AFL spectator sample. Both methods exhibited strengths and limitations for measuring alcohol consumption, with each presenting promising avenues of inquiry for further research in a larger sample. We suggest that a combination of the two methods will inform the most meaningful approaches for prevention and intervention strategies to reduce harmful drinking among sport spectators.

Short Summary

This study reconfirms the relationship between sport spectatorship and risky alcohol consumption and demonstrates the feasibility of monitoring alcohol consumption in real-time while watching sport using both EMA and transdermal monitors. Next steps are to investigate the relationship between sport-watching, heavy drinking, and harms to inform prevention and intervention strategies.

Disclosures and Acknowledgements

This research was supported by a La Trobe University internal grant. Kelly van Egmond was supported by the full fee research scholarship (LTUFFRS) and postgraduate research scholarship (LTUPRS) from La Trobe University, Cassandra J. C. Wright PhD, was funded by a NHMRC Early Career Fellowship (1161246), and Amy Pennay is supported by an Australian Research Council Discovery Early Career Researcher Award (190101074).

All authors contributed in a significant way to the manuscript and authors have read and approved the final manuscript.

The authors report no declarations of interest.

Introduction

International research has identified a strong connection between heavy drinking and sports spectatorship (Estrada and Tryggvesson 2001; Nelson and Wechsler 2003; Palmer 2011). Sports spectators consume more alcohol than non-sports spectators (Nelson and Wechsler 2003), and studies have identified a link between excessive alcohol use and alcohol-related harm, violence, and crime among sport spectators (Glassman et al. 2007; Kalist and Lee 2016; Kingsland et al. 2013; Palmer 2011). High rates of alcohol related ambulance and emergency department attendances are recorded after major Australian Rules Football (AFL) matches (Lloyd et al. 2013), which is the most popular sport in Australia (Australian Bureau of Statistics 2010). While qualitative work has investigated practices and cultures of drinking while watching AFL (Palmer 2011; Palmer and Thompson 2007), to our knowledge, no studies have investigated quantities of alcohol consumed while watching AFL matches. This is important because we need to know how much, and in what ways, AFL spectators are consuming alcohol to meaningfully inform prevention and intervention approaches.

To measure quantities of alcohol consumed, alcohol researchers most commonly use self-reported data (Dawson 1998). However, most self-reports are completed retrospectively over longer periods (e.g., typical drinking in the last month, or amount consumed yesterday/last night) and are subject to recall bias, which makes them ill-suited to measure drinking at an event level (Kuntsche and Labhart 2012). Given these limitations, researchers have used breathalyser data to measure intoxication among sports spectators. For example, in the US, when 747 baseball spectators were cross-sectionally breathalysed upon entrance to the match and then again during the match, 41% tested positive for alcohol, with 8.4% testing at or above the US legal breath alcohol concentration (BrAC) driving limit of 0.08% (Wolfe et al. 1998). Comparable results were found in a sample of 4420 Swedish Premier Football League spectators, with 46.8% testing positive for alcohol during the match, and 8.9% testing above 0.1% BrAC (Durbeej et al. 2017). While providing an objective measurement of intoxication; however, breathalysers only measure alcohol at a specific time-point, missing information on drinking patterns. They also require active researcher engagement and are thus costly and time intensive to administer.

Ecological Momentary Assessments (EMA) surveys can overcome some of these burdens, by asking participants to report data on their drinking events and potentially other behaviours in the moment (Kuntsche and Labhart 2013a, 2014). It has been observed that EMA methods record higher alcohol consumption than retrospective methods, as well as higher agreement with breathalyser data (Mun et al. 2021). Notably, EMA surveys can be collected using smartphone applications, making them affordable and convenient to implement (Kuntsche and Labhart 2013a). However, collecting real-time EMA data requires active and regular participation, which can in turn influence participants' willingness to engage (Piasecki 2019).

Transdermal alcohol monitors, such as the Secure Continuous Remote Alcohol Monitor Continuous Alcohol Monitoring (SCRAM-CAM™; Alcohol Monitoring Systems Inc., Highlands Ranch, CO), are a promising tool to monitor drinking objectively and continuously over several days without active participation from the participant. These monitors measure alcohol consumption by analysing the alcohol that is secreted through skin. SCRAM-CAMs are not as accurate as breathalysers. They have been found to detect approximately 73% of 690 self-reported drinking episodes (Barnett, Meade, and Glynn 2014), and 65% of 324 self-reported drinking days (Karns-Wright et al. 2018). However, previous research using SCRAM-CAMs has reported user discomfort (Caluzzi et al. 2019), and given they were originally developed to monitor drinking among offenders (Flango and Cheesman 2009; Voas et al. 2011), some participants have reported concern about wearing the monitor in public (Marques and McKnight 2007). Finally, due to the ethanol transportation through the skin being physiologically complex, they can exhibit low sensitivity to lower-level alcohol consumption (Roache et al. 2019), and significant delays in alcohol detection have been found, with lag times reported up to four hours (Fairbairn and Kang 2019; Karns-Wright et al. 2017; Leffingwell et al. 2013).

Studies measuring real-time alcohol consumption with the use of EMA and transdermal monitors have found a correspondence of 73% (Mun et al. 2021) and 86% (Simons et al. 2015) and a significant correlation between the two methods (Mun et al. 2021). To date, only Norman et al., (2020) have tested the feasibility of both methods when measuring alcohol consumption during an event (a music festival), recommending the use of both measurements to provide the most

comprehensive overview of intoxication. How this translates to sport spectators who are watching a sporting event in real-time, and not wanting to miss the events of the game, is unclear.

It is important to understand the way in which sports spectators consume alcohol while watching sporting matches to inform prevention and intervention initiatives. However, there are clearly challenges with regards to the best way to collect and measure alcohol consumption while watching sport. In order to inform future research, the current study investigates the use of both EMA surveys and SCRAM-CAM monitors in: 1) feasibility of measuring alcohol consumption during sporting event drinking occasions, 2) correspondence in alcohol measurement, and 3) user experience.

Methods

Study recruitment

Facebook and Instagram advertisements targeted adults in metropolitan Melbourne (where most AFL matches are played). Participants completed a screening survey to determine eligibility. To complement this approach and to specifically target AFL spectators who attend matches, we used a street intercept approach at Melbourne's largest AFL stadium, the Melbourne Cricket Ground (MCG). Six researchers engaged spectators entering the MCG to watch an AFL match over two days (four matches in total).

Our predetermined sample size for this feasibility study was 15 participants, due to our possession of this many SCRAM-CAM monitors. Participants who completed the screening survey were contacted by phone for further in-depth screening and study information, where the EMA and SCRAM-CAM monitor methodology was described in detail.

Inclusion criteria were: being aged 18 years or older; regular consumption of alcohol while watching AFL (≥2 standard drinks during one single occasion in the preceding month while watching their AFL team play); watching their AFL team at least fortnightly; owning a smartphone and being willing to use it for the study; and being able to read and understand English. Exclusion criteria included medical conditions that prohibited use of the SCRAM-CAMs (e.g. circulation problems, leg ulcers, tendonitis, diabetes, history of swelling, neuropathy, deep vein thrombosis, or a nickel or metal

allergy). Written consent was collected from participants prior to the start of the first AFL match. The study was approved by the Human Research Ethics Committee at La Trobe University (HEC18524).

Study procedure

Participants were asked to nominate three consecutive weeks where they planned to watch their AFL teams' matches either in-person at the stadium or at another location (e.g. home or the pub). Twelve participants participated during all three matches, two participants participated during two matches, and one participant during one match, resulting in data collection for 15 participants and a total of 41 days a match took place. Participants wore a SCRAM-CAM ankle monitor for two to three days during the weekends of the three nominated matches and downloaded a smartphone application for the EMA component (LifeData™ Inc.). The EMA app was set up in such a way that it would send reminder notifications to complete each survey for the nominated matches (tailored to each participant). AFL matches were played on Friday nights, Saturdays, and Sundays, with the earliest match time beginning 1.10pm and the latest being 7.50pm. AFL matches have four quarters of approximately half an hour duration (total match time is approximately two and a half hours including breaks between quarters). An initial survey was sent each week on Friday to ask participants which match they would be watching that weekend; their response would consequently trigger their match-day surveys at the right time. Surveys were sent to participants ten minutes before the match, at quarter time break, half-time break, three-quarter time break, after the match, and then every two hours until midnight. The maximum number of surveys a participant would receive was 10 per day. Each weekend, participants attended an appointment where a researcher fitted and removed a SCRAM-CAM monitor on the day of the match, and a second appointment to remove the monitor the day after the match. Participants were reimbursed AUD$50 per session, provided at the monitor removal appointment.

Measures

Self-reported events

The survey prior to the match asked whether the participant had already consumed any drinks and how many standard drinks (10g of ethanol) they had consumed up until that time. An image with examples of standard drinks was given to improve reporting accuracy. All following surveys asked the number of standard drinks since the last survey. Out of the total 41 days participants that observed an AFL match, missing data occurred for 11 matches (26.2%), due to missing the initial survey which then did not trigger match-day surveys (unknown whether this was caused by non-completion of the participant or technical issues with the EMA survey application), resulting in a total of 30 self-reported events from 13 participants.

The number of drinks was a sum of the reported standard drinks across all surveys completed over the self-reported event (before, during and after the AFL match). The only previous available study looking at correspondence between daily self-reports and TAC data classified both the self-reported drinking and TAC data into categories to improve interpretation (Karns-Wright et al. 2018). Using these same methods, drinking was categorised as follows: None: no drinking; Moderate: >0 drinks and <=5 drinks for men (<=4 drinks for women); and Heavy: >5 drinks for men (>4 drinks for women).

TAC events

The SCRAM-CAM ankle bracelet was fitted and secured by a researcher to ensure contact with the skin without causing severe discomfort or restricted blood flow. To deal with the delays in TAC measurement, a TAC event was measured in a noon-to-noon timeframe. TAC events were identified following the nine TAC research rules as developed by Roache et al., (2019). A TAC event is defined as a non-zero (positive) TAC reading preceded and followed by at least two zero TAC readings (Roache et al. 2019). During the 41 days participants reported viewing an AFL match, a total of 45 TAC events were identified, with some multiple TAC events within a single day. Following the TAC research rules (Roache et al. 2019), 417 individual TAC datapoints were excluded because of an inter-

reading interval smaller than 20 minutes instead of the regular 30-minute interval (*n* =397), or due to steep reading-to-reading slope rises (> 0.182 g/dl/hour) or drops (< -0.126 g/dl/hour) (*n* = 20). Twelve full TAC events were removed due to: an implausibly high start TAC reading (*n* = 1), the event consisted of a single TAC reading (*n* = 9), the event had all negative slopes after the first TAC reading (*n* = 1), and because the peak TAC was below 0.01 g/dl, spanning for more than 240 minutes (*n* = 1). This resulted in a total of 33 TAC events from 15 participants.

As with self-reported events, TAC events were classified following the categories used by Karns-Wright et al., (2018): None: 1 non-zero TAC readings; Low: 3 or more TAC readings > 0 but no readings > .01 g/dl; Moderate: ≥ 3 TAC readings above 0 and ≥ 1 TAC reading above 0.01 g/dl but < 2 readings above 0.02 g/dl; Heavy: 2 or more TAC readings > 0.02 g/dl.

AMS TAC events identification

The SCRAM-CAM developers, AMS, review the TAC data after it has been downloaded to the secure online server. AMS only reviews positive TAC readings when the TAC event contains ≥ 3 TAC readings above 0.02 g/dl (See Barnett et al., 2014; Roache et al., 2019 for a further description). To compare our findings with AMS findings we used the reports on the secure online server informing on how many TAC events were identified for the period a SCRAM-CAM was active.

Exit survey

After the final AFL match, at the monitor removal appointment participants were asked to complete an exit survey to understand their experiences of both the EMA and SCRAM-CAM components. This consisted of five open-ended questions which took approximately five minutes to complete.

Data Analysis

From the 30 self-reported events (*n*=13) and 33 TAC events (*n*=15), we only included data where both EMA and TAC findings were available, resulting in a final analytic sample of 29 drinking events from *n* =13 participants. The data was processed in R version 4.0.2. and then further analysed using IBM SPSS Statistics for Windows, version 27 (IBM Corp., Armonk, N.Y., USA). Descriptive

statistics including mean, range and total number of self-reported drinks, peak TAC, and area under the TAC curve (AUC), were calculated to measure alcohol consumption during the drinking event. To study the time windows of the drinking events, the time of a self-reported drinking event was measured from the first survey the participant reported alcohol consumption. The time window of the TAC event was measured from the zero TAC value proceeding the first positive TAC value, to the next zero TAC value or the last point of the TAC event. The intraclass correlation coefficient (ICC) was used to test correspondence. A two-way random-effect model based on single ratings and absolute agreement assessed the correspondence between the self-reported and TAC events. Interpretation was as follows: below 0.40, poor; between 0.40 and 0.59, fair; between 0.60 and 0.74, good; 0.75 or above, excellent (Cicchetti 1994). AUC was computed using the trapezoidal rule (Dodd and Pepe 2003), which means that the AUC is the sum of the trapezoids under the curve. Pearson's correlation coefficients were computed to analyse the relationship between the TAC curve characteristics (peak and AUC) and the total self-reported number of drinks per event. A Pearson's *r* correlation of 0.10, 0.30, 0.50, and 0.70, was considered to indicate a small, medium, large, and very large effect (Cohen 2013; Maher, Markey, and Ebert-May 2013). To study the user experience of both the EMA and the SCRAM-CAM, the content of the exit surveys was analysed thematically and the frequency of themes were collated (Züll 2016).

Results

Participant demographics

Table 1 shows the participant demographics of the final sample (N = 13).

Drinking events reported and identified

The average number of drinks reported across the 29 drinking events was 5.0 standard drinks (Table 2). These drinks were reported to be consumed in 2 hours and 15 minutes on average. While the transdermal drinking events lasted an average of 5.5 hours, they could last up to 27.5 hours.

Correspondence between self-reported and TAC events

Overall, the ICC between the self-reported and TAC events was considered to be good at r=0.62.

Out of the 29 matched events, positive TAC events (low, moderate, or heavy) occurred for 20 (69.0%) events and self-reported drinking (moderate and heavy) occurred for 25 events (81.0%) (Table 3). This means that for nine events, there was no TAC event (SCRAM-CAM did not detect any alcohol) and for four events the participants did not report any drinking (Table 3).

When participants did not report any drinking, moderate TAC events were detected for 40% of the events. When heavy drinking was reported, TAC detected heavy drinking for 88.9% of events and with one occasion of moderate drinking detected (11.1%). All self-reported heavy drinking events were therefore detected by the SCRAM-CAM.

Using the AMS detection criteria, only 7 (26.9%) drinking events were detected, and all these drinking events corresponded with heavy self-reported drinking events.

Correlations between TAC characteristics and self-reported number of drinks

Correlations between AUC and the total number of self-reported drinks were positive and large-sized; r=0.549. Peak TAC values also showed a large-sized correlation with the self-reported number of drinks; r=0.673 (Figure 1).

Exit survey

Participants were generally positive about their experiences in the project due to their interest in the technological aspects of the study. The main critical feedback offered was divided into those who found the SCRAM-CAM uncomfortable and those who had issues with either the survey application or the questions in the EMA component. Three participants had no issues with the SCRAM-CAM monitors, three indicated it was very uncomfortable, and the remainder noted it was somewhat uncomfortable (although most noted that they had adjusted to a level of comfortability by the second week; they did not wear the monitor in the days between the weekends). Two participants had difficulty getting the survey application to function properly for them, one participant said the questions became harder to answer after drinking more alcohol, and another suggested that it was

annoying to have to answer questions during the match. Almost all participants suggested that the reimbursement was an important part of their participation, given the time required for participation including time and travel for fitting and removing the monitor.

