JMIR Publications
The Leading eHealth Publisher

AlcoWatch – Exploring the feasibility of using smartwatch-based Ecological Momentary Assessment for high temporal density longitudinal measurement of alcohol use.

Chris Stone, Sally Adams, Robyn E Wootton, Andy Skinner

Submitted to: JMIR Formative Research
on: June 12, 2024

Preprint JMIR Publications

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on it's website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressively prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Preprint
JMIR Publications

AlcoWatch – Exploring the feasibility of using smartwatch-based Ecological Momentary Assessment for high temporal density longitudinal measurement of alcohol use.

Chris Stone[1, 2]; Sally Adams[3]; Robyn E Wootton[2, 4, 5, 6]; Andy Skinner[1]

[1]Integrative Cancer Epidemiology Programme Bristol Medical School University of Bristol Bristol GB
[2]School of Psychological Science University of Bristol Bristol GB
[3]School of Psychology University of Birmingham Birmingham GB
[4]Nic Waals Institute Lovisenberg Diaconal Hospital Oslo NO
[5]MRC Integrative Epidemiology Unit Bristol Medical School University of Bristol Bristol GB
[6]PsychGen Centre of Genetic Epidemiology and Mental Health Norwegian Institute of Public Health Oslo NO

Corresponding Author:
Chris Stone
Integrative Cancer Epidemiology Programme
Bristol Medical School
University of Bristol
School of Psychological Science
12a Priory Road
Bristol
GB

Abstract

Background: Smartwatch-based ecological momentary assessment (EMA) methods have previously been developed to minimise participant burden and maximise engagement and compliance. In this study we explore the feasibility of using these smartwatch-based EMA methods to capture longitudinal, high temporal density self-report data about alcohol consumption in a non-clinical population selected to represent high and low Social-Economic Position (SEP) groups.

Objective: To assess the feasibililty of using smartwatch-based ecological momentary assessment methods to capture self-report data about alcohol consumption.

Methods: Thirty-two participants from the Avon Longitudinal Study of Parents and Children (13 high, 19 low SEP) wore a smartwatch running a custom-developed EMA app for 3 months between Oct 2019 and June 2020. Every day over a 12 week period, participants were asked five times a day about any alcoholic drinks they had consumed in the previous two hours, and the context in which they were consumed. They were also asked if they had missed recording any alcoholic drinks the day before. As a comparison, participants also completed fortnightly online diaries of alcohol consumed using the Timeline Followback (TLFB) method. At the end of the study participants completed a semi-structured interview about their experiences.

Results: The compliance rate for all participants who started the study for the smartwatch ?EMA method decreased from around 70% in week 1 to 45% in week 12, compared with the online TLFB method which was flatter at around 50% over the 12 weeks. The compliance for all participants still active for the smartwatch ?EMA method was much flatter, around 70% for the whole 12 weeks, while for the online TLFB method it varied between 50% and 80% over the same period. The completion rate for the smartwatch ?EMA method varied around 80% across the twelve weeks. Within high and low SEP groups there was considerable variation in compliance and completion at each week of the study. However, almost all point estimates for both smartwatch EMA and online TLFB indicated lower levels of engagement for low SEP participants. All participants scored 'experiences of using' the two methods equally highly, with 'willingness to use again' slightly higher for smartwatch EMA.

Conclusions: Our findings demonstrate the feasibility and acceptability of using smartwatch EMA methods for capturing alcohol consumption. These methods have the benefits of capturing higher temporal density, longitudinal data on alcohol consumption, with greater participant engagement and less missing data, and with the potential to be less susceptible to the recall errors than established methods such as TLFB. Future studies should explore the factors impacting participant attrition (the biggest reason for reduced engagement), and the validity of alcohol data captured with these methods. The consistent pattern of lower engagement in low SEP than high SEP participants indicates further work is also warranted to explore the impact and causes of these differences.

(JMIR Preprints 12/06/2024:63184)
DOI: https://doi.org/10.2196/preprints.63184

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.
2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?
✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Preprint JMIR Publications

Original Manuscript

Preprint JMIR Publications

AlcoWatch – Exploring the feasibility of using smartwatch-based Ecological Momentary Assessment for high temporal density longitudinal measurement of alcohol use.

Chris Stone[1,2], Sally Adams[3], Robyn E. Wootton[2,4,5,6], & Andy Skinner[1],

1. Integrative Cancer Epidemiology Programme, Bristol Medical School, University of Bristol
2. School of Psychological Science, University of Bristol
3. School of Psychology, University of Birmingham.
4. Nic Waals Institute, Lovisenberg Diaconal Hospital, Oslo, Norway
5. MRC Integrative Epidemiology Unit, Bristol Medical School, University of Bristol
6. PsychGen Centre for Genetic Epidemiology and Mental Health, Norwegian Institute of Public Health, Oslo, Norway.

Corresponding author: Chris Stone – chris.stone@bristol.ac.uk

Keywords

Smartwatch, Ecological Momentary Assessment, EMA, Alcohol, ALSPAC

Funding

This work was funded by a UKRI Innovation Fellowship to Skinner (MR/S003894/1), and further supported by the Integrative Cancer Epidemiology Programme, funded by Cancer Research UK (grant number C18281/A29019). The MRC Integrative Epidemiology Unit is supported by the Medical Research Council and the University of Bristol (grants MC_UU_00032/05 and MC_UU_00032/07). Wootton is funded by a postdoctoral fellowship from the South-Eastern Norway Regional Health Authority (2020024).
The UK Medical Research Council and Wellcome (Grant ref: 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors and Andy Skinner will serve as guarantor for the contents of this paper.

Conflict of Interests

None

Abstract

Background: Smartwatch-based ecological momentary assessment (EMA) methods have previously been developed to minimise participant burden and maximise engagement and compliance. In this study we explore the feasibility of using these smartwatch-based EMA methods to capture longitudinal, high temporal density self-report data about alcohol consumption in a non-clinical population selected to represent high and low Social-Economic Position (SEP) groups.

Methods: Thirty-two participants from the Avon Longitudinal Study of Parents and Children (13 high, 19 low SEP) wore a smartwatch running a custom-developed EMA app for 3 months between Oct 2019 and June 2020. Every day over a 12 week period, participants were asked five times a day about any alcoholic drinks they had consumed in the previous two hours, and the context in which they were consumed. They were also asked if they had missed recording any alcoholic drinks the day before. As a comparison, participants also completed fortnightly online diaries of alcohol consumed using the Timeline Followback (TLFB) method. At the end of the study participants completed a semi-structured interview about their experiences.

Results: The compliance rate for all participants who started the study for the smartwatch μEMA method decreased from around 70% in week 1 to 45% in week 12, compared with the online TLFB method which was flatter at around 50% over the 12 weeks. The compliance for all participants still active for the smartwatch μEMA method was much flatter, around 70% for the whole 12 weeks, while for the online TLFB method it varied between 50% and 80% over the same period. The completion rate for the smartwatch μEMA method varied around 80% across the twelve weeks. Within high and low SEP groups there was considerable variation in compliance and completion at each week of the study. However, almost all point estimates for both smartwatch EMA and online TLFB indicated lower levels of engagement for low SEP participants. All participants scored 'experiences of using' the two methods equally highly, with 'willingness to use again' slightly higher for smartwatch EMA.

Conclusions: Our findings demonstrate the feasibility and acceptability of using smartwatch EMA methods for capturing alcohol consumption. These methods have the benefits of capturing higher temporal density, longitudinal data on alcohol consumption, with greater participant engagement and less missing data, and with the potential to be less susceptible to the recall errors than established methods such as TLFB. Future studies should explore the factors impacting participant attrition (the biggest reason for reduced engagement), and the validity of alcohol data captured with these methods. The consistent pattern of lower engagement in low SEP than high SEP participants indicates further work is also warranted to explore the impact and causes of these differences.