Discussion

According to our EMA data, alcohol was consumed during 24 (83.0%) of the 29 events, with the SCRAM-CAM detecting alcohol for 19 (69.0%) events. Participants reported drinking an average of 5.0 standard drinks over 2.3 hours while watching AFL, which is considered risky drinking (defined by the Australian government as more than 4 standard drinks in one sitting (NHMRC 2020)). This is congruent with studies measuring moderate to high levels of intoxication as measured with breathalysers among sport spectators in Sweden and the United States (Durbeej et al. 2017; Wolfe et al. 1998).

There are several considerations to bear in mind regarding the feasibility of the research methods tested here. When using the AMS criteria to detect TAC events, all moderate drinking events were missed. Using the TAC criteria as recommended by Karns-Wright et al., (2018) increased the sensitivity to lower alcohol consumption and resulted in the SCRAM-CAM detecting more than half of the moderate self-reported drinking events. In total, of the 24 self-reported drinking events, a positive TAC event was detected for 16 (66.6%) of these events. Good correspondence between self-reported and TAC events was identified, with an ICC coefficient of 0.62. This correspondence is lower than the previously reported correspondence between EMA and TAC of 73-86% (Mun et al. 2021; Simons et al. 2015), however, this is the first study to test correspondence between the two methods during a sport event. It is important to note that a significant amount of moderate drinking was still missed by the SCRAM-CAM and higher correspondence was found with heavy drinking events. These findings suggest that the SCRAM-CAM monitors may only be useful for studies concerned with detecting heavier drinking practices.

We observed positive and large-sized correlation coefficients between peak TAC and AUC with self-reported number of drinks (r=0.55, and r=0.67 respectively). Though the correlation between self-

reported number of drinks and TAC measurements is classified as large, this might be driven by an outlier effect (participant 4 and 8 (Figure 1)) and by the high correspondence in the heavy drinking category. However, removing participant 4 and 8 increased the correlation even further (r=0.736 and 0.852, for peak and AUC respectively), suggesting no outlier effect. Similarly, the lower correspondence we identified may be due to almost half of the self-reported moderate drinking events going undetected by the SCRAM-CAM. Wearing the SCRAM-CAM tight around the ankle can be uncomfortable for participants (Caluzzi et al. 2019), resulting in varied distance between the monitor and the skin through which the alcohol could evaporate and go undetected by the SCRAM-CAM.

On the other hand, during heavy drinking episodes the SCRAM-CAM can potentially measure alcohol when the ability to self-report is decreasing. Previous research has found that heavy drinking can increase errors in self-reported drinking (Davis, Thake, and Vilhena 2010; Livingston and Callinan 2015). Indeed, for 40.0% of the detected TAC events categorized as moderate participants reported they did not drink, and for a further 18.0% of the detected TAC events categorized as heavy, participants reported only moderate drinking events. It is unclear whether these discrepancies resulted from problems with self-reporting or false-positive TAC events (e.g., alcohol being spilled near the monitor or alcohol in the carpet). Further research is needed to determine the magnitude and direction of these disagreements to come to stronger conclusions.

In line with previously reported limitations of the use of EMA surveys (Piasecki 2019), the main issue reported by participants was the response burden. Some participants had trouble with the functionality of the application, completing surveys with increased alcohol consumption, and having to complete the surveys when watching an AFL match. However, due to a limitation in the design of the EMA surveys, most missing data (72.0%) was due to participants missing the first survey, resulting in subsequent surveys not being triggered. This could have been solved if participants had the option to initiate the surveys during a game themselves. Unfortunately, due to the choice of EMA application software, this was not possible.

Given that the EMA application could easily be downloaded on participants' smartphones without any support from the researchers, EMA surveys are likely to be more suitable than SCRAM-CAMs for

collecting data from sports spectators in larger studies. In contrast, SCRAM-CAMs were burdensome to fit and remove for each session, and their high cost meant we had to limit our sample size. Although we could have left the monitors on participants for the full three-week period, we decided against this due to our previous studies showing discomfort with sleep, exercise and challenges keeping the device dry (i.e., when showering, etc.). Moreover, TAC detection shows substantial delays due to the way ethanol is excreted through the skin (Karns-Wright et al., 2018). This makes it harder to know when a participant has started drinking. Though participants reported discomfort, congruent with both Australian festival attendees (Caluzzi et al. 2019) and a young adult population in the United States (Marques and McKnight 2007), they adjusted to a certain level of comfort by the second week.

This study provides a first investigation of quantitative methods measuring alcohol consumed in an AFL spectator sample, however some limitations are present. Notably, the participant sample of only 29 usable drinking events with matched EMA and TAC data from 13 participants. For this reason, we decided to focus solely on effect sizes rather than significance levels. Further, this small sample size limited analysis on factors predictive of higher levels of drinking, such as gender, body mass index, and location of viewing (pub or in-person at the stadium). Given that TAC drinking events are difficult to interpret in terms of amount consumed, as TAC data does not relate directly to BrAC levels, we used the categories as reported previously (Karns-Wright et al. 2018). Using these categories could have influenced the correspondence. Researchers are currently developing software to convert TAC data to BrAC estimations which could lead to a better correspondence and understanding of the TAC data (van Egmond et al. 2020; Leffingwell et al. 2013; S E Luczak, Rosen, and Wall 2015). Finally, participants only self-reported their drinking up until midnight. While the next-day survey allowed them to report further drinking, it did not contain information about the time that the additional drinks were consumed.

Implications and future directions

Given higher rates of ambulance and emergency department presentations after AFL matches (Lloyd et al. 2013), accurate and nuanced information about the relationship between sports spectatorship,

alcohol consumption, and alcohol-related harms is necessary to inform meaningful approaches to prevention and intervention. In order to do so, a first step is to identify the best way to collect such data. Showing both a good correspondence and a large-sized correlation between the two methods, this preliminary study found that it is feasible to monitor alcohol consumption using both EMA and TAC monitors in an AFL spectator sample. Next steps are to investigate consumption patterns in a larger sample (allowing for additional covariates) and investigate the relationship between heavy drinking while watching sport and resultant experience of harms. Using EMA surveys will enable researchers to collect more information on drinking contexts, such as location or company, and any harms that might have occurred. Further, a clear benefit of the SCRAM-CAM is the ability to provide detailed information on intoxication levels and drinking patterns, by studying the absorption and elimination rates. This information will deepen our understanding of the time-sensitive relationship between AFL sport spectatorship, drinking patterns, contexts, and alcohol-related harms. We suggest that a combination of the two methods will inform the most meaningful approaches for prevention and intervention strategies to reduce harmful drinking among sport spectators.

Data Availability Statement

The data underlying this article cannot be shared publicly due to reasons relating to the privacy of the participants in this study. The data might be shared on reasonable request to the corresponding author.

References

Alcohol Monitoring Services. n.d. “Alcohol Monitoring Services.” *2022*. Retrieved February 5, 2022 (https://www.scramsystems.com/our-company/about-us/).

Allan, Julaine, Anton Clifford, Patrick Ball, Margaret Alston, and Peter Meister. 2012. “‘You’re Less Complete If You Haven’t Got a Can in Your Hand’: Alcohol Consumption and Related Harmful Effects in Rural Australia: The Role and Influence of Cultural Capital.” *Alcohol and Alcoholism* 47(5):624–29. doi: 10.1093/ALCALC/AGS074.

Anon. 2022. “BACtrack Skyn™.” *BACtrack*. Retrieved February 5, 2022 (https://skyn.bactrack.com/).

Australian Bureau of Statistics. 2010. “Spectator Attendance at Sporting Events, Australia (Cat. No. 4174.0).” *Canberra: ABS*.

Australian Institute of Health and Welfare. 2020. “National Drug Strategy Household Survey 2019.” *Canberra: AIHW* 90.

Australian Institute of Health and Welfare. 2018. *Australian Burden of Disease Study: Impact and Causes of Illness and Death in Australia 2018.*

Barnett, Nancy P., E. B. Meade, and T. R. Glynn. 2014. “Predictors of Detection of Alcohol Use Episodes Using a Transdermal Alcohol Sensor.” *Experimental and Clinical Psychopharmacology* 22(1):86–96. doi: 10.1037/a0034821.

Barnett, Nancy P., Jennifer Tidey, James G. Murphy, Robert Swift, and Suzanne M. Colby. 2011. “Contingency Management for Alcohol Use Reduction: A Pilot Study Using a Transdermal Alcohol Sensor.” *Drug and Alcohol Dependence* 118(2–3):391–99.

Berchtold, André. 2016. “Test–Retest: Agreement or Reliability?:” *Http://Dx.Doi.Org/10.1177/2059799116672875* 9. doi: 10.1177/2059799116672875.

Blair, A. L., A. L. Chiaf, E. K. Crockett, T. K. Teague, and J. M. Croff. 2021. “Validation of Hair

Ethyl Glucuronide Using Transdermal Monitoring and Self-Reported Alcohol Use in Women of Childbearing Potential." *Neuropsychopharmacology Reports* 41(2):144–51. doi: https://dx.doi.org/10.1002/npr2.12151.

Bland, Martin J., and Douglas G. Altman. 1986. "Statistical Methods for Assessing Agreement between Two Methods of Clinical Measurement." *The Lancet* 327(8476):307–10. doi: 10.1016/S0140-6736(86)90837-8.

Bond, Jason C., Thomas K. Greenfield, Deidre Patterson, and William C. Kerr. 2014. "Adjustments for Drink Size and Ethanol Content: New Results from a Self-Report Diary and Transdermal Sensor Validation Study." *Alcoholism: Clinical and Experimental Research* 38(12):3060–67. doi: http://dx.doi.org/10.1111/acer.12589.

Bühler, Mira, Sabine Vollstädt-Klein, Andrea Kobiella, Henning Budde, Laurence J. Reed, Dieter F. Braus, Christian Büchel, and Michael N. Smolka. 2010. "Nicotine Dependence Is Characterized by Disordered Reward Processing in a Network Driving Motivation." *Biological Psychiatry* 67(8):745–52. doi: 10.1016/j.biopsych.2009.10.029.

Callinan, Sarah. 2015. "How Big Is a Self-Poured Glass of Wine for Australian Drinkers?" *Drug and Alcohol Review* 34(2):207–10. doi: 10.1111/DAR.12154.

Caluzzi, Gabriel, Amy Pennay, Megan Cook, Cassandra Wright, Thomas Norman, and Emmanuel Kuntsche. 2019. "Transdermal Monitors to Assess Alcohol Consumption in Real-Time and Real-Life–a Qualitative Study on User-Experience." *Addiction Research & Theory* 27(4):354–61.

Cicchetti, Domenic V. 1994. "Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology." *Psychological Assessment* 6(4):284.

Cohen, B. Beth, and Daniel C. Vinson. 1995. "Retrospective Self-Report of Alcohol Consumption: Test-Retest Reliability by Telephone." *Alcoholism: Clinical and Experimental Research* 19(5):1156–61.

Cohen, Jacob. 2013. *Statistical Power Analysis for the Behavioral Sciences*. Academic press.

Collins, Allan C., Toni N. Yeager, M. E. Lebsack, and S. Scott Panter. 1975. "Variations in Alcohol Metabolism: Influence of Sex and Age." *Pharmacology Biochemistry and Behavior* 3(6):973–78.

Collins, R. Lorraine, Elizabeth T. Morsheimer, Saul Shiffman, Jean A. Paty, Maryann Gnys, and George D. Papandonatos. 1998. "Ecological Momentary Assessment in a Behavioral Drinking Moderation Training Program." *Experimental and Clinical Psychopharmacology* 6(3):306–15. doi: 10.1037/1064-1297.6.3.306.

Croff, J. M., M. L. Hartwell, A. L. Chiaf, E. K. Crockett, and I. J. Washburn. 2021. "Feasibility and Reliability of Continuously Monitoring Alcohol Use among Female Adolescents and Young Adults." *Drug & Alcohol Review* 40(7):1143–54. doi: https://dx.doi.org/10.1111/dar.13045.

Csikszentmihalyi, Mihaly, Reed Larson, M Csikszentmihalyi, and R Larson. 2014. "Validity and Reliability of the Experience-Sampling Method." *Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi* 35–54. doi: 10.1007/978-94-017-9088-8_3.

Davis, Christopher G., Jennifer Thake, and Natalie Vilhena. 2010. "Social Desirability Biases in Self-Reported Alcohol Consumption and Harms." *Addictive Behaviors* 35(4):302–11.

Dawson, Deborah A. 1998. "Measuring Alcohol Consumption: Limitations and Prospects for Improvement." *Addiction* 93(7):965.

Degenhardt, Louisa, Wayne Hall, and Michael Lynskey. 2001. "Alcohol, Cannabis and Tobacco Use among Australians: A Comparison of Their Associations with Other Drug Use and Use Disorders, Affective and Anxiety Disorders, and Psychosis." *Addiction (Abingdon, England)* 96(11):1603–14. doi: 10.1046/J.1360-0443.2001.961116037.X.

Dodd, Lori E., and Margaret S. Pepe. 2003. "Partial AUC Estimation and Regression." *Biometrics* 59(3):614–23.

Dougherty, Donald M., Nathalie Hill-Kapturczak, Yuanyuan Liang, Tara E. Karns, Sarah L. Lake, Sharon E. Cates, and John D. Roache. 2015. "The Potential Clinical Utility of Transdermal Alcohol Monitoring Data to Estimate the Number of Alcoholic Drinks Consumed." *Addictive Disorders & Their Treatment* 14(3):124–30. doi: 10.1097/ADT.0000000000000060.

Dulin, Patrick L., Corene E. Alvarado, James M. Fitterling, and Vivian M. Gonzalez. 2017. "Comparisons of Alcohol Consumption by Timeline Follow Back vs. Smartphone-Based Daily Interviews." *Addiction Research & Theory* 25(3):195–200.

Durbeej, Natalie, Tobias H. Elgán, Camilla Jalling, and Johanna Gripenberg. 2017. "Alcohol Intoxication at Swedish Football Matches: A Study Using Biological Sampling to Assess Blood Alcohol Concentration Levels among Spectators." *PLoS One* 12(11):e0188284.

van Egmond, Kelly, Cassandra J. C. Wright, Michael Livingston, and Emmanuel Kuntsche. 2020. "Wearable Transdermal Alcohol Monitors: A Systematic Review of Detection Validity, and Relationship Between Transdermal and Breath Alcohol Concentration and Influencing Factors." *Alcoholism: Clinical and Experimental Research* 44(10):1918–32.

Estrada, Felipe, and Kalle Tryggvesson. 2001. "A Part of the Game–Alcohol, Football Fans and Male Comradeship." *Nordic Studies on Alcohol and Drugs* 18(3):245–60.

Fairbairn, C. E., and D. Kang. 2019. "Temporal Dynamics of Transdermal Alcohol Concentration Measured via New-Generation Wrist-Worn Biosensor." *Alcoholism: Clinical and Experimental Research* 43(10):2060–69. doi: 10.1111/acer.14172.

Fairbairn, C. E., D. Kang, and N. Bosch. 2020. "Using Machine Learning for Real-Time BAC Estimation from a New-Generation Transdermal Biosensor in the Laboratory." *Drug & Alcohol Dependence* 216:108205. doi: https://dx.doi.org/10.1016/j.drugalcdep.2020.108205.

Fairbairn, C. E., I. G. Rosen, S. E. Luczak, and W. J. Venerable. 2019. "Estimating the Quantity and Time Course of Alcohol Consumption from Transdermal Alcohol Sensor Data: A Combined Laboratory-Ambulatory Study." *Alcohol* 01:1. doi:

https://dx.doi.org/10.1016/j.alcohol.2018.08.015.

Flango, Victor E., and Fred L. Cheesman. 2009. “Effectiveness of the SCRAM Alcohol Monitoring Device: A Preliminary Test.” *Drug Court Review* 6(2):109–34.