Introduction

Ecological Momentary Assessment (EMA) refers to a class of methods in which people are asked to answer questions about specific aspects of their health, feelings and behaviours as they go about their normal lives. As individuals typically complete these assessments throughout the day, rather than completing a survey and having to remember details from days or weeks ago, recall errors are significantly reduced[1]. One domain in which EMA methods have been used extensively is the study of modifiable health behaviours, including physical activity, diet, tobacco smoking, and alcohol consumption[2]. The benefits of EMA methods are particularly important for measuring alcohol consumption. Commonly used questionnaires and diary-based methods are known to suffer from systematic biases towards underreporting, capturing somewhere between 20% and 60% of the true level of consumption[3,4], with retrospective recall bias being a major factor[5]. Furthermore, recent studies have highlighted the importance of both longitudinal measures of alcohol consumption in order to help us explore relationships between alcohol consumption and conditions such as depression[6], and high temporal resolution measures of alcohol consumption to enable exploration of factors such as subjective responses (feeling of stimulation, feeling of sedation, liking the effects of alcohol, wanting more alcohol) in heavy drinkers in natural environments[7]. There is, therefore, a need for EMA methods that can be used to capture high temporal density measures of alcohol consumption over extended periods of time.

In response to the increased adoption of mobile and wearable technologies, EMA methods have been adapted for use on smartphones, and more recently, smartwatches. For smartwatch-based EMA methods, a particular area of focus has been reducing participant burden, so that even more frequent sampling over longer periods can be tolerated. Smartwatches are worn on the wrist so they are never beyond reach, which reduces time to access the device and respond to prompts. Also, because they are worn against the body, they enable the use of potentially more discrete haptic prompts for responses. One smartwatch EMA method aiming to minimise participant burden is microinteraction-based EMA (μEMA), which reduces prompts to brief questions with a limited set of answers that can be responded to with a single tap multiple times a day[8]. When compared with smartphone-based EMA over a 4 week period, smartwatch-based μEMA had better performance in terms of compliance rates (defined as the percentage of all scheduled prompts to which participants responded, regardless of the success of prompt delivery: 82% v 64%), and completion rates (defined as the percentage of scheduled prompts actually delivered to the participant to which they responded: 92% v 67%)[8].

Having demonstrated high levels of engagement in these initial 4 week explorations of smartwatch-based μEMA, the approach was subsequently used to capture self-report measures over longer periods in various settings. Beukenhorst and colleagues[9] explored the use of smartwatch-based μEMA methods in participants with osteoarthritis of the knee to capture self-report measures of pain and quality of life 4 or 5 times a day over 3 months. The compliance rate across all participants ranged from around 85% at the beginning of the study to around 30% at the end of the 3 months. The authors suggested the reduction over time was largely a result of participant attrition, possibly driven by technical factors including poor watch battery life. Considering just active participants, compliance remained above 75% throughout the 3 months. Ponnada and colleagues[10] explored using the same smartwatch-based μEMA methods for capturing self-report data of physical activity and affect in young adults over an even longer timeframe of 12 months. Preliminary findings at 6 months (when the authors published interim results of the then-unfinished study) report overall compliance of rate 67%, and completion rate of 79%[10]. These findings suggest that while it is necessary to be mindful of technical issues that can lead to attrition, smartwatch-based μEMA appears to be a viable method for collecting high temporal resolution self-report data over extended timeframes.

In terms of the effect of sample characteristics on EMA study compliance, a number of recent studies have reported mixed findings for age and sex[11-13]. One sample characteristic not explored as extensively is socio-economic position (SEP). For studies of alcohol use, SEP is of particular importance as there are known to be complex relationships between SEP, alcohol consumption and health outcomes (the ‘alcohol harm paradox’), in which individuals are more likely to have drank in the past week and drank 6 units of alcohol or more in one drinking episode if they are a high-income earning managerial/professional worker[14,15], but individuals with lower individual or neighbourhood SEP have an increased susceptibility to the negative and harmful health effects of alcohol consumption[16,17]. Understanding if SEP has a systematic effect on engagement with any new methods measuring alcohol use is therefore clearly important.

The objective of this study was to explore the feasibility of using smartwatch-based μEMA methods for the capture of high resolution self-reported data about alcohol use over extended periods of time. We asked participants to use a smartwatch-based μEMA system to record self-reported data about alcohol use and the context in which alcohol was used six times a day over a 3-month period. For comparison, we also asked all participants to record their alcohol use every two weeks using an online version of one of the currently most established methods for capturing data on alcohol consumption, the timeline followback (TLFB) method. We report engagement metrics (compliance and completion) for both smartwatch μEMA and TLFB methods. Qualitative experiences of participants (gathered in interviews at the end of the study) focused primarily on smartwatch μEMA are reported for all participants, and for high and low SEP groups. Finally, in line with recommendations from a recent evaluation of pressing issues in EMA studies[18], we also report the characteristics of missing data for the smartwatch μEMA for all participants and high and low SEP groups, to explore whether SEP impacts the nature of missing data in a systematic manner.

Methods

Study overview

The design and analyses of the study followed those detailed in the pre-registered study protocol[19]. Any deviations from the study protocol are detailed in the following sections. An overview of the study elements, with their timings throughout the study is shown in Figure 1. Data collection happened between September 2019 and June 2020.

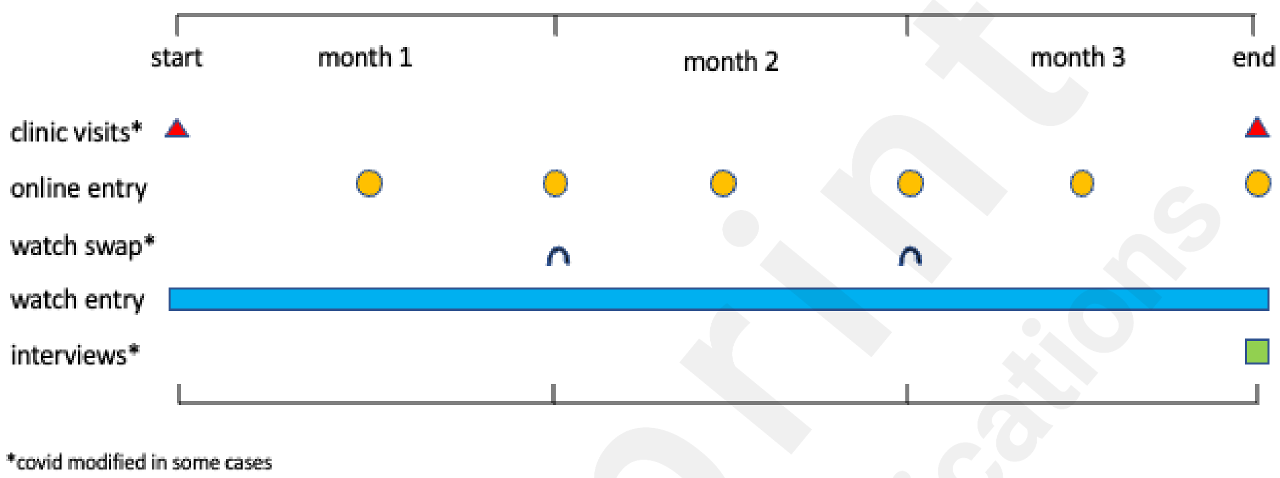

Figure 1. Overview of study elements with timings

Ethical approval

Ethical approval for the study was provided by the Avon Longitudinal Study of Parents and Children (ALSPAC) study Law and Ethics Committee (reference: 83643).

Participants

Participants were recruited from the first generation of children born in the ALSPAC study (ALSPAC-G1). During Phase I of ALSPAC enrolment, 14,541 pregnancies in the former Avon Health Authority in the south-west of England with expected dates of delivery between 1 April 1991 and 31 December 1992 were recruited. Of these initial pregnancies, there was a total of 14,676 foetuses, resulting in 14,062 live births and 13,988 children who were alive at 1 year of age. A further 906 pregnancies were recruited during Phases II, III and IV respectively, resulting in an additional 913 children being enrolled. The total sample size is 15,447 pregnancies, of which 14,901 were alive at 1 year of age.

Further details on the cohort profile, representativeness and phases of recruitment are described in three cohort-profile papers.[20-22]

The ALSPAC study website www.bristol.ac.uk/alspac/ contains details of all the data that is available through a fully searchable data dictionary and variable search tool (http://www.bris.ac.uk/alspac/researchers/our-data/). Informed consent for the use of data collected via questionnaires and clinics was obtained from participants following the recommendations of the ALSPAC Ethics and Law Committee at the time.

Study data were collected and managed using REDCap electronic data capture tools hosted at the University of Bristol. REDCap (Research Electronic Data Capture) is a secure, web-based software platform designed to support data capture for research studies[26].

Participants were selectively recruited by SEP using maternal highest education achievements, with those whose mothers had a degree or higher level of attainment recruited to the high SEP group, and those whose mothers had educational attainment less than degree level recruited to the low SEP group.

Inclusion criteria were:

- Consume at least 6 units of alcohol per week.
- Be able to wear and use smartwatch from 12:00 noon until 22:00 in the evening (the timings of questions in the current version of the smartwatch system cannot be altered to accommodate shift work, etc)
- Have access to the Internet to complete the TLFB assessments.

Exclusion criteria (standard exclusion criteria used in alcohol studies to prevent harm during pregnancy and to those at risk of addiction) were:

- Pregnant or breast-feeding (self-reported). Note that if a participant becomes pregnant during the study, they must withdraw from the study at this point.
- History of alcohol or drug addiction (self-reported).
- Strong familial history of alcoholism defined as one or more immediate relative (parent, sibling) or more than one other relative (e.g., cousin, grandparent) (self-reported).

Clinic visits

All participants began the study with a visit to an ALSPAC clinic. In this session (approximately 1 hour in duration) participants;

- were given details of the study and provided written consent,
- were given the smartwatch and shown how to use the EMA application and maintain the smartwatch (including instructions on keeping it charged, not getting it wet, and removing before contact sports),
- were given instructions on how to use the online TLFB diary, provided with logon details, and asked about any events (e.g. birthdays, nights out, etc) that should be entered to the diary as reminders,
- provided with contact details of who to call if they encountered any issues during the study.

Participants completed a second clinic visit at the end of the study. In this session (approximately 30 minutes) participants;

- returned the smartwatch,
- completed a brief semi-structured interview about their experiences using the smartwatch and online TLFB diary led by the researcher,
- provided written final consent,
- were given their cash reimbursement.

13 participants (8 low SEP, 5 high SEP) completed the second clinic visit in person.
Covid-19 restrictions then meant the remaining 19 participants (11 low SEP, 8 high SEP) completed

their second clinic visit remotely. In these remote clinic sessions participants were interviewed by ALSPAC researchers by telephone, and arrangements were made for the smartwatch to be collected from their home by courier.

All participants were reimbursed £15 in shopping vouchers for each of the three months of the study they completed, with a bonus £15 if they completed all three months, so a total of £60 if they completed three months.

Smartwatch data collection

We developed a bespoke μEMA smartwatch application, which incorporated the elements proposed by Intille et al[8] for reducing participant burden, including using haptic rather than audible prompts to minimise disruption, and keeping questions brief with a limited set of answers that can be responded to with a single tap to reduce the effort required to respond.

To optimise the design of our app we worked closely with the ALSPAC Original Cohort Advisory Panel (OCAP) – a committee representing the variety of backgrounds and opinions of ALSPAC participants. Using a think aloud approach, panel members walked through the μEMA application and articulated any issues with the visual design, wording, and navigation of the application. A number of common themes were identified, which led to the implementation of the following changes in the final design of the μEMA application:

- The introduction of the 'My usual' short-cut to speed up the recording of alcoholic drinks that are the same type, quantity and context.

- The introduction of the 'Back' option so participants could go back and correct previous entries in the current epoch.

Participants were provided with a TicWatch C2 smartwatch running Android Wear 2.6, onto which was loaded our μEMA application.

At five timepoints (14:00, 16:00, 18:00, 20:00 and 22:00) every day of the study participants were asked about any alcoholic drinks they had consumed in the past 2 hour period, and the context of these drinking events. At each timepoint, the following set of questions and response options were presented to the participant:

Q1: Did you drink alcohol in the last two hours? Response options [Yes| No | My Usual*].
Q2: What were you drinking? Response options [Beer/cider | Wine | Spirits].
Q3: What size was your drink? Response options: Beer/cider: [Half pint | 330 bottle | Pint], Wine: [125ml | 175ml |250ml], Spirits: [Single | Double | Free pour].
Q4: Who were you with? Response options [Alone | In company].
Q5: Where were you? Response options [At home | Elsewhere].
Q6 Any more alcoholic drinks to be recorded? Response options [Yes | No | My Usual*].

*On the first day of the study participants were asked if they wanted to set up a 'My Usual' short cut to simplify entering data about their mostly frequently consumed alcoholic drink. If they selected this option they were asked questions 2 - 5 to specify their usual drink. Then whenever they selected 'My Usual', those drink details were recorded for that timepoint.

Examples of these questions rendered in our μEMA application are shown in Figure 2. The full sequencing of the questions is illustrated in Figure S1 in Supplementary Materials.

Figure 2. Example smartwatch μEMA application questions

In addition to the questions throughout the day, participants were also asked at 12:00 every day if they had missed recording any drinks the previous day. If they responded Yes, they were asked about those drinks using the same set of questions above.

Participant data was stored on the smartwatch device itself rather than uploaded to a cloud-based store to minimise the risk of loss of personal data. Given the corresponding lack of remote monitoring, to mitigate the risk of technical issues stopping data collection without our knowledge during the data collection period we swapped smartwatches at the end of month 1 and month 2. However, Covid-19 restrictions meant this was not possible for 15 participants (7 low SEP, 8 high SEP), and they were instead instructed to continue using their current smartwatch until the end of the study.

Online data collection

The Timeline Followback method uses a calendar populated in advance with participant defined diary entries (e.g. birthdays, theatre, nights out) to aid recall of alcohol consumption. Online methods for the self-administration of TLFB have been validated[23], and previous studies suggest 14 days is optimum length for balancing quality of data captured and participant burden[24], so we asked participants to complete an online TLFB every 14 days recording their alcohol consumption in the previous two weeks.

We implemented TLFB using RedCAP[25], which was hosted on the ALSPAC application platform. Participants were given a five-day window to complete each TLFB two week entry and sent text message reminders to complete their entry at the beginning of the five day window. In each two-week entry participants indicated how many units of alcohol were consumed every day. Note that participants cannot view data captured in the smartwatch μEMA application in order to use this to complete the TLFB. In keeping with standard practice for TLFB, no context questions were asked. The format of the TLFB diary entry screen is shown in Figure S2 in Supplementary Materials.

Interviews

In the second clinic session at the end of the study (either in person or remotely by telephone) each participant completed a semi-structured interview exploring their experiences in the study.

For both smartwatch μEMA and online TLFB methods participants were asked:

Q1/6: Overall, how would you rate your experience of using the smartwatch/online system during the study, on a scale from 1 (I didn't like it at all) to 10 (I really liked it)?
Q2/7: If you were asked to use the smartwatch/online system again in another study, how likely would you be to say yes, on a scale from 1 (I wouldn't want to use it again) to 10 (I'd really like to use it again)?
Q3/8: If you agreed to use the smartwatch/online system again in another study, what's the maximum length of time you'd be willing to use it for?
Q4/9: What things did you like about using the smartwatch/online system?
Q5/10: What things did you dislike about using the smartwatch/online system?

Participants were asked additional questions specifically about their experiences using the smartwatch μEMA system:

Q11: How did you feel about having to charge the smartwatch every night?
Q12: Did you use any of the other functions on the smartwatch during the study (e.g. step counter, etc)?
Q13: Did you pair the smartwatch with any smartphones during the study?

Participants were also asked about their attitudes towards future developments of the smartwatch μEMA system:

Q14: How would you feel if the smartwatch system used data from its motion sensors to work out when the best time is to ask you the questions (e.g. so it doesn't ask you questions when you're driving)?
Q15: How would you feel if the smartwatch system used data from its GPS location sensor (which tells the watch where the person is) to work out when the best time is to ask you the questions (e.g. so it doesn't ask you questions when you're at work)?