Gibb, Kenneth A., Albert S. Yee, Charles C. Johnston, Steven D. Martin, and Richard M. Nowak. 1984. “Accuracy and Usefulness of a Breath Alcohol Analyzer.” *Annals of Emergency Medicine* 13(7):516–20. doi: 10.1016/S0196-0644(84)80517-X.

Gilligan, Conor, Kristen G. Anderson, Benjamin O. Ladd, Yun Ming Yong, and Michael David. 2019. “Inaccuracies in Survey Reporting of Alcohol Consumption.” *BMC Public Health* 19(1):1–11. doi: 10.1186/S12889-019-7987-3/TABLES/4.

Glassman, Tavis, Chudley E. Werch, Edessa Jobli, and Hui Bian. 2007. “Alcohol-Related Fan Behavior on College Football Game Day.” *Journal of American College Health* 56(3):255–60.

Gmel, Gerhard, Emmanuel Kuntsche, and Jürgen Rehm. 2011. “Risky Single-Occasion Drinking: Bingeing Is Not Bingeing.” *Addiction* 106(6):1037–45. doi: 10.1111/J.1360-0443.2010.03167.X.

Gmel, Gerhard, and Jürgen Rehm. 2004. “Measuring Alcohol Consumption:” *Contemporary Drug Problems* 31(3):467–540. doi: 10.1177/009145090403100304.

Greenfield, Thomas K., and William C. Kerr. 2008. “Alcohol Measurement Methodology in Epidemiology: Recent Advances and Opportunities.” *Addiction (Abingdon, England)* 103(7):1082. doi: 10.1111/J.1360-0443.2008.02197.X.

Gunn, R. L., J. A. Steingrimsson, J. E. Merrill, T. Souza, and N. Barnett. 2021. “Characterising Patterns of Alcohol Use among Heavy Drinkers: A Cluster Analysis Utilising Alcohol Biosensor Data.” *Drug & Alcohol Review* 40(7):1155–64. doi: https://dx.doi.org/10.1111/dar.13306.

Hammersley, Martyn. 1987. “Some Notes on the Terms ‘Validity’ and ‘Reliability.’” *British Educational Research Journal* 13(1):73–82. doi: 10.1080/0141192870130107.

Heale, Roberta, and Alison Twycross. 2015. “Validity and Reliability in Quantitative Studies.” *Evidence-Based Nursing* 18(3):66–67. doi: 10.1136/EB-2015-102129.

Hill-Kapturczak, Nathalie, Sarah L. Lake, John D. Roache, Sharon E. Cates, Yuanyuan Liang, and Donald M. Dougherty. 2014. “Do Variable Rates of Alcohol Drinking Alter the Ability to Use Transdermal Alcohol Monitors to Estimate Peak Breath Alcohol and Total Number of Drinks?” *Alcoholism: Clinical and Experimental Research* 38(10):2517–22. doi: http://dx.doi.org/10.1111/acer.12528.

Hill-Kapturczak, Nathalie, John D. Roache, Yuanyuan Liang, Tara E. Karns, Sharon E. Cates, and Donald M. Dougherty. 2015. “Accounting for Sex-Related Differences in the Estimation of Breath Alcohol Concentrations Using Transdermal Alcohol Monitoring.” *Psychopharmacology* 232(1):115–23. doi: 10.1007/s00213-014-3644-9.

Hufford, Michael R., Alan L. Shields, Saul Shiffman, Jean Paty, and Mark Balabanis. 2002. “Reactivity to Ecological Momentary Assessment: An Example Using Undergraduate Problem Drinkers.” *Psychology of Addictive Behaviors* 16(3):205–11. doi: 10.1037/0893-164X.16.3.205.

Hurzeler, Tristan, Nicholas A. Buckley, Firouzeh Noghrehchi, Peter Malouf, Andrew Page, Jennifer L. Schumann, and Kate M. Chitty. 2021. “Alcohol-Related Suicide across Australia: A Geospatial Analysis.” *Australian and New Zealand Journal of Public Health* 45(4):394–99. doi: 10.1111/1753-6405.13122.

Institute, National Drug Research. n.d. “Social and Economic Costs of Substance Use.” *2021*.

Johnson, Richard D., Michael Morowitz, Anne F. Maddox, David J. Shearman, and North Terrace. 1991. “Cigarette Smoking and Rate of Gastric Emptying: Effect on Alcohol Absorption.” *BMJ : British Medical Journal* 302(6767):20. doi: 10.1136/BMJ.302.6767.20.

Jones, Andrew, Danielle Remmerswaal, Ilse Verveer, Eric Robinson, Ingmar H. A. Franken, Cheng K. Fred Wen, and Matt Field. 2018. “Compliance with Ecological Momentary Assessment Protocols in Substance Users: A Meta-Analysis.” doi: 10.1111/add.14503.

Kalist, David E., and Daniel Y. Lee. 2016. “The National Football League: Does Crime Increase on Game Day?” *Journal of Sports Economics* 17(8):863–82.

Karns-Wright, Tara E., Donald M. Dougherty, Nathalie Hill-Kapturczak, Charles W. Mathias, and John D. Roache. 2018. “The Correspondence between Transdermal Alcohol Monitoring and Daily Self-Reported Alcohol Consumption.” *Addictive Behaviors* 85:147–52. doi: 10.1016/j.addbeh.2018.06.006.

Karns-Wright, Tara E., John D. Roache, Nathalie Hill-Kapturczak, Yuanyuan Liang, Jillian Mullen, and Donald M. Dougherty. 2017. “Time Delays in Transdermal Alcohol Concentrations Relative to Breath Alcohol Concentrations.” *Alcohol and Alcoholism* 52(1):35–41. doi: http://dx.doi.org/10.1093/alcalc/agw058.

Kelly, Thomas M., John E. Donovan, Janet M. Kinnane, and David M. C. D. Taylor. 2002. “A COMPARISON OF ALCOHOL SCREENING INSTRUMENTS AMONG UNDER-AGED DRINKERS TREATED IN EMERGENCY DEPARTMENTS.” *Alcohol and Alcoholism* 37(5):444–50. doi: 10.1093/ALCALC/37.5.444.

Kerr, William C., Thomas K. Greenfield, Jennifer Tujague, and Stephan E. Brown. 2005. “A Drink Is a Drink? Variation in the Amount of Alcohol Contained in Beer, Wine and Spirits Drinks in a US Methodological Sample.” *Alcoholism: Clinical and Experimental Research* 29(11):2015–21. doi: 10.1097/01.ALC.0000187596.92804.BD.

Kerr, William C., Deidre Patterson, Mary Albert Koenen, and Thomas K. Greenfield. 2008. “Alcohol Content Variation of Bar and Restaurant Drinks in Northern California.” *Alcoholism, Clinical and Experimental Research* 32(9):1623–29. doi: 10.1111/J.1530-0277.2008.00741.X.

Khadjesari, Zarnie, Elizabeth Murray, Eleftheria Kalaitzaki, Ian R. White, Jim McCambridge, Christine Godfrey, and Paul Wallace. 2009. “Test–Retest Reliability of an Online Measure of Past Week Alcohol Consumption (the TOT-AL), and Comparison with Face-to-Face Interview.” *Addictive Behaviors* 34(4):337–42.

Kingsland, Melanie, Luke Wolfenden, Bosco C. Rowland, Karen E. Gillham, Vanessa J. Kennedy, Robyn L. Ramsden, Richard W. Colbran, Sarah Weir, and John H. Wiggers. 2013. “Alcohol Consumption and Sport: A Cross-Sectional Study of Alcohol Management Practices Associated with at-Risk Alcohol Consumption at Community Football Clubs.” *BMC Public Health* 13(1):1–9.

Kuntsche, E., and F. Labhart. 2014. “The Future Is Now--Using Personal Cellphones to Gather Data on Substance Use and Related Factors.” *Addiction* 109(7):1052–53. doi: 10.1111/ADD.12540.

Kuntsche, Emmanuel, and Raimondo Bruno. 2015. “Moody Booze: Introducing the Special Section on Affect Regulation and Substance Use.” *Drug and Alcohol Review* 34(6):569–70. doi: 10.1111/DAR.12349.

Kuntsche, Emmanuel, Paul Dietze, and Rebecca Jenkinson. 2014. “Understanding Alcohol and Other Drug Use during the Event.” *Drug and Alcohol Review* 33(4):335–37. doi: 10.1111/DAR.12171.

Kuntsche, Emmanuel, and Florian Labhart. 2012. “Investigating the Drinking Patterns of Young People over the Course of the Evening at Weekends.” *Drug and Alcohol Dependence* 124(3):319–24. doi: 10.1016/J.DRUGALCDEP.2012.02.001.

Kuntsche, Emmanuel, and Florian Labhart. 2013a. “Using Personal Cell Phones for Ecological Momentary Assessment: An Overview of Current Developments.” *European Psychologist* 18(1):3–11. doi: 10.1027/1016-9040/A000127.

Kuntsche, Emmanuel, and Florian Labhart. 2013b. “Using Personal Cell Phones for Ecological Momentary Assessment.” *European Psychologist*.

Labhart, Florian, Skanda Muralidhar, Benoit Massé, Lakmal Meegahapola, Emmanuel Kuntsche, and Daniel Gatica-Perez. 2021. “Ten Seconds of My Nights: Exploring Methods to Measure Brightness, Loudness and Attendance and Their Associations with Alcohol Use from Video Clips.” *PLOS ONE* 16(4):e0250443. doi: 10.1371/JOURNAL.PONE.0250443.

Labhart, Florian, Thanh Trung Phan, Daniel Gatica-Perez, and Emmanuel Kuntsche. 2021. “Shooting

Shots: Estimating Alcoholic Drink Sizes in Real Life Using Event-Level Reports and Annotations of Close-up Pictures." *Drug and Alcohol Review* 40(7):1228–38. doi: 10.1111/DAR.13212.

Labhart, Florian, Darshan Santani, Jasmine Truong, Flavio Tarsetti, Olivier Bornet, Sara Landolt, Daniel Gatica-Perez, and Emmanuel Kuntsche. 2017. "Development of the Geographical Proportional-to-Size Street-Intercept Sampling (GPSIS) Method for Recruiting Urban Nightlife-Goers in an Entire City." *Https://Doi.Org/10.1080/13645579.2017.1293928* 20(6):721–36. doi: 10.1080/13645579.2017.1293928.

Labhart, Florian, Flavio Tarsetti, Olivier Bornet, Darshan Santani, Jasmine Truong, Sara Landolt, Daniel Gatica-Perez, and Emmanuel Kuntsche. 2019. "Capturing Drinking and Nightlife Behaviours and Their Social and Physical Context with a Smartphone Application – Investigation of Users' Experience and Reactivity." *Https://Doi.Org/10.1080/16066359.2019.1584292* 28(1):62–75. doi: 10.1080/16066359.2019.1584292.

Lajunen, Timo, and Heikki Summala. 2003. "Can We Trust Self-Reports of Driving? Effects of Impression Management on Driver Behaviour Questionnaire Responses." *Transportation Research Part F: Traffic Psychology and Behaviour* 6(2):97–107.

Lansdorp, Bob, William Ramsay, Rashad Hamid, and Evan Strenk. 2019. "Wearable Enzymatic Alcohol Biosensor." *Sensors* 19(10):2380.

Laslett, Anne-Marie, Paul Catalano, Tanya Chikritzhs, Caroline Dale, Christopher Doran, Jason Ferris, Thameemul Jainullabudeen, Michael Livingston, Sharon Matthews, Janette Mugavin, Robin Room, Morgan Schlotterlein, and Claire Wilkinson. 2010. "The Range and Magnitude of Alcohol's Harm to Others." *Victoria: AER Centre for Alcohol Policy Research, Turning Point Alcohol and Drug Centre, Eastern Health*.

Lauckner, Carolyn, Erica Taylor, Darshti Patel, and Alexis Whitmire. 2019. "The Feasibility of Using Smartphones and Mobile Breathalyzers to Monitor Alcohol Consumption among People Living

with HIV/AIDS." *Addiction Science and Clinical Practice* 14(1):1–11. doi: 10.1186/S13722-019-0174-0/TABLES/3.

Lawson, B., K. Aguir, T. Fiorido, V. Martini-Laithier, R. Bouchakour, S. Burtey, C. Reynard-Carette, and M. Bendahan. 2019. "Skin Alcohol Perspiration Measurements Using MOX Sensors." *Sensors and Actuators, B: Chemical* 280:306–12. doi: 10.1016/j.snb.2018.09.082.

Leffingwell, Thad R., Nathaniel J. Cooney, James G. Murphy, Susan Luczak, Gary Rosen, Donald M. Dougherty, and Nancy P. J. Barnett. 2013. "Continuous Objective Monitoring of Alcohol Use: Twenty-first Century Measurement Using Transdermal Sensors." *Alcoholism: Clinical and Experimental Research* 37(1):16–22.

Li, B., R. S. Downen, Q. Dong, N. Tran, M. LeSaux, A. C. Meltzer, and Z. Li. 2021. "A Discreet Wearable IoT Sensor for Continuous Transdermal Alcohol Monitoring - Challenges and Opportunities." *IEEE Sensors Journal* 21(4):5322–30. doi: https://dx.doi.org/10.1109/jsen.2020.3030254.

Lin, K. C., D. Kinnamon, D. Sankhala, S. Muthukumar, and S. Prasad. 2019. "AWARE: A Wearable Awareness with Real-Time Exposure, for Monitoring Alcohol Consumption Impact through Ethyl Glucuronide Detection." *Alcohol* 81:93–99. doi: https://dx.doi.org/10.1016/j.alcohol.2018.10.006.

Livingston, Michael, and Sarah Callinan. 2015. "Underreporting in Alcohol Surveys: Whose Drinking Is Underestimated?" *Journal of Studies on Alcohol and Drugs* 76(1):158–64.

Lloyd, Belinda, Sharon Matthews, Michael Livingston, Harindra Jayasekara, and Karen Smith. 2013. "Alcohol Intoxication in the Context of Major Public Holidays, Sporting and Social Events: A Time–Series Analysis in M Elbourne, A Ustralia, 2000–2009." *Addiction* 108(4):701–9.

Luczak, S. E., and I. G. Rosen. 2014. "Estimating BrAC from Transdermal Alcohol Concentration Data Using the BrAC Estimator Software Program." *Alcoholism: Clinical and Experimental Research* 38(8):2243–52. doi: 10.1111/acer.12478.

Luczak, S E, I. G. Rosen, and T. L. Wall. 2015. “Development of a Real-Time Repeated-Measures Assessment Protocol to Capture Change over the Course of a Drinking Episode.” *Alcohol and Alcoholism* 50(2):180–87. doi: 10.1093/alcalc/agu100.

Luczak, Susan E., I. Gary Rosen, and Tamara L. Wall. 2015. “Development of a Real-Time Repeated-Measures Assessment Protocol to Capture Change over the Course of a Drinking Episode.” *Alcohol and Alcoholism* 50(2):180–87. doi: 10.1093/ALCALC/AGU100.

Maher, Jessica Middlemis, Jonathan C. Markey, and Diane Ebert-May. 2013. “The Other Half of the Story: Effect Size Analysis in Quantitative Research.” *CBE—Life Sciences Education* 12(3):345–51.

Marques, Paul R., and A. Scott McKnight. 2007. *Evaluating Transdermal Alcohol Measuring Devices*. United States. National Highway Traffic Safety Administration.

Marques, Paul R., and A. Scott McKnight. 2009. “Field and Laboratory Alcohol Detection with 2 Types of Transdermal Devices.” *Alcoholism: Clinical and Experimental Research* 33(4):703–11.

May, Carl, and David Foxcroft. 1995. “Minimizing Bias in Self-Reports of Health Beliefs and Behaviours.” *Health Education Research* 10(1):107–12.