Engagement measures

In terms of reporting engagement we follow the approach used by Intille and colleagues[8]. We report compliance rates (defined as the proportion of scheduled questions to which participants respond) both across all participants who started the study, and for participants still active in the study. This enables us to explore whether a drop off in compliance is driven by attrition or some other factor.

In addition, for smartwatch EMA we report completion rate (defined as the proportion of questions delivered to active participants by the μEMA application to which they respond). This helps us understand whether failure to respond is driven by factors that prevent the questions being delivered (e.g. the smartwatch being switched off or the battery flat) or participants' propensity to respond to questions that are presented. Completion rate is not reported for online TLFB as questions were not actively presented to participants, so there is no notion of prompts that were or were not delivered to participants.

For all measures we report these for all participants, and for low and high SEP groups.

Missing data

We explored whether prompts were more likely to be missed at certain times of the day by comparing rates of missingness across 2-hour epochs. We also explored whether prompts were more likely to be missed on certain days of the week, using the same approach. Again, we did this for all participants and for low and high SEP groups.

Results

Participants

Our original aim was to recruit 40 participants (20 low SEP, 20 high SEP). Covid-19 restrictions meant recruitment had to be halted prematurely in March 2020, at which point 32 participants had been recruited. 13 were high SEP (mean age 27, range 26-28, 7 female = 54%) and 19 were low SEP (mean age 28, range 26-28, 13 female = 68%).

Engagement

Figure 3 shows the compliance rate for all participants who started the study for the smartwatch μEMA and online TLFB methods. Figure 4 shows the compliance rate for all participants still active in the study (i.e. those that have not dropped out yet) for the smartwatch μEMA and online TLFB methods. Figure 5 shows completion rate for all participants still in the study for the smartwatch μEMA method (not appropriate for TLFB methods as described above). Figure 6 shows participant attrition rate (percentage of participants still actively participating) for smartwatch μEMA and online TLFB methods.

Preprint JMIR Publications

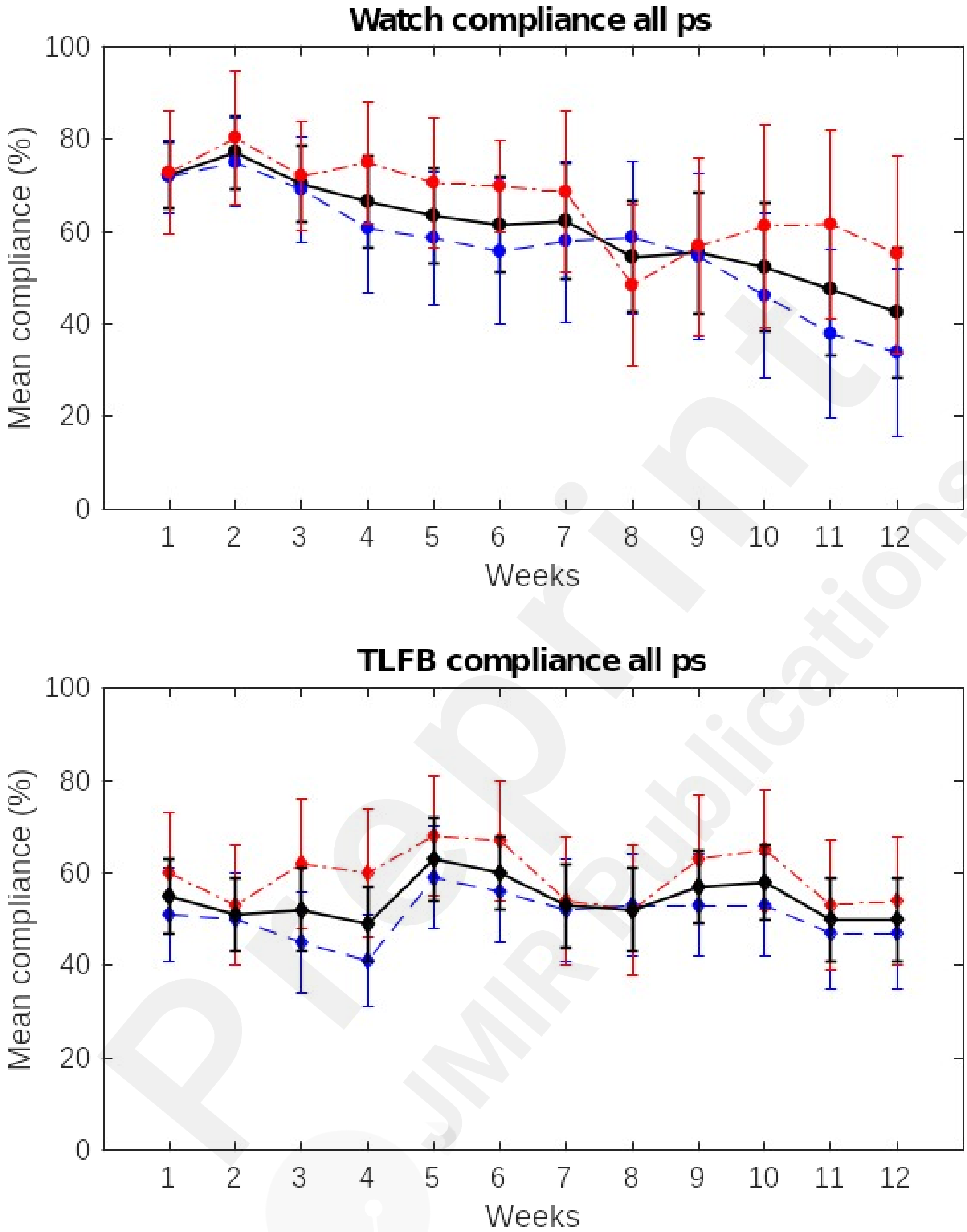

Figure 3. Smartwatch μEMA (top) and TLFB (bottom) compliance rates (mean and 95%CI) across all participants who started the study, for all (black solid line), high SEP (red dot dash line) and low SEP (blue dashed line) participants.