Meier, Patrick, and Helmut K. Seitz. 2008. “Age, Alcohol Metabolism and Liver Disease.” *Current Opinion in Clinical Nutrition & Metabolic Care* 11(1):21–26.

Miller, Elizabeth T., Dan J. Neal, Lisa J. Roberts, John S. Boer, Sally O. Cresslr, Jane Metrik, and G. Alan Marlatt. 2002. “Test-Retest Reliability of Alcohol Measures: Is There a Difference between Internet-Based Assessment and Traditional Methods?” *Psychology of Addictive Behaviors* 16(1):56–63.

Miller, Peter, Louise M c Donald, Stephen M c Kenzie, Kerry O’Brien, and Petra Staiger. 2013. “When the C Ats Are Away: The Impact of Sporting Events on Assault-and Alcohol-related Emergency Department Attendances.” *Drug and Alcohol Review* 32(1):31–38.

Monk, Rebecca Louise, Derek Heim, Adam Qureshi, and Alan Price. 2015. "'I Have No Clue What I Drunk Last Night' Using Smartphone Technology to Compare In-Vivo and Retrospective Self-Reports of Alcohol Consumption." *PLOS ONE* 10(5):e0126209. doi: 10.1371/JOURNAL.PONE.0126209.

Mun, Eun-Young, Xiaoyin Li, Michael S. Businelle, Emily T. Hébert, Zhengqi Tan, Nancy P. Barnett, and Scott T. Walters. 2021. "Ecological Momentary Assessment of Alcohol Consumption and Its Concordance with Transdermal Alcohol Detection and Timeline Follow-Back Self-report Among Adults Experiencing Homelessness." *Alcoholism: Clinical and Experimental Research* 45(4):864–76.

Nelson, Toben F., and Henry Wechsler. 2003. "School Spirits: Alcohol and Collegiate Sports Fans." *Addictive Behaviors* 28(1):1–11.

NHMRC. 2020. "Australian Guidelines to Reduce Health Risks from Drinking Alcohol | NHMRC." Retrieved September 10, 2021 (https://www.nhmrc.gov.au/about-us/publications/australian-guidelines-reduce-health-risks-drinking-alcohol).

Niall Bolger, and Jean-Philippe Laurenceau. 2013. "Intensive Longitudinal Methods: An Introduction to Diary and Experience." *Guilford, New York*. Retrieved January 15, 2022 (https://books.google.com.au/books?hl=en&lr=&id=5bD4LuAFq0oC&oi=fnd&pg=PP1&ots=jnjkZ0GU92&sig=yBv78VkSIjd4yYvE_AHdT8siCJA&redir_esc=y#v=onepage&q&f=false).

Norman, Thomas, Amy Peacock, Stuart G. Ferguson, Emmanuel Kuntsche, and Raimondo Bruno. 2020. "Combining Transdermal and Breath Alcohol Assessments, Real-Time Drink Logs and Retrospective Self-Reports to Measure Alcohol Consumption and Intoxication across a Multi-Day Music Festival." *Drug and Alcohol Review*. doi: 10.1111/DAR.13215.

Palmer, Catherine. 2011. "Key Themes and Research Agendas in the Sport-Alcohol Nexus." *Journal of Sport and Social Issues* 35(2):168–85.

Palmer, Catherine, and Kirrilly Thompson. 2007. "The Paradoxes of Football Spectatorship: On-Field

and Online Expressions of Social Capital among the 'Grog Squad.'" *Sociology of Sport Journal* 24(2):187–205.

Peacock, Amy, Raimondo Bruno, Frances H. Martin, and Andrea Carr. 2013. "The Impact of Alcohol and Energy Drink Consumption on Intoxication and Risk-Taking Behavior." *Alcoholism: Clinical and Experimental Research* 37(7):1234–42. doi: 10.1111/ACER.12086.

Piasecki, Thomas M. 2019. "Assessment of Alcohol Use in the Natural Environment." *Alcoholism: Clinical and Experimental Research* 43(4):564–77.

Riordan, Benjamin C., Damian Scarf, Saleh Moradi, Jayde A. M. Flett, Kate B. Carey, and Tamlin S. Conner. 2017. "The Accuracy and Promise of Personal Breathalysers for Research: Steps toward a Cost-Effective Reliable Measure of Alcohol Intoxication?" *Digital Health* 3:2055207617746752.

Riuttanen, Antti, Saara J. Jäntti, and Ville M. Mattila. 2020. "Alcohol Use in Severely Injured Trauma Patients." *Scientific Reports 2020 10:1* 10(1):1–8. doi: 10.1038/s41598-020-74753-y.

Roache, J. D., T. E. Karns-Wright, M. Goros, N. Hill-Kapturczak, C. W. Mathias, and D. M. Dougherty. 2019. "Processing Transdermal Alcohol Concentration (TAC) Data to Detect Low-Level Drinking." *Alcohol* 01:1. doi: https://dx.doi.org/10.1016/j.alcohol.2018.08.014.

Roache, John D., Tara E. Karns, Nathalie Hill-Kapturczak, Jillian Mullen, Yuanyuan Liang, Richard J. Lamb, and Donald M. Dougherty. 2015. "Using Transdermal Alcohol Monitoring to Detect Low-Level Drinking." *Alcoholism: Clinical and Experimental Research* 39(7):1120–27. doi: http://dx.doi.org/10.1111/acer.12750.

Sakai, Joseph T., Susan K. Mikulich-Gilbertson, Robert J. Long, and Thomas J. Crowley. 2006. "Validity of Transdermal Alcohol Monitoring: Fixed and Self-regulated Dosing." *Alcoholism: Clinical and Experimental Research* 30(1):26–33.

Saldich, Emily B., Chunming Wang, I. Gary Rosen, Jay Bartroff, and Susan E. Luczak. 2021. "Effects of Stomach Content on the Breath Alcohol Concentration-Transdermal Alcohol Concentration

Relationship." *Drug and Alcohol Review*. doi: 10.1111/DAR.13267.

Schafer, Theodore, and A. Walls Joseph L. 2006. *Models for Intensive Longitudinal Data* . Oxford University Press.

Shiffman, Saul. 2009. "Ecological Momentary Assessment (EMA) in Studies of Substance Use." *Psychological Assessment* 21(4):486.

Shiffman, Saul, Arthur A. Stone, and Michael R. Hufford. 2008. "Ecological Momentary Assessment." *Annu. Rev. Clin. Psychol.* 4:1–32.

Simons, Jeffrey S., Thomas A. Wills, Noah N. Emery, and Russell M. Marks. 2015. "Quantifying Alcohol Consumption: Self-Report, Transdermal Assessment, and Prediction of Dependence Symptoms." *Addictive Behaviors* 50:205–12. doi: http://dx.doi.org/10.1016/j.addbeh.2015.06.042.

Sirlanci, M., I. G. Rosen, T. L. Wall, and S. E. Luczak. 2018. "Applying a Novel Population-Based Model Approach to Estimating Breath Alcohol Concentration (BrAC) from Transdermal Alcohol Concentration (TAC) Biosensor Data." *Alcohol* 20:20. doi: https://dx.doi.org/10.1016/j.alcohol.2018.09.005.

Sobell, Linda C., Mark B. Sobell, Gloria I. Leo, and Anthony Cancilla. 1988. "Reliability of a Timeline Method: Assessing Normal Drinkers' Reports of Recent Drinking and a Comparative Evaluation across Several Populations." *British Journal of Addiction* 83(4):393–402.

Sobell, Mark B., Linda C. Sobell, Felix Klajner, Daniel Pavan, and Ellen Basian. 1986. "The Reliability of a Timeline Method for Assessing Normal Drinker College Students' Recent Drinking History: Utility for Alcohol Research." *Addictive Behaviors* 11(2):149–61.

Stockwell, Tim, Susan Donath, Mark Cooper-Stanbury, Tanya Chikritzhs, Paul Catalano, and Cid Mateo. 2004. "Under-Reporting of Alcohol Consumption in Household Surveys: A Comparison of Quantity-Frequency, Graduated-Frequency and Recent Recall." *Addiction (Abingdon, England)* 99(8):1024–33. doi: 10.1111/J.1360-0443.2004.00815.X.

Swette, L., A. Griffith, and A. La Conti. 1997. “Potential and Diffusion Controlled Solid Electrolyte Sensor for Continuous Measurement of Very Low Levels of Transdermal Alcohol.” *Giner, Inc., Waltham, MA. Appl. No. 840802*.

Swift, Robert. 2003. “Direct Measurement of Alcohol and Its Metabolites.” *Addiction* 98(SUPPL. 2):73–80. doi: 10.1046/J.1359-6357.2003.00605.X.

Swift, Robert M., Christopher S. Martin, Larry Swette, Anthony LaConti, and Nancy Kackley. 1992. “Studies on a Wearable, Electronic, Transdermal Alcohol Sensor.” *Alcoholism: Clinical and Experimental Research* 16(4):721–25. doi: 10.1111/J.1530-0277.1992.TB00668.X.

Trull, Timothy J., and Ulrich Ebner-Priemer. 2014. “The Role of Ambulatory Assessment in Psychological Science.” *Current Directions in Psychological Science* 23(6):466–70. doi: 10.1177/0963721414550706.

Vestal, Robert E., Elizabeth A. McGuire, Jordan D. Tobin, Reubin Andres, Arthur H. Norris, and Esteban Mezey. 1977. “Aging and Ethanol Metabolism.” *Clinical Pharmacology & Therapeutics* 21(3):343–54. doi: 10.1002/CPT1977213343.

Voas, Robert B., Robert L. DuPont, Stephen K. Talpins, and Corinne L. Shea. 2011. “Towards a National Model for Managing Impaired Driving Offenders.” *Addiction* 106(7):1221–27. doi: 10.1111/J.1360-0443.2010.03339.X.

Voogt, Carmen, Emmanuel Kuntsche, Marloes Kleinjan, Evelien Poelen, and Rutger Engels. 2014. “Using Ecological Momentary Assessment to Test the Effectiveness of a Web-Based Brief Alcohol Intervention Over Time Among Heavy-Drinking Students: Randomized Controlled Trial.” *J Med Internet Res 2014;16(1):E5 Https://Www.Jmir.Org/2014/1/E5* 16(1):e2817. doi: 10.2196/JMIR.2817.

Wolfe, Jeannette, Ricardo Martinez, and Warren A. Scott. 1998. “Baseball and Beer: An Analysis of Alcohol Consumption Patterns among Male Spectators at Major-League Sporting Events.” *Annals of Emergency Medicine* 31(5):629–32.

Wray, Tyler B., Jennifer E. Merrill, and Peter M. Monti. 2014. “Using Ecological Momentary Assessment (EMA) to Assess Situation-Level Predictors of Alcohol Use and Alcohol-Related Consequences.” *Alcohol Research: Current Reviews* 36(1):19.

Wright, Cassandra, Paul M. Dietze, Paul A. Agius, Emmanuel Kuntsche, Michael Livingston, Oliver C. Black, Robin Room, Margaret Hellard, and Megan S. C. Lim. 2018. “Mobile Phone-Based Ecological Momentary Intervention to Reduce Young Adults’ Alcohol Use in the Event: A Three-Armed Randomized Controlled Trial.” *JMIR MHealth and UHealth* 6(7):e149.

Züll, C. 2016. “Open-Ended Questions.” *GESIS Survey Guidelines* 3.

Table 4

Participant Demographics

	N = 13
Sex (% female)	2 (15.4%)
Age (years)	32.1 (13.1)
Ethnicity	
Caucasian	12 (92.3%)
Frequency of watching an AFL match	
Once a week or more	11 (84.6%)
Once a fortnight to once a month	2 (15.4%)
Level of risky drinking	
Lower risk	0 (0.0%)
Moderate risk	4 (30.8%)
High risk	6 (46.2%)
Possible dependence	3 (23.1%)

Note. This table shows the demographics of the final sample N = 13 participants. Data is reported as the number of participants (percentage of final sample), except from age which is reported as the mean (standard deviatio).

*Level of risky drinking classification is determined by using AUDIT-C scores:

Women; 0-2, 3-5, 6-7, 8-12, is categorized as low, moderate, high risk, and possible dependence, respectively.

Men; 0-3, 4-5, 6-7, 8-12, is categorized as low, moderate, high risk, and possible dependence, respectively.

Table 5

Drinking events reported and identified: means and range.

	Mean (SD)	**Median**	**Range**
EMA			
Number of drinks	5.0 (4.9)	2.0	0 - 20
Time of event (hrs)	2.3 (2.7)	1.5	0 - 9
TAC			
Peak	0.051 (0.085)	0.009	0.000 - 0.316
AUC	0.352 (0.873)	0.020	0.000 – 3.900
Time of event (hrs)	5.6 (6.4)	4	0 – 27.5

Note. Shown are means with standard deviations in brackets.

Table 6

Agreement between self-reported and TAC drinking categories.

		Self-reported events category			**Total**
		None	Moderate	Heavy	
TAC events category	None	2 (40.0%)	7 (46.7%)	0 (0.0%)	**9 (31.0%)**
	Low	1 (20.0%)	3 (20.0%)	0 (0.0%)	**4 (13.8%)**
	Moderate	2 (40.0%)	2 (13.3%)	1 (11.1%)	**5 (17.2%)**
	Heavy	0 (0.0%)	3 (20.0%)	8 (88.9%)	**11 (37.9%)**
	Total	**5 (17.2%)**	**15 (51.7%)**	**9 (31.0%)**	**29**
	AMS confirmed	0 (0.0%)	0 (0.0%)	7 (77.8%)	**7**

Note. Numbers represent the number of drinking events detected by the SCRAM-CAM, or self-reported. Percentages show the percentage of TAC events in a certain self-reported category. Pearson Chi-square analysis: $\chi 2$ (6) = 17.1, p=0.009), phi coefficient is considered strong (φ = 0.767).

TAC events were classified into the following categories: None: >1 but < 2 non-zero TAC points; Low: 3 or more TAC readings > 0, but no readings > .01 g/dl; Moderate: ≥ 3 TAC points above 0 and ≥ 1 TAC point above 0.01 g/dl but < 2 points above 0.02 g/dl; Heavy: 2 or more readings > 0.02 g/dl.

Self-reported Classifications: None: no drinking; Moderate:>0 drinks but<5 drinks for men/4 drinks for women; and Heavy:>5 drinks for men/>4 drinks for women.

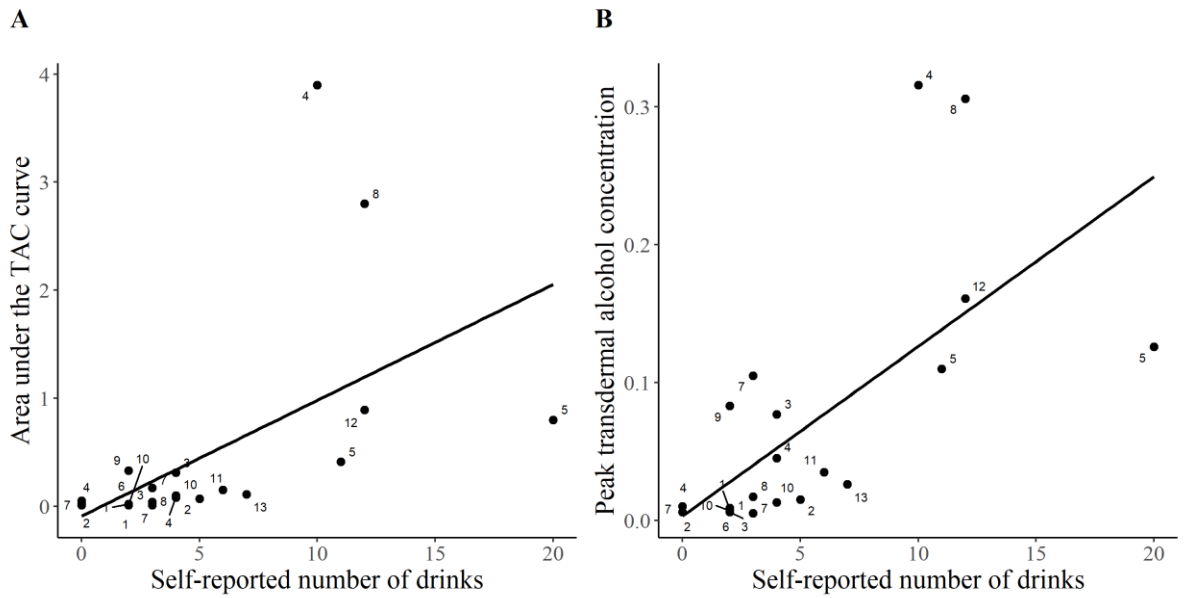

Figure 5. **A.** Scatterplot of AUC values on self-reported number of drinks**, B.** Scatterplot of TAC peak values on self-reported number of drinks.