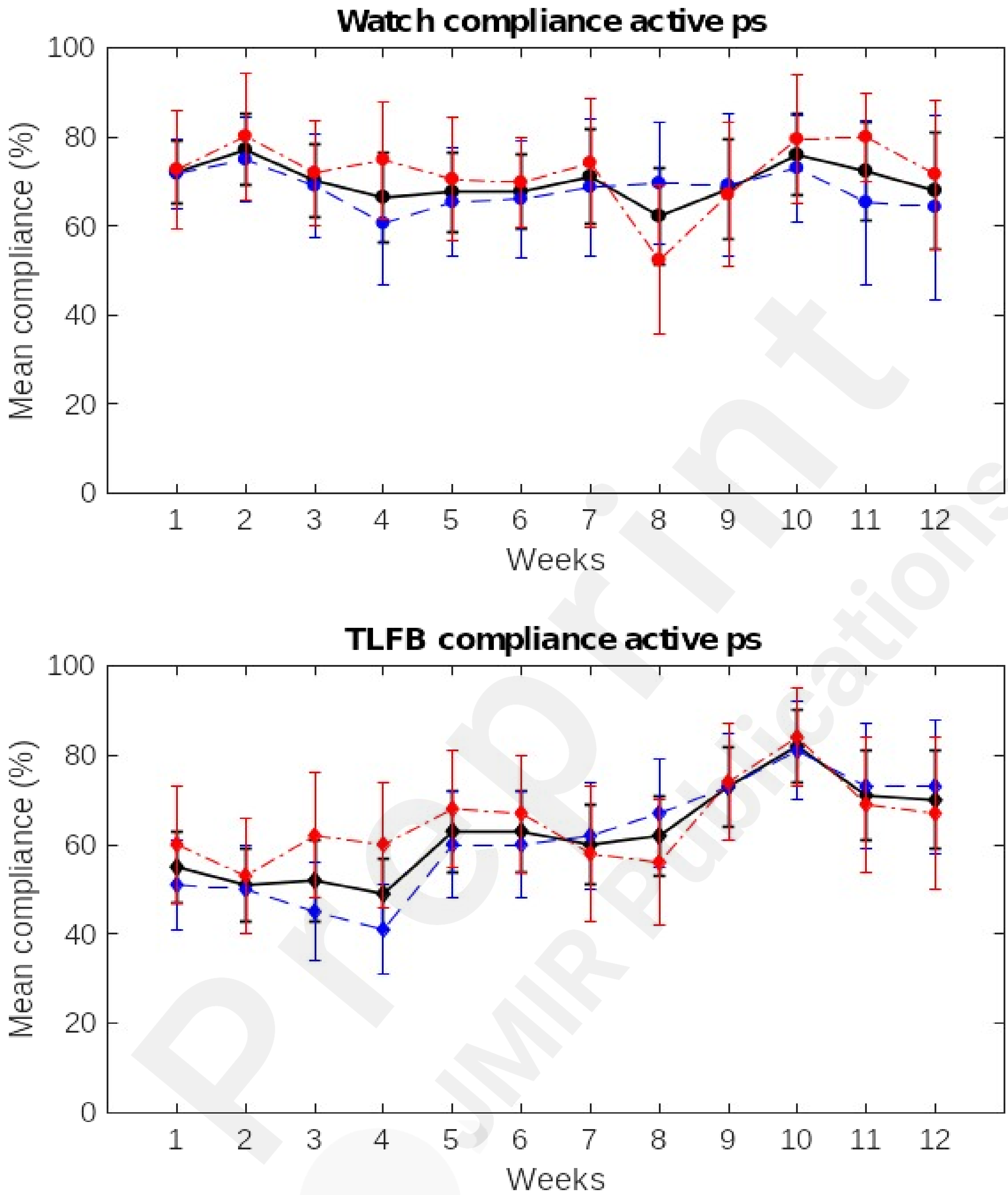

Figure 4. Smartwatch μEMA (top) and TLFB (bottom) compliance rates (mean and 95%CI) across all participants still active in the study for all (black solid line), high SEP (red dot dash line) and low SEP (blue dashed line) participants.

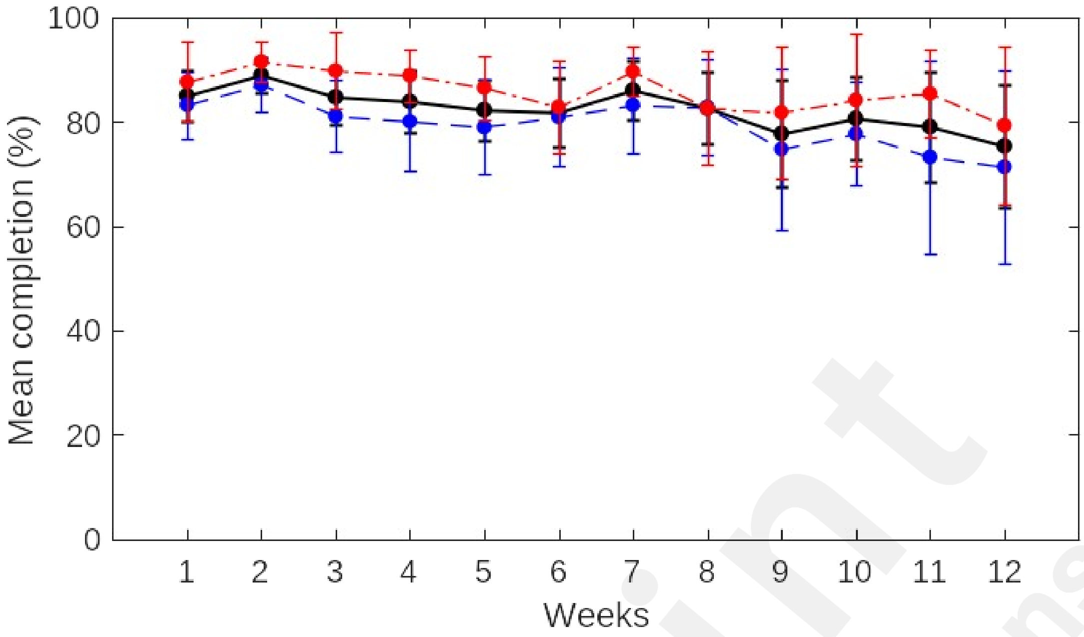

Figure 5. Smartwatch μEMA completion rate (mean and 95%CI) for all (black solid line), high SEP (red dot dash line) and low SEP (blue dashed line) participants.

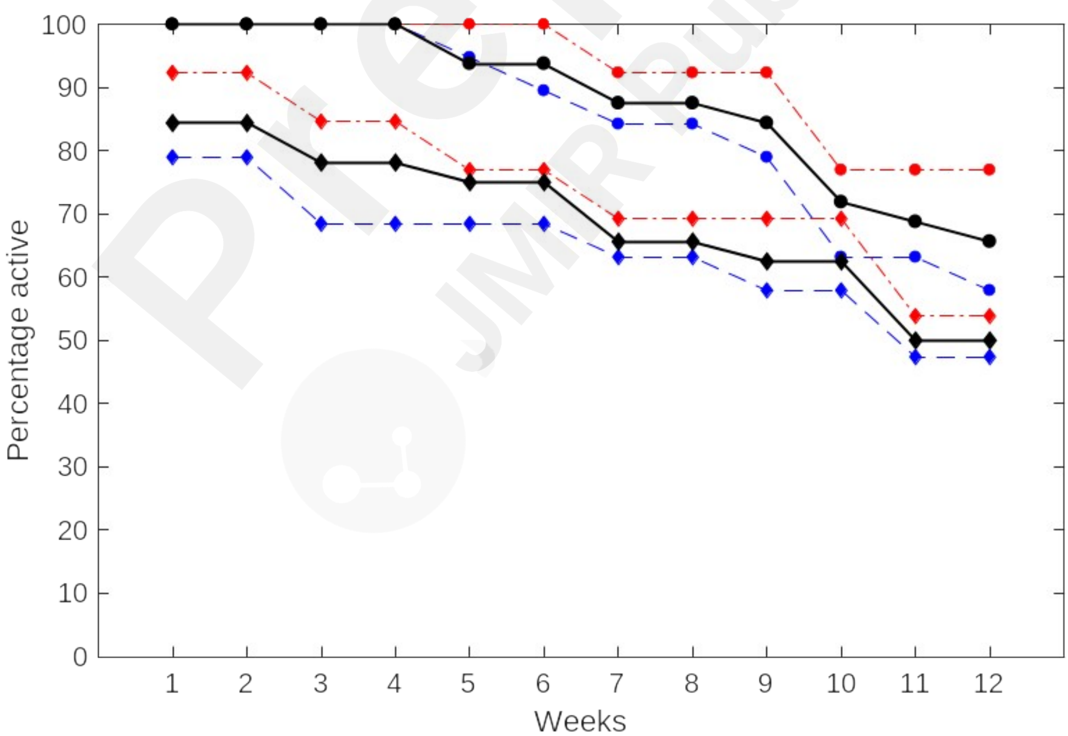

Figure 6. Participant attrition for smartwatch μEMA (circles) and online TLFB (diamonds) for all (black solid line), high SEP (red dot dash line) and low SEP (blue dashed line) participants.

Qualitative feedback

Preprint JMIR Publications

All participants completed interview questions about both online and watch methods for data collection. Participants' responses to the questions 'How would you rate your experience of using the smartwatch EMA/online TLFB from 1 (did not like) to 10 (really liked)?', 'How willing would you be to use the smartwatch μEMA /online TLFB again in another study from 1 (would not) to 10 (really like to)?', and 'How long would you be willing to use it for (in months)?' are shown in Table 1.

Table 1. Participants' ratings of the extent to which they liked using, would be willing to use again, and for how long, for the smartwatch μEMA (Watch) and online (TLFB) methods (medians with interquartile ranges).