8. Discussion

The aim of my research was to advance our understanding of TAC monitoring technology, and to provide further evidence on the test–retest reliability and validity of the SCRAM-CAM ankle bracelet and the ION RAP wristband in both controlled and naturalistic environments. Test–retest reliability refers to the consistency of a measure, and includes two notions: reliability and agreement (Berchtold 2016; Heale and Twycross 2015). High reliability means that when identical monitors are fitted in parallel on the same participant and in the same conditions, they produce similar measurement patterns (Berchtold 2016; Cohen and Vinson 1995; Heale and Twycross 2015). Agreement requires the monitors to provide similar (in the best case, identical) values (Berchtold 2016; Bland and Altman 1986; Riordan et al. 2017). Validity, specifically construct or convergent validity, is the extent to which wearable monitors accurately measure the amount of alcohol (Heale and Twycross 2015). Determining validity involves studying the agreement between the monitored result and a reference measure of alcohol consumption (e.g., BrAC or self-reports) (Csikszentmihalyi et al. 2014; Heale and Twycross 2015; Sakai et al. 2006). (Hammersley 1987; Heale and Twycross 2015; Sakai et al. 2006).

To achieve my aims I completed three major research tasks, as outlined in this thesis:

1) systematically review the literature on test–retest reliability and validity of TAC monitors and identify research gaps;

2) study the test–retest reliability and validity of the SCRAM-CAM ankle bracelet and conduct the first controlled laboratory study of a prototype of the ION RAP wristband; and

3) investigate the validity of the SCRAM-CAM under naturalistic settings by comparing the TAC data against ecological momentary assessments.

Although TAC monitors have been available for almost two decades, rapid advances in technology have occurred in the past 10 years. In Chapter 3, I summarised and discussed the findings

of studies published after 2013 that validated TAC monitors. My systematic review in 2019 identified 13 validation studies of three TAC monitors: the SCRAM, WrisTAS and the BACtrack Skyn. Since the systematic review in Chapter 3 was conducted, six more validation studies have been published, and another four such studies are reported in this thesis. In the following sections, I integrate the findings presented in this thesis with the broader literature, including the validation studies published since the systematic review. I also discuss the strengths and limitations of my research, and give recommendations for further research.

Key findings

The key findings that emerged from Chapters 3–7 are summarised in Table 8-1.

Table 8-7. Overview of key findings of Chapters 3–7

Ch.	Aims	Findings
3	Systematic literature review	• Correlations between TAC and BrAC and/or self-reported data ranged from 0.59 to 0.91 for the WrisTAS, 0.56–0.94 for the SCRAM-CAM, and 0.77–0.79 for the BACtrack Skyn. • All monitors, but especially the SCRAM-CAM, detected low-to-moderate-level drinking episodes poorly. • Failure rates for the WrisTAS and BACtrack Skyn were high. • No researchers have investigated TAC monitor test–retest reliability using repeated measures.
4	Test reliability of two SCRAM-CAMs in parallel	• Despite a positive relationship between TAC measurements from SCRAM-CAMs worn on the left and right ankle, the correlations were lower than expected (0.55–0.70), suggesting significant measurement error. • This relationship was unaffected by individual characteristics such as sex and BMI.
5	Testing agreement of SCRAM-CAM TAC, and reliability and agreement of the ION RAP TAC	• SCRAM-CAM: mean absolute difference of 0.010 g/dL, with 14% of the TAC measurements showing differences higher than 0.020 g/dL. • Food consumption and drinking rate had no effect on the level of agreement between left and right SCRAM-CAM TAC. • ION RAP reliability was found to be excellent, with an ICC of 0.81.

		• ION RAP: Mean difference of 131 (SD = 94.4) nA, 30% of the points fell above our acceptable limit.
6	Validation of ION RAP and SCRAM-CAM against BrAC	• Failure rate of 61.5% for the ION RAP wristband. • No significant differences in delays for ION RAP and SCRAM-CAM; ~45 minutes to first detection of alcohol and ~1.5 hours behind BrAC peak levels to reach TAC peak levels. • SCRAM-CAM TAC showed small to large correlations to BrAC (peak: r = 0.193, AUC: r = 0.466). • ION RAP showed close to zero correlations to BrAC (peak: r = -0.092, AUC: r = 0.016). • maximal cross-correlations between BrAC and TAC were large (SCRAM-CAM: r = 0.63, r = ION RAP: 0.67).
7	Compare EMA SCRAM-CAM TAC over the course of a day	• The correspondence between reported and identified drinking events was good, with an ICC of 0.62. • TAC values were highly correlated to the self-reported number of drinks (r=0.55–0.67). • Participants reported experiencing discomfort while wearing the SCRAM-CAM ankle bracelet. • Participants who completed EMA surveys during AFL games reported them to be burdensome.

Note. TAC = Transdermal Alcohol Concentration, BrAC = Breath Alcohol Concentration, BMI = Body Mass Index, SCRAM-CAM= Secure Continuous Alcohol Monitoring Continuous Alcohol Monitoring, EMA = Ecological Momentary Assessment, AUC = Area Under the Curve, ICC = Intra-Class Correlation.

Discussion of the main findings and limitations of transdermal alcohol monitoring

Test–retest reliability

Test–retest reliability is the extent to which applying the same tool to the same participants twice under the same circumstances results in similar measurements and conclusions (Heale and Twycross 2015; Khadjesari et al. 2009; Miller et al. 2002). This can be assessed through two concepts: reliability and agreement. In Chapter 3, I presented a review of the literature on TAC monitors, finding that the TAC as measured by the SCRAM was a valid measure of alcohol consumption, with medium to large correlations to the self-reported number of drinks and BrAC, but little research has assessed test–retest reliability. Swift et al. (1992) found that the TAC values as measured by two WrisTAS monitors worn simultaneously were highly correlated (peak TAC r = .71, AUC r = .94). However, no researcher (at the time of writing and to the best of my knowledge), has studied the reliability of currently available wearable monitors. Swift et al. (1992) chose to place the monitor on different locations on the arm, while more recently published studies (Luczak and Rosen 2014; Saldich et al. 2021) involved participants wearing TAC monitors on both the left and right wrist. These studies reported observed differences in peak TAC measurements; however, they did not actually analyse this difference statistically, or report the correlations between their datasets. The work presented in Chapter 4 revealed that the TAC data produced by the left and right SCRAM-CAMs were reliable, with ICC coefficients ranging between 0.56 and 0.72, but less reliable than expected. Given that these monitors measured TAC in the same conditions, on the same participants, with the only difference being placement on the right or left leg, we expected that the correlation coefficients would have been closer to one, as has been reported for breathalyser data (r = 0.99) (Riordan et al. 2017) and self-reported data (up to r = 0.99) (Khadjesari et al. 2009; Miller et al. 2002; Sobell et al. 1986, 1988).

My systematic review found only one validation study of a new-generation wristband (the BACtrack Skyn). Two studies of other monitors – the Milo ION™ (Lansdorp et al. 2019) and the Metal OXide (MOX) sensor (Lawson et al. 2019) – have been published, but did not calculate test–retest reliability. In the research presented in Chapter 5, therefore, I assessed the test–retest reliability of a prototype of the ION RAP (ION Wearables Inc. [previously Milo ION]). The ION RAP TAC data (left and right wrists) I obtained was highly correlated, and thus the device had good reliability. These correlation coefficients give evidence of how similar the alcohol measurement patterns are as taken by both the left and right monitor during the drinking event; however, they do not indicate the reproducibility of the data. So, I then assessed the agreement between TAC measures for both the SCRAM-CAM and the ION RAP, obtaining results that point to the need for improved TAC measurement. For the SCRAM-CAMs, there was an absolute mean difference between the TAC measures of 0.010 g/dL, with the difference of 14% of the matched left and right TAC measurements being higher than 0.020 g/dL, which is of a similar magnitude to the consumption of several drinks. The ION RAP left and right readings showed an absolute difference of 131 nA, with 30% of paired readings (for the same participant) differing by more than our predetermined threshold. The discrepancies between the left–right measurements for both monitors were found to increase with higher TAC levels. These results suggest that TAC monitors lack precision, a problem observed to be exacerbated at higher TAC levels.

The lower-than-expected reliability measures for the SCRAM-CAM, and the substantial disagreement between the paired left and right TAC readings as measured by both the SCRAM-CAM and the ION RAP, are concerning from a measurement standpoint because a variation of several drinks can result in very different conclusions in research and clinical settings (e.g., whether an individual performed risky drinking behaviour [consuming four standard drinks in a single drinking occasion] or non-risky drinking behaviour).

Finally, in Chapter 5 I reported a high failure rate of 61.5% for the ION RAP but 0% for the SCRAM-CAM, leading me to conclude that the SCRAM-CAM is more reliable than the ION RAP.

This aligns with Fairbairn and Kang (2019), who reported a much higher failure rate for the Skyn than the SCRAM-CAM (18–38% vs 2%), and hence recommended the SCRAM-CAM for use in research.

Validity

Previous studies validated TAC monitors by comparing the TAC data to either BrAC or self-reported data. The validity of TAC monitors is defined as the extent to which its outputs accurately measure the amount of alcohol, and as such includes studying the correspondence between TAC and other measures of alcohol consumption (e.g., BrAC or self-reports), but also the sensitivity and specificity of TAC monitors (Hammersley 1987; Heale and Twycross 2015; Sakai et al. 2006). Previous researchers reported correlations between TAC and BrAC and/or self-reported data ranging from 0.59 to 0.91 for the WrisTAS (Bond et al. 2014; Swift 2003; Swift et al. 1992), 0.56–0.94 for the SCRAM-CAM (Barnett et al. 2011; Dougherty et al. 2015; Fairbairn & Kang 2019; Hill-Kapturczak et al. 2014; Sakai et al. 2006) and 0.77–0.79 for the BACtrack Skyn (Fairbairn & Kang 2019). Sensitivity was observed to increase with higher TAC levels, and researchers consistently found the monitors had high specificity – especially the SCRAM devices (Karns-Wright et al. 2018; Marques & McKnight 2009; Roache et al. 2015, 2019). It is important to note the limitations of self-reports as the "gold standard" or reference measure; research has found self-reported measures to be unreliable due to biases related to cognitive decline at high alcohol consumption, social desirability, and underreporting (Davis et al. 2010; Livingston & Callinan 2015).

Chapter 6 describes my comparison of the validity of the SCRAM-CAM and a prototype of the ION RAP, relative to BrAC measurements. A key challenge when using TAC monitors to measure alcohol consumption is the delay in alcohol measurement. While breathalysers can detect alcohol almost immediately after consumption, in my research both the SCRAM-CAM and ION RAP detected alcohol only after 45 minutes on average, longer than the 23-minute delay reported in previous work (Fairbairn & Kang 2019). I reported that TAC peaked 1.5 hours after BrAC on average, with previous research reporting delays up to 4.5 hours (Fairbairn & Kang 2019; Sakai et al. 2006). In one study, Skyn TAC devices recorded peak values over an hour before SCRAM-CAMs

(Fairbairn & Kang 2019). The mechanisms behind the difference in latencies that were observed in these initial studies are unknown, but could be related to the devices used (sampling frequencies or the method of measurement) or to body positioning. In my study presented in Chapter 6, I did not find shorter lag times for a wristband than an ankle bracelet for the ION RAP. Thus, my work suggests that, independent of body positioning and technology (fuel cell or enzymatic detection), TAC monitors show consistent lag times due to the complexity of ethanol transport through the skin. Note that Fairbairn and Kang (2019) placed the Skyn on the inside of the wrist, while I placed the ION RAP on the outside of the wrist. Further research should investigate whether placing the ION RAP on the inside of the wrist (where the blood vessels are closer to the skin surface) reduces the latency of alcohol measurement.

The study presented in Chapter 6 also assessed the cross-correlation between the SCRAM-CAM and ION RAP TAC to BrAC; measuring the similarity of the TAC and BrAC curves over time by analysing the correlations between the two curves at different time lags, in order to estimate the displacement between the two curves. I found maximal cross-correlation coefficients ranging from 0.63 to 0.67, suggesting that the SCRAM-CAM and ION RAP TAC curves over time are very similar to the BrAC curves over time. However, I observed lower correlations between extracted TAC and BrAC curve characteristics (peak TAC: $r = .19$, AUC: $r = .47$) for the SCRAM-CAM than those published previously (Fairbairn & Kang 2019; Sakai et al. 2006). This may be due to two differences in protocols: first, previous studies used a restricted weight-based alcohol administration protocol intended to bring the participants up to a BrAC level of 0.08 (Fairbairn & Kang 2019; Sakai et al. 2006). Participants in my study drank the same amount of alcohol independent of their weight or gender (40 grams of alcohol in total; four Australian standard drinks). This would have caused a larger variation in BrAC peaks than in earlier research, and thus more divergence between the TAC and BrAC data. Restricted dosing is helpful when aiming to control for factors that could influence alcohol detection, such as gender; however, the dose-unrestricted results reported in Chapter 6 gave a better understanding of the validity of these devices in real-time alcohol measurement. Secondly, unlike in previously published studies (Fairbairn & Kang 2019; Sakai et al. 2006), I did not use a

control group that received no alcohol. Using a control group will result in a higher correlation between TAC and BrAC because neither the breathalyser nor the TAC monitor will measure any alcohol, resulting in a perfect correlation for this group. This can skew the results in favour of a higher correspondence between the measures. When pursuing other aims – for example, to test whether the TAC monitor is measuring alcohol or environmental noise –a control group could be important. However, since the SCRAM-CAM has been available, multiple studies have validated it in both controlled and naturalistic settings. An especially striking finding from the study described in Chapter 6 was the close to zero correlations between the ION RAP TAC and BrAC values (peak TAC: r = -0.09, AUC: r = 0.02). This, in combination with the high failure rates, suggests that the SCRAM-CAM is not only the more reliable monitor, but produces higher validity data than the ION RAP.

I compared SCRAM-CAM TAC data to the number of drinks self-reported via EMA (see Chapter 7), and found a good correspondence between the number of drinking events identified by the SCRAM-CAM and reported through EMA surveys (r = 0.62). Further, the TAC measures were highly correlated (peak TAC: r=0.673, AUC: r = 0.549) with the self-reported number of drinks. This is in accordance with the results in Chapter 3, and previously reported correlations of 0.79–0.94 between AUC and self-reported data (Barnett et al., 2011; Leffingwell et al., 2013), with a positive linear relationship between TAC and self-reported drinks (Dougherty, Karns et al., 2015; Hill-Kapturczak et al., 2014). Similarly, a recently published study found correlation coefficients ranging from 0.46 to 0.78 for TAC and EMA self-report data, with much lower correlations between TLFB results and those from EMA and TAC (Mun et al. 2021). This suggests that EMA surveys and TAC monitors are more valid than traditional retrospective surveys, such as the TLFB, in alcohol consumption measurement during a given timeframe. However, in Chapter 7, I also showed that the SCRAM-CAM missed almost half of the self-reported moderate drinking events (0–5 standard drinks), and using the SCRAM-CAM developer's (AMS) criteria resulted in missing all moderate drinking events, measuring only the heavy ones (>5 standard drinks). This is consistent with the literature reviewed in Chapter 3, in which, depending on the detection criteria used, the sensitivity of TAC monitors to measure low alcohol levels was shown to be very low. Using the AMS detection

criteria can miss up to five standard drinks (Karns-Wright et al. 2018; Roache et al. 2015, 2019). Additionally, a more recently published study that compared the TAC from the WrisTAS to the estimated BAC (eBAC) calculated from self-reported data documented larger differences between TAC and eBAC when blood alcohol levels rose (Croff et al. 2021). However, as described earlier, self-reported measures as used in Croff et al. (2021) tend to be not very reliable due to biases related to memory, social desirability and motivation that are even more pronounced at higher levels of alcohol consumption (Davis et al. 2010; Livingston and Callinan 2015). This means that this particular study could not determine conclusively whether higher discrepancies between the TAC and eBAC at higher levels of alcohol consumption are due to problems related to self-report or TAC.

Other validation research using the WrisTAS has been published recently. Blair et al. (2021) reported correlations between TAC and eBAC (calculated using self-reported measures) ranging between 0.33 and 0.57, which increased when blood alcohol concentration rose (Blair et al. 2021). These correlations are lower than the previously reported 0.62 (Bond et al. 2014) and the 85.7% detection rate (Simons et al. 2015) reported in Chapter 3. Neither Blair et al. (2021) nor Croff et al. (2021) reported monitor failures. (Note that WrisTAS monitors are no longer commercially available, which is why they receive little attention in this thesis.)

Factors influencing TAC detection ability and the relationship between TAC and BrAC

The factor with the most substantial effect on TAC readings is the number of drinks consumed (as reported or as controlled in a laboratory alcohol administration study), which is logical because the BAC increases with the amount of alcohol consumed. This is clear from the results reported in Chapter 3, and is supported by recent research that found consistently lower number of self-reported drinks for TAC drinking episodes with lower peak TAC and AUC levels as compared to episodes with higher peak TAC and AUC (Gunn et al. 2021). This phenomenon is reflected in Chapter 7, where I reported a positive relationship between SCRAM-CAM TAC and the self-reported number of drinks using EMA surveys. My work expanded on these findings because it examined the

effects of four factors: two individual (gender and BMI) and two contextual factors (food consumption and drinking rate), described in further detail below.

Gender

As reported in Chapter 3, debate about the effect of gender on TAC readings is ongoing. The study in Chapter 4 reported no significant effect of gender on the reliability of the TAC data measured by the left and right SCRAM-CAM, in line with recent studies reporting no differences in the peak TAC levels (Mun et al. 2021) and AUC values between women and men (Croff et al. 2021). In contrast, gender was found to play an important role in the TAC–BrAC relationship; because gender affects BrAC measures but not TAC measures, regression models converting TAC into BrAC estimates need to include a gender variable to account for this difference in effect (Dougherty et al. 2015; Hill-Kapturczak et al. 2015).

BMI

In Chapter 3 I discussed Barnett et al.'s (2014), finding that TAC was lower in participants with a higher BMI. However, when adjusting for gender and number of drinks, the effect of BMI disappeared, and only the number of drinks had a significant effect on TAC (Barnett et al. 2014). As reported in Chapter 4, the correlation measures between TAC measured on the left and right ankle were significantly lower in participants with high than low BMIs, supporting the contention that BMI affects TAC measurement. However, the small sample in this study and the small amount of research available on this topic suggest more research is needed to confirm the effect of BMI on TAC detection.

Food consumption

Chapter 5 describes my examination of the effects of food consumption after drinking alcohol; I found eating had no effect on the level of agreement between TAC data recorded by the left and right SCRAM-CAMs. This result contrasts with those of Saldich et al. (2021), who found no clear patterns in the ratios of the peak and AUC across different food consumption conditions, but reductions in peak TAC and AUC when a meal was consumed before instead of after drinking. The

lack of effect of food consumption in my study may thus be related to the timing of the meal; again, further work is required on this topic.

Drinking rate

In Chapter 5, I reported no observed effects of drinking rate on the relationship between the left and right SCRAM-CAM TAC readings. This finding was in line with that of Hill-Kapturczak et al. (2014), who asked participants to drink up to five beers (each contained 4.6% alcohol by volume, total of 12oz per beer) at their own pace and found no difference in the relationship between TAC and BrAC. No existing research shows whether drinking rate affects TAC readings independently, so this question will be left for future studies to answer.

Limitations and strengths of TAC monitors

The research presented in this thesis shows that substantial improvements in TAC technology are required before we can rely on it to measure alcohol consumption in real time. Most research has focused on the SCRAM-CAM, and although they are still the most reliable monitors available, they are far from perfect in reliability and agreement and in their delay in alcohol detection and time to reach peak alcohol levels (Karns-Wright et al. 2017). SCRAM-CAMs are also costly and uncomfortable (Caluzzi et al. 2019). The WrisTAS monitors have high failure rates (Bond et al. 2014; Simons et al. 2015) and were recently (unknown when exactly) taken off the market completely. Newer-generation wristbands, including the ION RAP and Skyn, seem promising, but research shows high failure rates for both monitors, and in my research the ION RAP TAC readings had close to zero correlation to BrAC and substantial discrepancies when measured twice in the same conditions. However, my research involved ION RAP prototypes rather than fully developed products, so it is premature to make any recommendations about their use or otherwise here. The lack of research on their reliability and validity in real-world drinking situations makes it difficult to know their ability to accurately estimate BAC in these situations. Further, because these wrist monitors are still in development and expensive, it would be difficult and costly to use them with large sample sizes. With commercial release and wider use of these monitors, prices are likely to decrease and larger studies

will become more feasible. Finally, new monitors have been developed and evaluated since I completed my systematic review (Chapter 3), including the Metal OXide (MOX) sensor (Lawson et al. 2019), AWARE (Lin et al. 2019) and the IoT sensor (Li et al. 2021). As for the ION RAP at the time of my systematic review, published research on the reliability and validity of these new devices is not yet available.

Even though current TAC monitors have substantial limitations, they do enable accurate detection of alcohol use, and can thus be used as abstinence monitors in populations in which self-reported drinking would be unreliable. Nonetheless, objective and precise estimates of BAC in near-real time remain crucial for research in alcohol consumption. The limitations of self-reports (including EMA) and breathalysers for research purposes have been covered thoroughly in this thesis (see Chapter 1). TAC monitors can measure alcohol consumption without action required from the participant, even at levels of intoxication that would make active research participation challenging. So, for situations like festivals, sporting events or other heavy drinking events or populations, TAC monitors have advantages over other commonly employed methods of alcohol consumption measurement. Finally, new wrist-worn TAC monitors are being developed, including some with smartphone integration, allowing individuals to examine their real-time intoxication levels at will. Wider use and further development of small, discreet wrist-worn monitors, including the ION RAP, Skyn, MOX, AWARE, and IoT sensor, will reduce costs and improve performance, enabling them to produce precise estimates of BAC in near-real time and be applied to larger populations.

Limitations and strengths of TAC research

The research literature on TAC monitors is still far from comprehensive. Studies of reliability and validity, including those that form part of this thesis, have a mean sample size of about 25 participants, with the largest sample consisting of 61 participants (Karns-Wright et al. 2017; Roache et al. 2015). One study involved one of the authors as the only participant (Luczak and Rosen 2014). This means that most studies of TAC monitors and the relationship between TAC and BrAC has been well powered, and the high failure rates of wristbands have further reduced sample sizes and study power. For this reason, research can only determine effect sizes, not the statistical significance of

those effects. However, it is important to note that even though my laboratory alcohol-administration study only included 22 participants, all of them came into the laboratory for three full days of data collection. During each of these days, the SCRAM-CAMs measured TAC every 30 minutes, and the ION RAP every ~five seconds, producing a substantial amount of data. Multiple devices per participant further increased the amount of data. So, although statistical power may be low in the studies focusing on inter-individual relationships or the effects of the differences in the sessions (e.g., drinking rate), for the TAC sampling level I had sufficient power to test overall reliability and agreement. Another strength of my research was that it involved both research in a controlled laboratory setting and in naturalistic settings, covering different aspects of the reliability and validity of TAC monitoring.

A lack of standardised data cleaning procedures for TAC results makes comparability across validation studies of TAC monitors difficult. Further, most studies of TAC monitors have relied on relatively conventional analysis (regression, etc.). Before starting the main laboratory study in this thesis, substantial time should have been put into trialling the monitors, ensuring consistency in strap tightness, cleaning of the face plates and refining the calibration time to maximise TAC data quality. Even though data from 10 sessions in the laboratory alcohol-administration study had to be discarded due to insufficient time being allowed for calibration of the SCRAM-CAMs, the devices were calibrated and adjusted to the participants in the other 55 sessions, enabling the monitor to adjust to the participant. Moreover, I employed data cleaning procedures recommended in previous research (Roache et al., 2019). These procedures were demonstrated to improve the sensitivity of the SCRAM-CAM and to quantify lower levels of alcohol consumption.

Research on the TAC to BrAC relationship to date has been mostly based on the SCRAM-CAM, which, as noted, has a range of limitations. Additionally, regression models can estimate BrAC, but if they include parameters such as time-to-peak, which can only be calculated via post-hoc measurement of TAC over the drinking event, they are not helpful for the conversion of TAC in real time. For real-time alcohol measurement and TAC interpretation, researchers are developing machine learning algorithms such as that used in the BrAC Estimator Software (Fairbairn et al. 2019;

Fairbairn, Kang, & Bosch 2020; Sirlanci et al. 2018). This software has shown promise as a tool for the conversion of TAC to BrAC, but is not available to the public yet. Finally, it is unknown how this software performs using TAC data as measured by the new-generation wristbands, which means that the exact relationship between TAC and BrAC remains unclear.

Recommendations for further research

More comprehensive validation of TAC monitoring technology is necessary, including studies using much larger samples across multiple contexts enabling determination of the effects of both individual and contextual factors. These larger datasets will also be necessary to improve the ability of machine learning applications, such as the BrAC Estimator Software, to convert TAC values into interpretable and accurate BrAC estimates. In particular, more research in naturalistic drinking events and with current and new devices is critical. Future research should put substantial time into trialling their monitors, ensuring consistency in strap tightness, cleaning of the face plates and refining the calibration time in order to generate high-quality TAC data. Finally, researchers in the TAC technology space should develop and agree on standardised data cleaning and processing protocols to improve the comparability of future studies.

Conclusions

The results of the research presented herein suggest that wearable TAC monitors still have substantial room for improvement in reliability and validity, and are not sufficiently reliable and valid to measure quantities of alcohol consumed in real time. SCRAM-CAMs are the most reliable and valid monitors available today, but have their own limitations. Newer, wrist-worn devices, including the next-generation ION RAP and the Skyn, may become similarly effective with further development. Despite their flaws, TAC monitors do enable detection of whether alcohol has been consumed or not and may be especially useful for research involving heavy drinkers and in situations where self-reported drinking would be especially unreliable and breathalysers too burdensome to capture a full drinking event. In this way, TAC monitors can complement self-reports or

breathalysers, and – dependent on study design, participants and research aims – a combination of these methods can meaningfully inform alcohol consumption prevention and intervention strategies.

References

Alcohol Monitoring Services. 2022. “Alcohol Monitoring Services.” (https://www.scramsystems.com/our-company/about-us/).

Allan, Julaine, Anton Clifford, Patrick Ball, Margaret Alston, and Peter Meister. 2012. “‘You’re Less Complete If You Haven’t Got a Can in Your Hand’: Alcohol Consumption and Related Harmful Effects in Rural Australia: The Role and Influence of Cultural Capital.” *Alcohol and Alcoholism* 47(5):624–29. doi: 10.1093/ALCALC/AGS074.

BACtrack Skyn. 2022. “BACtrack Skyn™.” (https://skyn.bactrack.com/).

Australian Bureau of Statistics. 2010. “Spectator Attendance at Sporting Events, Australia (Cat. No. 4174.0).” *Canberra: ABS*.

Australian Institute of Health and Welfare. 2020. “National Drug Strategy Household Survey 2019.” *Canberra: AIHW* 90.

Australian Institute of Health and Welfare. 2018. *Australian Burden of Disease Study: Impact and Causes of Illness and Death in Australia 2018.*

Barnett, Nancy P., E. B. Meade, and T. R. Glynn. 2014. “Predictors of Detection of Alcohol Use Episodes Using a Transdermal Alcohol Sensor.” *Experimental and Clinical Psychopharmacology* 22(1):86–96. doi: 10.1037/a0034821.

Barnett, Nancy P., Jennifer Tidey, James G. Murphy, Robert Swift, and Suzanne M. Colby. 2011. “Contingency Management for Alcohol Use Reduction: A Pilot Study Using a Transdermal Alcohol Sensor.” *Drug and Alcohol Dependence* 118(2–3):391–99.

Berchtold, André. 2016. “Test–Retest: Agreement or Reliability?:” *Http://Dx.Doi.Org/10.1177/2059799116672875* 9. doi: 10.1177/2059799116672875.

Blair, A. L., A. L. Chiaf, E. K. Crockett, T. K. Teague, and J. M. Croff. 2021. “Validation of Hair Ethyl Glucuronide Using Transdermal Monitoring and Self-Reported Alcohol Use in Women of Childbearing Potential.” *Neuropsychopharmacology Reports* 41(2):144–51. doi: https://dx.doi.org/10.1002/npr2.12151.

Bland, Martin J., and Douglas G. Altman. 1986. “Statistical Methods for Assessing Agreement between Two Methods of Clinical Measurement.” *The Lancet* 327(8476):307–10. doi: 10.1016/S0140-6736(86)90837-8.

Bond, Jason C., Thomas K. Greenfield, Deidre Patterson, and William C. Kerr. 2014. “Adjustments for Drink Size and Ethanol Content: New Results from a Self-Report Diary and Transdermal Sensor Validation Study.” *Alcoholism: Clinical and Experimental Research* 38(12):3060–67. doi: http://dx.doi.org/10.1111/acer.12589.

Bühler, Mira, Sabine Vollstädt-Klein, Andrea Kobiella, Henning Budde, Laurence J. Reed, Dieter F. Braus, Christian Büchel, and Michael N. Smolka. 2010. “Nicotine Dependence Is Characterized by Disordered Reward Processing in a Network Driving Motivation.” *Biological Psychiatry* 67(8):745–52. doi: 10.1016/j.biopsych.2009.10.029.

Callinan, Sarah. 2015. “How Big Is a Self-Poured Glass of Wine for Australian Drinkers?” *Drug and Alcohol Review* 34(2):207–10. doi: 10.1111/DAR.12154.

Caluzzi, Gabriel, Amy Pennay, Megan Cook, Cassandra Wright, Thomas Norman, and Emmanuel Kuntsche. 2019. “Transdermal Monitors to Assess Alcohol Consumption in Real-Time and Real-Life–a Qualitative Study on User-Experience.” *Addiction Research & Theory* 27(4):354–61.

Cicchetti, Domenic V. 1994. “Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology.” *Psychological Assessment* 6(4):284.

Cohen, B. Beth, and Daniel C. Vinson. 1995. “Retrospective Self-Report of Alcohol Consumption: Test-Retest Reliability by Telephone.” *Alcoholism: Clinical and Experimental Research*

19(5):1156–61.

Cohen, Jacob. 2013. *Statistical Power Analysis for the Behavioral Sciences*. Academic press.

Collins, Allan C., Toni N. Yeager, M. E. Lebsack, and S. Scott Panter. 1975. “Variations in Alcohol Metabolism: Influence of Sex and Age.” *Pharmacology Biochemistry and Behavior* 3(6):973–78.