	All SEP		Low SEP		High SEP	
	Watch	TLFB	Watch	TLFB	Watch	TLFB
Experience of using	8 (6-8)	7 (6-8)	8 (6-8)	7 (5-9)	8 (7-8)	8 (5-8)
Willing to use again	10 (7-10)	8 (7-10)	10 (7-10)	9 (7-10)	9 (6.5-10)	8 (6-10)
How long for (months)?	3 (3-12)	4 (3-12)	6 (3-12)	6 (3-12)	3 (2-5)	3 (2-5)

From participants' responses to questions 'What things did you like and dislike about using the smartwatch system?' distinct and recurrent (i.e. mentioned by two or participants) themes were identified, and are shown in Table 2, grouped by SEP. The equivalent for the online TLFB system is shown in Table S1 in supplementary materials. Also shown in Table 2 are participants' responses to the questions 'How did you feel about having to charge the smartwatch every night?', 'Did you pair the smartwatch with any smartphones during the study?', and 'Did you use any of the other functions on the smartwatch during the study (e.g. step counter, etc)?'

Table 2. Participants' likes and dislikes of the smartwatch μEMA method, attitudes to daily charging, attempts to pair with phone, and other watch features used, for low and high SEP participants.

	Low SEP	High SEP
Watch liked	n=19	n=13
Quick and easy	9 (47%)	6 (46%)
Being prompted	1 (5%)	4 (30%)
Having a watch	0 (0%)	2 (16%)
Liked the smartwatch	2 (10%)	2 (16%)
Additional smartwatch features	2 (10%)	0 (0%)
No need for extra device	0 (0%)	2 (16%)
Always to hand	4 (21%)	2 (16%)
Watch disliked	n=19	n=13
Timing/frequency of prompts	4 (21%)	3 (23%)
Technical issues	3 (16%)	4 (30%)
Battery life	4 (21%)	2 (16%)
Watch too big	3 (16%)	2 (16%)
Unsure how to handle missing drinks	1 (5%)	2 (16%)
Already have a watch/smartwatch	2 (10%)	1 (8%)
Attitude to needing to charge every night	n=15	n=11

No issues	11 (73%)	8 (73%)
Frustrating	4 (27%)	2 (18%)
Attempted to pair with smartphone	n=15	n=11
Did not attempt	13 (87%)	9 (82%)
Attempted but failed	2 (13%)	0 (0%)
Attempted and succeeded	0 (0%)	2 (18%)
Other watch features used	n=15	n=11
None	11 (73%)	5 (45%)
Step counter	3 (20%)	3 (27%)
Timer	1 (7%)	2 (18%)

From participants' responses to the question 'How would you feel if the smartwatch system used data from its GPS location sensor to work out when the best time is to ask you the questions (e.g. so it doesn't ask you questions when you are at work)?' and 'How would you feel if the smartwatch system used data from its motion sensors to work out when the best time is to ask you the questions (e.g. so it doesn't ask you questions when you are driving)?' all distinct themes were identified (even if raised by only one participant) and are shown in Table 3.

Note that because of a technical issue with the questionnaire delivery 6 participants (2 High SEP [1 female], 4 Low SEP [2 female) were not asked the questions about smartwatch charging, pairing and other functions, or about the potential use of GPS and motion sensors.

Table 3. Participants' views on using GPS data and motion sensor data to optimise smartwatch μEMA, for low and high SEP participants.

	Low SEP	High SEP
Attitudes to using GPS in future	n=15	n=11
No issues	13 (80%)	7 (64%)
Depends on context	3 (20%)	1 (9%)
OK if consented	0 (0%)	1 (9%)
Concerned about battery life	0 (0%)	1 (9%)
Scary	2 (13%)	0 (0%)
Intrusive	2 (13%)	0 (0%)
Attitudes to using motion sensors in future		
No issues	15 (100%)	9 (82%)
Concerns about reliability	0 (0%)	1 (9%)
Uncomfortable with idea	0 (0%)	1 (9%)

Missing data

The rates of missingness in all, low and high SEP participant groups for the smartwatch μEMA method in participants still active in study are show for days of the week in Table 4. There was considerable variability in each of the three groups, with overlapping confidence intervals for all

Preprint JMIR Publications

days.

Table 4. Percentage missingness by day of the week for all, high and low SEP groups (mean and 95% CI).

Weekday	All Participants (n=32)	High SEP (n=13)	Low SEP (n=19)
Monday	28.3 (21.2, 35.5)	23.8 (14.4, 33.3)	31.4 (20.8, 42.0)
Tuesday	25.6 (18.8, 32.4)	24.9 (15.9, 33.9)	26.1 (15.8, 36.5)
Wednesday	25.0 (19.4, 30.6)	23.6 (14.0, 33.2)	26.0 (18.5, 33.4)
Thursday	26.5 (21.1, 31.2)	24.2 (15.3, 33.1)	28.1 (20.6, 35.5)
Friday	25.2 (19.1, 31.2)	24.9 (14.6, 35.1)	25.3 (17.0, 33.7)
Saturday	28.0 (21.3, 34.6)	25.8 (15.0, 36.6)	29.5 (20.2, 38.8)
Sunday	28.2 (21.0, 35.4)	29.6 (18.2, 41.0)	27.2 (17.0, 37.4)

The rates of missingness in all, low and high SEP participant groups in participants still active in study are show for epochs within days in Table 5. Again there was considerable variability within each of the groups, with overlapping confidence intervals in each epoch. However, there was a noticeably lower level of overlap comparing the first four epochs (12:00-14:00,14:00-16:00,16:00-18:00,18:00-20:00) with the last epoch (20:00-22:00), which had the highest level of missing data.

Table 5. Percentage missingness by 2-hour epoch for all, high and low SEP groups (mean and 95% CI).

Epoch	All Participants (n=32)	High SEP (n=13)	Low SEP (n=19)
12:00 -14:00	26.7 (20.4, 32.9)	26.1 (17.3, 34.9)	27.0 (17.6, 36.4)
14:00 -16:00	23.6 (17.7, 29.5)	22.2 (13.7, 30.8)	24.5 (15.9, 33.2)
16:00 - 18:00	24.7 (18.5, 30.8)	23.1 (14.1, 32.0)	25.8 (16.8, 34.7)
18:00 - 20:00	26.0 (20.4, 31.5)	24.2 (15.5, 32.8)	27.2 (19.3, 35.1)
20:00 - 22:00	32.5 (25.5, 40.0)	30.7 (20.1, 41.2)	33.8 (23.5, 44.0)

Discussion

Principle findings

Comparing smartwatch μEMA and online TLFB engagement metrics

Using the most stringent measure of engagement (compliance rate for all participants who started

Preprint JMIR Publications

the study), compliance for the smartwatch μEMA method decreased gradually over the twelve weeks from around 70% in week 1 to approximately 45% in week 12, while compliance for the online TLFB method was flatter at around 50% over the twelve weeks. Using this metric, the smartwatch μEMA method performed better than the online TLFB method until around week 9, and from there on, the performance of the two methods was broadly similar.

Considering the compliance rates for all participants still active in the study (removing the effects of participant attrition), the compliance for the smart μEMA method was much flatter, staying around 70% for the whole 12 weeks. For the online TLFB method, compliance increased from approximately 50% in week 1 to around 80% in week 10, before dropping back to 70% in week 12. As before, the performance of the smartwatch μEMA method was better than that of the online TLFB method until around week 9, from which point there were only minor differences between them. Of note here is the increasing compliance with the online TLFB method in the first ten weeks of the study. While we have no firm explanation for this, one factor that may have contributed is that Covid-19 lockdowns would have meant participants spent more time at home as the study progressed, which may have improved ease of access to the online TLFB system. Access to the smartwatch μEMA would not be altered to the same degree, as the method was specifically designed with ease of access in mind, so a similar increase in compliance would not be expected.

Turning to completion rate of the smartwatch μEMA method (as discussed previously this is not applicable to the online TLFB method), for the first 8 weeks this was again broadly flat at around 85%, dropping to around 75% for the last 4 weeks. This demonstrates that, when the μEMA application is functioning correctly and the smartwatch device is kept charged so that prompts are delivered, participants showed a high level of propensity to respond. This declined only slightly over the 12-week period, suggesting longer data collection periods may be possible with acceptable levels of completion rates using this method.