Collins, R. Lorraine, Elizabeth T. Morsheimer, Saul Shiffman, Jean A. Paty, Maryann Gnys, and George D. Papandonatos. 1998. “Ecological Momentary Assessment in a Behavioral Drinking Moderation Training Program.” *Experimental and Clinical Psychopharmacology* 6(3):306–15. doi: 10.1037/1064-1297.6.3.306.

Croff, J. M., M. L. Hartwell, A. L. Chiaf, E. K. Crockett, and I. J. Washburn. 2021. “Feasibility and Reliability of Continuously Monitoring Alcohol Use among Female Adolescents and Young Adults.” *Drug & Alcohol Review* 40(7):1143–54. doi: https://dx.doi.org/10.1111/dar.13045.

Csikszentmihalyi, Mihaly, Reed Larson, M Csikszentmihalyi, and R Larson. 2014. “Validity and Reliability of the Experience-Sampling Method.” *Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi* 35–54. doi: 10.1007/978-94-017-9088-8_3.

Davis, Christopher G., Jennifer Thake, and Natalie Vilhena. 2010. “Social Desirability Biases in Self-Reported Alcohol Consumption and Harms.” *Addictive Behaviors* 35(4):302–11.

Dawson, Deborah A. 1998. “Measuring Alcohol Consumption: Limitations and Prospects for Improvement.” *Addiction* 93(7):965.

Degenhardt, Louisa, Wayne Hall, and Michael Lynskey. 2001. “Alcohol, Cannabis and Tobacco Use among Australians: A Comparison of Their Associations with Other Drug Use and Use Disorders, Affective and Anxiety Disorders, and Psychosis.” *Addiction (Abingdon, England)* 96(11):1603–14. doi: 10.1046/J.1360-0443.2001.961116037.X.

Dodd, Lori E., and Margaret S. Pepe. 2003. “Partial AUC Estimation and Regression.” *Biometrics*

59(3):614–23.

Dougherty, Donald M., Nathalie Hill-Kapturczak, Yuanyuan Liang, Tara E. Karns, Sarah L. Lake, Sharon E. Cates, and John D. Roache. 2015. “The Potential Clinical Utility of Transdermal Alcohol Monitoring Data to Estimate the Number of Alcoholic Drinks Consumed.” *Addictive Disorders & Their Treatment* 14(3):124–30. doi: 10.1097/ADT.0000000000000060.

Dulin, Patrick L., Corene E. Alvarado, James M. Fitterling, and Vivian M. Gonzalez. 2017. “Comparisons of Alcohol Consumption by Timeline Follow Back vs. Smartphone-Based Daily Interviews.” *Addiction Research & Theory* 25(3):195–200.

Durbeej, Natalie, Tobias H. Elgán, Camilla Jalling, and Johanna Gripenberg. 2017. “Alcohol Intoxication at Swedish Football Matches: A Study Using Biological Sampling to Assess Blood Alcohol Concentration Levels among Spectators.” *PLoS One* 12(11):e0188284.

van Egmond, Kelly, Cassandra J. C. Wright, Michael Livingston, and Emmanuel Kuntsche. 2020. “Wearable Transdermal Alcohol Monitors: A Systematic Review of Detection Validity, and Relationship Between Transdermal and Breath Alcohol Concentration and Influencing Factors.” *Alcoholism: Clinical and Experimental Research* 44(10):1918–32.

Estrada, Felipe, and Kalle Tryggvesson. 2001. “A Part of the Game–Alcohol, Football Fans and Male Comradeship.” *Nordic Studies on Alcohol and Drugs* 18(3):245–60.

Fairbairn, C. E., and D. Kang. 2019. “Temporal Dynamics of Transdermal Alcohol Concentration Measured via New-Generation Wrist-Worn Biosensor.” *Alcoholism: Clinical and Experimental Research* 43(10):2060–69. doi: 10.1111/acer.14172.

Fairbairn, C. E., D. Kang, and N. Bosch. 2020. “Using Machine Learning for Real-Time BAC Estimation from a New-Generation Transdermal Biosensor in the Laboratory.” *Drug & Alcohol Dependence* 216:108205. doi: https://dx.doi.org/10.1016/j.drugalcdep.2020.108205.

Fairbairn, C. E., I. G. Rosen, S. E. Luczak, and W. J. Venerable. 2019. “Estimating the Quantity and Time Course of Alcohol Consumption from Transdermal Alcohol Sensor Data: A Combined

Laboratory-Ambulatory Study." *Alcohol* 01:1. doi: https://dx.doi.org/10.1016/j.alcohol.2018.08.015.

Flango, Victor E., and Fred L. Cheesman. 2009. "Effectiveness of the SCRAM Alcohol Monitoring Device: A Preliminary Test." *Drug Court Review* 6(2):109–34.

Gibb, Kenneth A., Albert S. Yee, Charles C. Johnston, Steven D. Martin, and Richard M. Nowak. 1984. "Accuracy and Usefulness of a Breath Alcohol Analyzer." *Annals of Emergency Medicine* 13(7):516–20. doi: 10.1016/S0196-0644(84)80517-X.

Gilligan, Conor, Kristen G. Anderson, Benjamin O. Ladd, Yun Ming Yong, and Michael David. 2019. "Inaccuracies in Survey Reporting of Alcohol Consumption." *BMC Public Health* 19(1):1–11. doi: 10.1186/S12889-019-7987-3/TABLES/4.

Glassman, Tavis, Chudley E. Werch, Edessa Jobli, and Hui Bian. 2007. "Alcohol-Related Fan Behavior on College Football Game Day." *Journal of American College Health* 56(3):255–60.

Gmel, Gerhard, Emmanuel Kuntsche, and Jürgen Rehm. 2011. "Risky Single-Occasion Drinking: Bingeing Is Not Bingeing." *Addiction* 106(6):1037–45. doi: 10.1111/J.1360-0443.2010.03167.X.

Gmel, Gerhard, and Jürgen Rehm. 2004. "Measuring Alcohol Consumption:" *Contemporary Drug Problems* 31(3):467–540. doi: 10.1177/009145090403100304.

Greenfield, Thomas K., and William C. Kerr. 2008. "Alcohol Measurement Methodology in Epidemiology: Recent Advances and Opportunities." *Addiction (Abingdon, England)* 103(7):1082. doi: 10.1111/J.1360-0443.2008.02197.X.

Gunn, R. L., J. A. Steingrimsson, J. E. Merrill, T. Souza, and N. Barnett. 2021. "Characterising Patterns of Alcohol Use among Heavy Drinkers: A Cluster Analysis Utilising Alcohol Biosensor Data." *Drug & Alcohol Review* 40(7):1155–64. doi: https://dx.doi.org/10.1111/dar.13306.

Hammersley, Martyn. 1987. "Some Notes on the Terms 'Validity' and 'Reliability.'" *British Educational Research Journal* 13(1):73–82. doi: 10.1080/0141192870130107.

Heale, Roberta, and Alison Twycross. 2015. “Validity and Reliability in Quantitative Studies.” *Evidence-Based Nursing* 18(3):66–67. doi: 10.1136/EB-2015-102129.

Hill-Kapturczak, Nathalie, Sarah L. Lake, John D. Roache, Sharon E. Cates, Yuanyuan Liang, and Donald M. Dougherty. 2014. “Do Variable Rates of Alcohol Drinking Alter the Ability to Use Transdermal Alcohol Monitors to Estimate Peak Breath Alcohol and Total Number of Drinks?” *Alcoholism: Clinical and Experimental Research* 38(10):2517–22. doi: http://dx.doi.org/10.1111/acer.12528.

Hill-Kapturczak, Nathalie, John D. Roache, Yuanyuan Liang, Tara E. Karns, Sharon E. Cates, and Donald M. Dougherty. 2015. “Accounting for Sex-Related Differences in the Estimation of Breath Alcohol Concentrations Using Transdermal Alcohol Monitoring.” *Psychopharmacology* 232(1):115–23. doi: 10.1007/s00213-014-3644-9.

Hufford, Michael R., Alan L. Shields, Saul Shiffman, Jean Paty, and Mark Balabanis. 2002. “Reactivity to Ecological Momentary Assessment: An Example Using Undergraduate Problem Drinkers.” *Psychology of Addictive Behaviors* 16(3):205–11. doi: 10.1037/0893-164X.16.3.205.

Hurzeler, Tristan, Nicholas A. Buckley, Firouzeh Noghrehchi, Peter Malouf, Andrew Page, Jennifer L. Schumann, and Kate M. Chitty. 2021. “Alcohol-Related Suicide across Australia: A Geospatial Analysis.” *Australian and New Zealand Journal of Public Health* 45(4):394–99. doi: 10.1111/1753-6405.13122.

Institute, National Drug Research. 2021. “Social and Economic Costs of Substance Use.”

Johnson, Richard D., Michael Morowitz, Anne F. Maddox, David J. Shearman, and North Terrace. 1991. “Cigarette Smoking and Rate of Gastric Emptying: Effect on Alcohol Absorption.” *BMJ : British Medical Journal* 302(6767):20. doi: 10.1136/BMJ.302.6767.20.

Jones, Andrew, Danielle Remmerswaal, Ilse Verveer, Eric Robinson, Ingmar H. A. Franken, Cheng K. Fred Wen, and Matt Field. 2018. “Compliance with Ecological Momentary Assessment Protocols in Substance Users: A Meta-Analysis.” doi: 10.1111/add.14503.

Kalist, David E., and Daniel Y. Lee. 2016. “The National Football League: Does Crime Increase on Game Day?” *Journal of Sports Economics* 17(8):863–82.

Karns-Wright, Tara E., Donald M. Dougherty, Nathalie Hill-Kapturczak, Charles W. Mathias, and John D. Roache. 2018. “The Correspondence between Transdermal Alcohol Monitoring and Daily Self-Reported Alcohol Consumption.” *Addictive Behaviors* 85:147–52. doi: 10.1016/j.addbeh.2018.06.006.

Karns-Wright, Tara E., John D. Roache, Nathalie Hill-Kapturczak, Yuanyuan Liang, Jillian Mullen, and Donald M. Dougherty. 2017. “Time Delays in Transdermal Alcohol Concentrations Relative to Breath Alcohol Concentrations.” *Alcohol and Alcoholism* 52(1):35–41. doi: http://dx.doi.org/10.1093/alcalc/agw058.

Kelly, Thomas M., John E. Donovan, Janet M. Kinnane, and David M. C. D. Taylor. 2002. “A Comparison of Alcohol Screening Instruments Among Under-Aged Drinkers Treated in Emergency Departments.” *Alcohol and Alcoholism* 37(5):444–50. doi: 10.1093/ALCALC/37.5.444.

Kerr, William C., Thomas K. Greenfield, Jennifer Tujague, and Stephan E. Brown. 2005. “A Drink Is a Drink? Variation in the Amount of Alcohol Contained in Beer, Wine and Spirits Drinks in a US Methodological Sample.” *Alcoholism: Clinical and Experimental Research* 29(11):2015–21. doi: 10.1097/01.ALC.0000187596.92804.BD.

Kerr, William C., Deidre Patterson, Mary Albert Koenen, and Thomas K. Greenfield. 2008. “Alcohol Content Variation of Bar and Restaurant Drinks in Northern California.” *Alcoholism, Clinical and Experimental Research* 32(9):1623–29. doi: 10.1111/J.1530-0277.2008.00741.X.

Khadjesari, Zarnie, Elizabeth Murray, Eleftheria Kalaitzaki, Ian R. White, Jim McCambridge, Christine Godfrey, and Paul Wallace. 2009. “Test–Retest Reliability of an Online Measure of Past Week Alcohol Consumption (the TOT-AL), and Comparison with Face-to-Face Interview.” *Addictive Behaviors* 34(4):337–42.

Kingsland, Melanie, Luke Wolfenden, Bosco C. Rowland, Karen E. Gillham, Vanessa J. Kennedy, Robyn L. Ramsden, Richard W. Colbran, Sarah Weir, and John H. Wiggers. 2013. “Alcohol Consumption and Sport: A Cross-Sectional Study of Alcohol Management Practices Associated with at-Risk Alcohol Consumption at Community Football Clubs.” *BMC Public Health* 13(1):1–9.

Kuntsche, E., and F. Labhart. 2014. “The Future Is Now--Using Personal Cellphones to Gather Data on Substance Use and Related Factors.” *Addiction* 109(7):1052–53. doi: 10.1111/ADD.12540.

Kuntsche, Emmanuel, and Raimondo Bruno. 2015. “Moody Booze: Introducing the Special Section on Affect Regulation and Substance Use.” *Drug and Alcohol Review* 34(6):569–70. doi: 10.1111/DAR.12349.

Kuntsche, Emmanuel, Paul Dietze, and Rebecca Jenkinson. 2014. “Understanding Alcohol and Other Drug Use during the Event.” *Drug and Alcohol Review* 33(4):335–37. doi: 10.1111/DAR.12171.

Kuntsche, Emmanuel, and Florian Labhart. 2012. “Investigating the Drinking Patterns of Young People over the Course of the Evening at Weekends.” *Drug and Alcohol Dependence* 124(3):319–24. doi: 10.1016/J.DRUGALCDEP.2012.02.001.

Kuntsche, Emmanuel, and Florian Labhart. 2013a. “Using Personal Cell Phones for Ecological Momentary Assessment: An Overview of Current Developments.” *European Psychologist* 18(1):3–11. doi: 10.1027/1016-9040/A000127.

Kuntsche, Emmanuel, and Florian Labhart. 2013b. “Using Personal Cell Phones for Ecological Momentary Assessment.” *European Psychologist*.

Labhart, Florian, Skanda Muralidhar, Benoit Massé, Lakmal Meegahapola, Emmanuel Kuntsche, and Daniel Gatica-Perez. 2021. “Ten Seconds of My Nights: Exploring Methods to Measure Brightness, Loudness and Attendance and Their Associations with Alcohol Use from Video Clips.” *PLOS ONE* 16(4):e0250443. doi: 10.1371/JOURNAL.PONE.0250443.

Labhart, Florian, Thanh Trung Phan, Daniel Gatica-Perez, and Emmanuel Kuntsche. 2021. “Shooting

Shots: Estimating Alcoholic Drink Sizes in Real Life Using Event-Level Reports and Annotations of Close-up Pictures." *Drug and Alcohol Review* 40(7):1228–38. doi: 10.1111/DAR.13212.

Labhart, Florian, Darshan Santani, Jasmine Truong, Flavio Tarsetti, Olivier Bornet, Sara Landolt, Daniel Gatica-Perez, and Emmanuel Kuntsche. 2017. "Development of the Geographical Proportional-to-Size Street-Intercept Sampling (GPSIS) Method for Recruiting Urban Nightlife-Goers in an Entire City." *Https://Doi.Org/10.1080/13645579.2017.1293928* 20(6):721–36. doi: 10.1080/13645579.2017.1293928.

Labhart, Florian, Flavio Tarsetti, Olivier Bornet, Darshan Santani, Jasmine Truong, Sara Landolt, Daniel Gatica-Perez, and Emmanuel Kuntsche. 2019. "Capturing Drinking and Nightlife Behaviours and Their Social and Physical Context with a Smartphone Application – Investigation of Users' Experience and Reactivity." *Https://Doi.Org/10.1080/16066359.2019.1584292* 28(1):62–75. doi: 10.1080/16066359.2019.1584292.

Lajunen, Timo, and Heikki Summala. 2003. "Can We Trust Self-Reports of Driving? Effects of Impression Management on Driver Behaviour Questionnaire Responses." *Transportation Research Part F: Traffic Psychology and Behaviour* 6(2):97–107.

Lansdorp, Bob, William Ramsay, Rashad Hamid, and Evan Strenk. 2019. "Wearable Enzymatic Alcohol Biosensor." *Sensors* 19(10):2380.