Comparing engagement metrics for low and high SEP groups

Although there is considerable variation within groups, and the differences between the groups in most cases are relatively small (ranging between approximately 5 to 20%), almost all of the point estimates in participant compliance, active participant compliance, completion rate and attrition for both smartwatch μEMA and online TLFB methods indicate lower engagement for low SEP participants than high SEP participants.

Patterns of missing data

There was considerable variability in all three groups (all SEP, low SEP and high SEP) with overlapping confidence intervals when comparing levels of missing data by days of the week and by 2-hour epochs in a day in each group. Perhaps the most noticeable pattern was the lower overlap in confidence intervals observed in the all participant group when comparing the first four epochs (12:00-14:00,14:00-16:00,16:00-18:00,18:00-20:00) with the last epoch (20:00-22:00), which had the highest level of missing data. It may be possible that modifications to the μEMA application, such as making bedtime configurable for individual participants, would improve engagement and reduce the levels of missing data in this epoch.

In terms of differences between high and low SEP groups, again there was considerable overlap of confidence intervals in the groups for all days and epochs. However, the point estimates for proportion of data missing for low SEP participants were slightly higher than those for high SEP participants for every epoch, and for all days except Friday and Sunday.

Qualitative feedback

Across all participants, the overall experience of using both methods was scored highly, with a slight preference for smartwatch μEMA (8/10) compared with online TLFB (7/10). The low SEP group scored the smartwatch μEMA method slightly higher (8/10) than the online TLFB method (7/10). The high SEP scored both methods the same (8/10).

For willingness to use again, again both methods scored highly, with all participants expressing a preference for smartwatch μEMA (10/10) compared with online TLFB (8/10). This same pattern was seen in both the high SEP group (9/10 vs 8/10) and low SEP group (10/10 vs 9/10)

In terms of the amount of time participants would be willing to use the methods again, for all participants online TLFB scored slightly longer (4 months compared with 3 months for smartwatch μEMA). Within high and low SEP groups there were no differences between the two methods, however there were differences between groups, with the low SEP group scoring both methods 6 months, and the high SEP group scoring both methods 3 months.

Focusing on the smartwatch μEMA method, the factor liked far more than any other was the ease and speed of use. There was a broader spread of disliked factors, which included the timing and frequency of prompts, technical issues with the system, battery life, and the size of the smartwatch. In both high and low SEP groups 73% of participants had no issues with the need for daily charging, while some (27% low SEP, 18% high SEP) found it frustrating. In both SEP groups more than 80% of participants did not try to pair the smartwatch to their own smartphone, in line with study instructions to not attempt this. In the low SEP group the 13% who attempted to pair the smartwatch with their smartphone failed, while in the high SEP group the 18% who attempted succeeded. High SEP participants (55%) explored other features of the smartwatch more than low SEP participants (27%), with step counter and timer the features most explored.

In terms of attitudes to using additional sensors on the smartwatch to enhance data capture, the low SEP group was more accepting of the use of both GPS (80% low SEP had no issues compared with 64% of high SEP) and motion sensors (100% low SEP compared with 82% of high SEP).

Comparison with Prior Work

Alcohol EMA studies

In Perski and colleagues' 2022 systematic review[2] and meta-analysis of EMA studies of health behaviours, adherence (defined as the 'average percentage of EMA assessments completed out of the available EMA assessments'), was pooled across 175 alcohol EMA studies. The median percentage adherence was 84%, with an interquartile range 77 to 91%, with studies using smartphone-based EMA methods having overall better levels of adherence than other EMA methods (e.g. online, paper and pen). No studies using wrist-worn wearable EMA methods were included. From the engagement metrics we used, completion rate seemed the closest in nature to their adherence metric. This was around 80% across the 12 weeks of our study, so within the range they reported.

In his review specifically of alcohol EMA methods, Piaseck[26] identified a number of studies that have used what are referred to as high resolution EMA methods (mainly smartphone-based), which like the current study prompted participants for responses throughout the day, and which varied in overall duration from a few days up to eight weeks. These had 'response rates' ranging from 63 to 90%. It was less clear which of our metrics to compare with this. Our completion rate (around 80% over the duration of the study) and active participants compliance rates (around 70%) sit within this range. Our all-participant compliance rate began around 70%, dropping to below 60% in week 8, so again was broadly within that range. Our study ran beyond 8 weeks, and our all-participant compliance declined to 45% at week 12, suggesting study duration may have greater impact on engagement for studies beyond a certain duration.

Recently, Howard and Lamb[27] used text message initiated survey smartphone-based EMA methods to study alcohol consumption in undergraduates. They reported high levels of compliance for active participants, starting at 89% in week 1, declining gradually to 70% in week 14. Our active participant compliance rates were lower, and varied around 70% across the 12 week period. While it is unclear if any one factor was responsible for this difference, possibilities include that the prompts and reminders delivered by phone in Howard and Lamb's study were audible and possibly more strident than the smartwatch haptic prompts in the current study, and that Howard and Lamb's rewards strategies were based on completion of individual surveys whereas ours was based on months active in the study, although previous meta-analyses have indicated differing reward strategies are not related to variations in engagement[11,12].

Smartwatch EMA studies

In a smartwatch EMA study with an assessment period close to that of the current study, Beukenhorst and colleagues[9] used smartwatch μEMA methods to capture self-report data about pain in patients with knee osteoarthritis, using 4 or 5 questions by day over 13 weeks. Their attrition rate was similar to that observed in the current study, with the number of active participants falling steadily to around 70% by week 10, although at the end of the 13 week period this had fallen to 42% (ours fell to 65% by week 12). The compliance rate for all participants was also similar to the current study, starting at around 75% then gradually falling off to around 50% by week 10 and 45% by week 12. The completion rates were also similar, starting at around 85% and keeping around 80% over the duration of the study. Participants also identified short battery life as a current barrier to smartwatch use.

In what may be the longest smartwatch EMA study to date, Ponnada and colleagues[10] captured self-report data of physical activity and affect in young adults over a period of 12 months. They used smartwatch-based μEMA methods to capture high density data over 24 4 day burst periods, spread throughout the 12 months. In their report of preliminary findings at the 6-month point, the mean overall compliance of μEMA in the burst periods for active participants was 67%, so similar to the 70% we observed across the 12 weeks of the current study. The mean overall completion reported at 6 months for these burst periods was 79%, so again similar to the rate observed in the current study, which began at 85% and reduced to 75% by week 12.

Effects of high and low SEP

In terms of the impact of SEP on engagement in alcohol EMA studies, Howard and Lamb[27] reported no significant differences in the numbers of students having parents with college education (similar to the SEP measure we used here), across their four responder categories ('poor', 'adequate', 'good', 'super'). However, it is noticeable that the proportions of students with college educated parents were higher in the 'good' and 'super' responder categories than in the 'poor' category, indicating higher SEP may potentially have played some role in increased engagement, as we observed in the current study.

Considering the impact of SEP on compliance more broadly in alcohol studies, Thern and Landbarg[28] reported that for self-reported factors such as self-rated health, health-related quality of life, and level of social support received, low SEP participants had approximately twice as much missing data as high SEP participants. While the current feasibility study cannot claim to provide evidence of higher levels of missing data in low SEP than high SEP participants, the point measures observed here do fit with this pattern.

Strengths and Limitations

To the best of our knowledge, this is the first study to explore the feasibility of using smartwatch-based μEMA methods to capture high temporal density data about alcohol consumption over extended periods.

While the smartwatch-based EMA app we produced was heavily based on the μEMA methods originally developed by Intille and colleagues[8], an important element of a true μEMA implementation is that only one question is presented to the participant at a time. In our implementation we chained questions to ask about type and size of drinks and the context in which they were consumed, and consequently our approach should perhaps be referred to as modified-μEMA. However there seems to have been little impact on engagement as a result of our modified approach.

Unfortunately, Covid-19 meant participant recruitment had to be halted early, reducing our sample size, and giving an unequal mix of high and low SEP participants. It also means that the results here may not be as generalisable as those seen observed outside of a global pandemic.