Laslett, Anne-Marie, Paul Catalano, Tanya Chikritzhs, Caroline Dale, Christopher Doran, Jason Ferris, Thameemul Jainullabudeen, Michael Livingston, Sharon Matthews, Janette Mugavin, Robin Room, Morgan Schlotterlein, and Claire Wilkinson. 2010. "The Range and Magnitude of Alcohol's Harm to Others." *Victoria: AER Centre for Alcohol Policy Research, Turning Point Alcohol and Drug Centre, Eastern Health*.

Lauckner, Carolyn, Erica Taylor, Darshti Patel, and Alexis Whitmire. 2019. "The Feasibility of Using Smartphones and Mobile Breathalyzers to Monitor Alcohol Consumption among People Living

with HIV/AIDS." *Addiction Science and Clinical Practice* 14(1):1–11. doi: 10.1186/S13722-019-0174-0/TABLES/3.

Lawson, B., K. Aguir, T. Fiorido, V. Martini-Laithier, R. Bouchakour, S. Burtey, C. Reynard-Carette, and M. Bendahan. 2019. "Skin Alcohol Perspiration Measurements Using MOX Sensors." *Sensors and Actuators, B: Chemical* 280:306–12. doi: 10.1016/j.snb.2018.09.082.

Leffingwell, Thad R., Nathaniel J. Cooney, James G. Murphy, Susan Luczak, Gary Rosen, Donald M. Dougherty, and Nancy P. J. Barnett. 2013. "Continuous Objective Monitoring of Alcohol Use: Twenty-first Century Measurement Using Transdermal Sensors." *Alcoholism: Clinical and Experimental Research* 37(1):16–22.

Li, B., R. S. Downen, Q. Dong, N. Tran, M. LeSaux, A. C. Meltzer, and Z. Li. 2021. "A Discreet Wearable IoT Sensor for Continuous Transdermal Alcohol Monitoring - Challenges and Opportunities." *IEEE Sensors Journal* 21(4):5322–30. doi: https://dx.doi.org/10.1109/jsen.2020.3030254.

Lin, K. C., D. Kinnamon, D. Sankhala, S. Muthukumar, and S. Prasad. 2019. "AWARE: A Wearable Awareness with Real-Time Exposure, for Monitoring Alcohol Consumption Impact through Ethyl Glucuronide Detection." *Alcohol* 81:93–99. doi: https://dx.doi.org/10.1016/j.alcohol.2018.10.006.

Livingston, Michael, and Sarah Callinan. 2015. "Underreporting in Alcohol Surveys: Whose Drinking Is Underestimated?" *Journal of Studies on Alcohol and Drugs* 76(1):158–64.

Lloyd, Belinda, Sharon Matthews, Michael Livingston, Harindra Jayasekara, and Karen Smith. 2013. "Alcohol Intoxication in the Context of Major Public Holidays, Sporting and Social Events: A Time–Series Analysis in M Elbourne, A Ustralia, 2000–2009." *Addiction* 108(4):701–9.

Luczak, S. E., and I. G. Rosen. 2014. "Estimating BrAC from Transdermal Alcohol Concentration Data Using the BrAC Estimator Software Program." *Alcoholism: Clinical and Experimental Research* 38(8):2243–52. doi: 10.1111/acer.12478.

Luczak, S E, I. G. Rosen, and T. L. Wall. 2015. “Development of a Real-Time Repeated-Measures Assessment Protocol to Capture Change over the Course of a Drinking Episode.” *Alcohol and Alcoholism* 50(2):180–87. doi: 10.1093/alcalc/agu100.

Luczak, Susan E., I. Gary Rosen, and Tamara L. Wall. 2015. “Development of a Real-Time Repeated-Measures Assessment Protocol to Capture Change over the Course of a Drinking Episode.” *Alcohol and Alcoholism* 50(2):180–87. doi: 10.1093/ALCALC/AGU100.

Maher, Jessica Middlemis, Jonathan C. Markey, and Diane Ebert-May. 2013. “The Other Half of the Story: Effect Size Analysis in Quantitative Research.” *CBE—Life Sciences Education* 12(3):345–51.

Marques, Paul R., and A. Scott McKnight. 2007. *Evaluating Transdermal Alcohol Measuring Devices*. United States. National Highway Traffic Safety Administration.

Marques, Paul R., and A. Scott McKnight. 2009. “Field and Laboratory Alcohol Detection with 2 Types of Transdermal Devices.” *Alcoholism: Clinical and Experimental Research* 33(4):703–11.

May, Carl, and David Foxcroft. 1995. “Minimizing Bias in Self-Reports of Health Beliefs and Behaviours.” *Health Education Research* 10(1):107–12.

Meier, Patrick, and Helmut K. Seitz. 2008. “Age, Alcohol Metabolism and Liver Disease.” *Current Opinion in Clinical Nutrition & Metabolic Care* 11(1):21–26.

Miller, Elizabeth T., Dan J. Neal, Lisa J. Roberts, John S. Boer, Sally O. Cresslr, Jane Metrik, and G. Alan Marlatt. 2002. “Test-Retest Reliability of Alcohol Measures: Is There a Difference between Internet-Based Assessment and Traditional Methods?” *Psychology of Addictive Behaviors* 16(1):56–63.

Miller, Peter, Louise M c Donald, Stephen M c Kenzie, Kerry O’Brien, and Petra Staiger. 2013. “When the C Ats Are Away: The Impact of Sporting Events on Assault-and Alcohol-related Emergency Department Attendances.” *Drug and Alcohol Review* 32(1):31–38.

Monk, Rebecca Louise, Derek Heim, Adam Qureshi, and Alan Price. 2015. "'I Have No Clue What I Drunk Last Night' Using Smartphone Technology to Compare In-Vivo and Retrospective Self-Reports of Alcohol Consumption." *PLOS ONE* 10(5):e0126209. doi: 10.1371/JOURNAL.PONE.0126209.

Mun, Eun-Young, Xiaoyin Li, Michael S. Businelle, Emily T. Hébert, Zhengqi Tan, Nancy P. Barnett, and Scott T. Walters. 2021. "Ecological Momentary Assessment of Alcohol Consumption and Its Concordance with Transdermal Alcohol Detection and Timeline Follow-Back Self-report Among Adults Experiencing Homelessness." *Alcoholism: Clinical and Experimental Research* 45(4):864–76.

Nelson, Toben F., and Henry Wechsler. 2003. "School Spirits: Alcohol and Collegiate Sports Fans." *Addictive Behaviors* 28(1):1–11.

NHMRC. 2020. "Australian Guidelines to Reduce Health Risks from Drinking Alcohol | NHMRC." Retrieved September 10, 2021 (https://www.nhmrc.gov.au/about-us/publications/australian-guidelines-reduce-health-risks-drinking-alcohol).

Niall Bolger, and Jean-Philippe Laurenceau. 2013. "Intensive Longitudinal Methods: An Introduction to Diary and Experience." *Guilford, New York*. Retrieved January 15, 2022 (https://books.google.com.au/books?hl=en&lr=&id=5bD4LuAFq0oC&oi=fnd&pg=PP1&ots=jnjkZ0GU92&sig=yBv78VkSIjd4yYvE_AHdT8siCJA&redir_esc=y#v=onepage&q&f=false).

Norman, Thomas, Amy Peacock, Stuart G. Ferguson, Emmanuel Kuntsche, and Raimondo Bruno. 2020. "Combining Transdermal and Breath Alcohol Assessments, Real-Time Drink Logs and Retrospective Self-Reports to Measure Alcohol Consumption and Intoxication across a Multi-Day Music Festival." *Drug and Alcohol Review*. doi: 10.1111/DAR.13215.

Palmer, Catherine. 2011. "Key Themes and Research Agendas in the Sport-Alcohol Nexus." *Journal of Sport and Social Issues* 35(2):168–85.

Palmer, Catherine, and Kirrilly Thompson. 2007. "The Paradoxes of Football Spectatorship: On-Field

and Online Expressions of Social Capital among the 'Grog Squad.'" *Sociology of Sport Journal* 24(2):187–205.

Peacock, Amy, Raimondo Bruno, Frances H. Martin, and Andrea Carr. 2013. "The Impact of Alcohol and Energy Drink Consumption on Intoxication and Risk-Taking Behavior." *Alcoholism: Clinical and Experimental Research* 37(7):1234–42. doi: 10.1111/ACER.12086.

Piasecki, Thomas M. 2019. "Assessment of Alcohol Use in the Natural Environment." *Alcoholism: Clinical and Experimental Research* 43(4):564–77.

Riordan, Benjamin C., Damian Scarf, Saleh Moradi, Jayde A. M. Flett, Kate B. Carey, and Tamlin S. Conner. 2017. "The Accuracy and Promise of Personal Breathalysers for Research: Steps toward a Cost-Effective Reliable Measure of Alcohol Intoxication?" *Digital Health* 3:2055207617746752.

Riuttanen, Antti, Saara J. Jäntti, and Ville M. Mattila. 2020. "Alcohol Use in Severely Injured Trauma Patients." *Scientific Reports 2020 10:1* 10(1):1–8. doi: 10.1038/s41598-020-74753-y.

Roache, J. D., T. E. Karns-Wright, M. Goros, N. Hill-Kapturczak, C. W. Mathias, and D. M. Dougherty. 2019. "Processing Transdermal Alcohol Concentration (TAC) Data to Detect Low-Level Drinking." *Alcohol* 01:1. doi: https://dx.doi.org/10.1016/j.alcohol.2018.08.014.

Roache, John D., Tara E. Karns, Nathalie Hill-Kapturczak, Jillian Mullen, Yuanyuan Liang, Richard J. Lamb, and Donald M. Dougherty. 2015. "Using Transdermal Alcohol Monitoring to Detect Low-Level Drinking." *Alcoholism: Clinical and Experimental Research* 39(7):1120–27. doi: http://dx.doi.org/10.1111/acer.12750.

Sakai, Joseph T., Susan K. Mikulich-Gilbertson, Robert J. Long, and Thomas J. Crowley. 2006. "Validity of Transdermal Alcohol Monitoring: Fixed and Self-regulated Dosing." *Alcoholism: Clinical and Experimental Research* 30(1):26–33.

Saldich, Emily B., Chunming Wang, I. Gary Rosen, Jay Bartroff, and Susan E. Luczak. 2021. "Effects of Stomach Content on the Breath Alcohol Concentration-Transdermal Alcohol Concentration

Relationship." *Drug and Alcohol Review*. doi: 10.1111/DAR.13267.

Schafer, Theodore, and A. Walls Joseph L. 2006. *Models for Intensive Longitudinal Data* . Oxford University Press.

Shiffman, Saul. 2009. "Ecological Momentary Assessment (EMA) in Studies of Substance Use." *Psychological Assessment* 21(4):486.

Shiffman, Saul, Arthur A. Stone, and Michael R. Hufford. 2008. "Ecological Momentary Assessment." *Annu. Rev. Clin. Psychol.* 4:1–32.

Simons, Jeffrey S., Thomas A. Wills, Noah N. Emery, and Russell M. Marks. 2015. "Quantifying Alcohol Consumption: Self-Report, Transdermal Assessment, and Prediction of Dependence Symptoms." *Addictive Behaviors* 50:205–12. doi: http://dx.doi.org/10.1016/j.addbeh.2015.06.042.

Sirlanci, M., I. G. Rosen, T. L. Wall, and S. E. Luczak. 2018. "Applying a Novel Population-Based Model Approach to Estimating Breath Alcohol Concentration (BrAC) from Transdermal Alcohol Concentration (TAC) Biosensor Data." *Alcohol* 20:20. doi: https://dx.doi.org/10.1016/j.alcohol.2018.09.005.

Sobell, Linda C., Mark B. Sobell, Gloria I. Leo, and Anthony Cancilla. 1988. "Reliability of a Timeline Method: Assessing Normal Drinkers' Reports of Recent Drinking and a Comparative Evaluation across Several Populations." *British Journal of Addiction* 83(4):393–402.

Sobell, Mark B., Linda C. Sobell, Felix Klajner, Daniel Pavan, and Ellen Basian. 1986. "The Reliability of a Timeline Method for Assessing Normal Drinker College Students' Recent Drinking History: Utility for Alcohol Research." *Addictive Behaviors* 11(2):149–61.

Stockwell, Tim, Susan Donath, Mark Cooper-Stanbury, Tanya Chikritzhs, Paul Catalano, and Cid Mateo. 2004. "Under-Reporting of Alcohol Consumption in Household Surveys: A Comparison of Quantity-Frequency, Graduated-Frequency and Recent Recall." *Addiction (Abingdon, England)* 99(8):1024–33. doi: 10.1111/J.1360-0443.2004.00815.X.

Swette, L., A. Griffith, and A. La Conti. 1997. “Potential and Diffusion Controlled Solid Electrolyte Sensor for Continuous Measurement of Very Low Levels of Transdermal Alcohol.” *Giner, Inc., Waltham, MA. Appl. No. 840802*.

Swift, Robert. 2003. “Direct Measurement of Alcohol and Its Metabolites.” *Addiction* 98(SUPPL. 2):73–80. doi: 10.1046/J.1359-6357.2003.00605.X.

Swift, Robert M., Christopher S. Martin, Larry Swette, Anthony LaConti, and Nancy Kackley. 1992. “Studies on a Wearable, Electronic, Transdermal Alcohol Sensor.” *Alcoholism: Clinical and Experimental Research* 16(4):721–25. doi: 10.1111/J.1530-0277.1992.TB00668.X.

Trull, Timothy J., and Ulrich Ebner-Priemer. 2014. “The Role of Ambulatory Assessment in Psychological Science.” *Current Directions in Psychological Science* 23(6):466–70. doi: 10.1177/0963721414550706.

Vestal, Robert E., Elizabeth A. McGuire, Jordan D. Tobin, Reubin Andres, Arthur H. Norris, and Esteban Mezey. 1977. “Aging and Ethanol Metabolism.” *Clinical Pharmacology & Therapeutics* 21(3):343–54. doi: 10.1002/CPT1977213343.

Voas, Robert B., Robert L. DuPont, Stephen K. Talpins, and Corinne L. Shea. 2011. “Towards a National Model for Managing Impaired Driving Offenders.” *Addiction* 106(7):1221–27. doi: 10.1111/J.1360-0443.2010.03339.X.

Voogt, Carmen, Emmanuel Kuntsche, Marloes Kleinjan, Evelien Poelen, and Rutger Engels. 2014. “Using Ecological Momentary Assessment to Test the Effectiveness of a Web-Based Brief Alcohol Intervention Over Time Among Heavy-Drinking Students: Randomized Controlled Trial.” *J Med Internet Res 2014;16(1):E5 Https://Www.Jmir.Org/2014/1/E5* 16(1):e2817. doi: 10.2196/JMIR.2817.

Wolfe, Jeannette, Ricardo Martinez, and Warren A. Scott. 1998. “Baseball and Beer: An Analysis of Alcohol Consumption Patterns among Male Spectators at Major-League Sporting Events.” *Annals of Emergency Medicine* 31(5):629–32.

Wray, Tyler B., Jennifer E. Merrill, and Peter M. Monti. 2014. “Using Ecological Momentary Assessment (EMA) to Assess Situation-Level Predictors of Alcohol Use and Alcohol-Related Consequences.” *Alcohol Research: Current Reviews* 36(1):19.

Wright, Cassandra, Paul M. Dietze, Paul A. Agius, Emmanuel Kuntsche, Michael Livingston, Oliver C. Black, Robin Room, Margaret Hellard, and Megan S. C. Lim. 2018. “Mobile Phone-Based Ecological Momentary Intervention to Reduce Young Adults’ Alcohol Use in the Event: A Three-Armed Randomized Controlled Trial.” *JMIR MHealth and UHealth* 6(7):e149.

Züll, C. 2016. “Open-Ended Questions.” *GESIS Survey Guidelines* 3.