Our participants were recruited from a UK birth cohort study[20]. Recruiting from a cohort study has the benefit that in future work we will be able to explore new methods for combining the high temporal density longitudinal data on alcohol consumption from our new smartwatch EMA methods with the wide variety of phenotypic and genetic data available for these participants, in order to explore the impact of alcohol consumption on health in new ways. One potential downside of recruiting from a cohort of this kind is that participants are used to regular assessments, so may possibly have been more motivated to engage with the study than participants not involved in cohort studies[20].

There is increasing awareness that the wide range of adherence metrics reported in EMA studies can make meaningful comparisons of findings across the EMA literature difficult[10,11]. We selected the adherence metrics reported here primarily to enable comparisons with existing smartwatch EMA studies, but acknowledge these do not always enable straightforward comparisons with other existing EMA studies.

Recommendation for Future Studies

The biggest factor impacting adherence in the current study was participant attrition. While feedback from participants in the end of study interview provided some clues as to what may have caused this, further studies to explore reasons for attrition in these new methods would be beneficial.

The consistency of the differences we observed between high and low SEP groups in compliance, completion and attrition suggests further work to explore the scale of these differences, and which factors are driving these, is warranted. It also indicates that the design of future systems of this kind would benefit from early-stage involvement of users from a range of SEPs. In addition to SEP, future work of this kind should consider other demographic factors, such as age, that may impact engagement with, and the quality of data collected using these new methods.

This study has focused on feasibility aspects of using new smartwatch-based μEMA methods for capturing data on alcohol consumptions. Future studies should also explore the validity of the data captured, and how both feasibility and validity vary for other health related behaviours beyond alcohol consumption.

As a community developing new EMA methods, it is important for us to strive to use metrics that enable the widest, most meaningful comparisons of our findings. In terms of adherence, this is now more complex than ever, as new EMA methods using latest technologies clearly benefit from using specific metrics (e.g. completion and compliance) that enable researchers to tease apart technical issues and participant propensity to respond. Now may be a good time to revisit previous work that has proposed standards for reporting in EMA studies [29,30].

Preprint JMIR Publications

Conclusions

In this study participants had higher levels of engagement (higher compliance and completion rates and lower levels of attrition) when using smartwatch-based μEMA methods then when using online Timeline Followback methods for recording alcohol consumption. These levels of engagement compared well with previous studies of new smartwatch μEMA methods. There was little difference in participants' ratings of their experience using the methods, or the length of time they would be willing to use either method again, though they did express a preference for using the smartwatch μEMA method rather than the online TLFB. This demonstrates the feasibility and acceptability of using smartwatch μEMA methods for capturing alcohol consumption.

These methods have the benefits of capturing higher temporal density, longitudinal data on alcohol consumption, together with data about the context of drinking events, that, as with EMA methods in general, are less susceptible to the recall and bias errors that are known to affect other methods for capturing data on alcohol consumption (though it is worth remembering that there are also large effects of social desirability bias in alcohol reporting which EMA methods will not address, so some underreporting may still exist).

Future studies should explore the factors impacting participant attrition (the biggest reason for reduced adherence), and the validity of alcohol data captured with these methods. While we would not claim our study provides firm evidence of differences in engagement between high and low SEP participants, we would suggest the consistent pattern of lower point measures of engagement observed in low SEP than high SEP participants indicates further work is warranted exploring the existence, scale and causes of such differences.

Acknowledgements

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses.

Preprint JMIR Publications

References

1. Shiffman, S., Stone, A.A., Hufford, M.R. (2008). Ecological momentary assessment. *Annu Rev Clin Psychol. (4),*1-32.
2. Perski, O., Keller, J., Kale, D., Asare, B. Y., Schneider, V., Powell, D., Naughton, F., Ten Hoor, G., Verboon, P., & Kwasnicka, D. (2022). Understanding health behaviours in context: A systematic review and meta-analysis of ecological momentary assessment studies of five key health behaviours. *Health Psychology Review*, *16*(4), 576–601.
3. Embree, B. G., & Whitehead, P. C. (1993). Validity and reliability of self-reported drinking behavior: dealing with the problem of response bias. *Journal of Studies on Alcohol*, *54*(3), 334–344.
4. Midanik L. (1982). The validity of self-reported alcohol consumption and alcohol problems: a literature review. *British Journal of Addiction*, *77*(4), 357–382.
5. Gmel, G., & Daeppen, J. B. (2007). Recall bias for seven-day recall measurement of alcohol consumption among emergency department patients: implications for case-crossover designs. *Journal of Studies on Alcohol and Drugs*, *68*(2), 303–310.
6. Hammerton, G., Lewis, G., Heron, J., Fernandes, G., Hickman, M., & Lewis, G. (2023). The association of alcohol dependence and consumption during adolescence with depression in young adulthood, in England: a prospective cohort study. *The Lancet. Psychiatry*, *10*(7), 490–498.
7. Fridberg, D. J., Lee, Z., Fischer, A. M., Cursio, J. F., & King, A. C. (2023). Assessing real-time positive subjective effects of alcohol using high-resolution ecological momentary assessment in risky versus light drinkers. *Alcohol, Clinical & Experimental Research*, 10.1111/acer.15130.
8. Intille, S., Haynes, C., Maniar, D., Ponnada, A., & Manjourides, J. (2016). μEMA: Microinteraction-based Ecological Momentary Assessment (EMA) Using a Smartwatch. *Proceedings of the ... ACM International Conference on Ubiquitous Computing. UbiComp (Conference)*, *2016*, 1124–1128.
9. Beukenhorst, A., Howells, K., Cook, L., McBeth, J., O'Neill, T., Parkes, M., Sanders, C., Sergeant, J., Weihrich, K., Dixon, W. (2020). Engagement and Participant Experiences With Consumer Smartwatches for Health Research: Longitudinal, Observational Feasibility Study. *JMIR Mhealth Uhealth, 8*(1):e14368.
10. Ponnada, A., Wang, S., Chu, D., Do, B., Dunton, G., & Intille, S. (2022). Intensive Longitudinal Data Collection Using Microinteraction Ecological Momentary Assessment: Pilot and Preliminary Results. *JMIR Formative Research*, *6*(2), e32772.
11. Ottenstein, C., & Werner, L. (2022). Compliance in Ambulatory Assessment Studies: Investigating Study and Sample Characteristics as Predictors. *Assessment*, *29*(8), 1765-1776.
12. Jones, A., Remmerswaal, D., Verveer, I., Robinson, E., Franken, I. H. A., Wen, C. K. F., & Field, M. (2019). Compliance with ecological momentary assessment protocols in substance users: a meta-analysis. *Addiction (Abingdon, England)*, *114*(4), 609–619.
13. Rintala, A., Wampers, M., Myin-Germeys, I., & Viechtbauer, W. (2020). Momentary predictors of compliance in studies using the experience sampling method. *Psychiatry Research*, *286*, 112896.
14. https://www.gov.uk/government/statistics/alcohol-specific-deaths-in-uk-registered-in-2018
15. Grittner, U., Kuntsche, S., Gmel, G., & Bloomfield, K. (2013). Alcohol consumption and social inequality at the individual and country levels--results from an international study. *European Journal of Public Health*, *23*(2), 332–339.
16. Mäkelä, P., & Paljärvi, T. (2008). Do consequences of a given pattern of drinking vary by socioeconomic status? A mortality and hospitalisation follow-up for alcohol-related causes of the Finnish Drinking Habits Surveys. *Journal of Epidemiology and Community Health*, *62*(8), 728–733.
17. Mackenbach, J. P., Kulhánová, I., Bopp, M., Borrell, C., Deboosere, P., Kovács, K., Looman, C. W., Leinsalu, M., Mäkelä, P., Martikainen, P., Menvielle, G., Rodríguez-Sanz, M., Rychtaříková, J., & de Gelder, R. (2015). Inequalities in Alcohol-Related Mortality in 17 European Countries: A Retrospective Analysis of Mortality Registers. *PLoS Medicine*, *12*(12), e1001909.

18. Stone, A. A., Schneider, S., & Smyth, J. M. (2023). Evaluation of Pressing Issues in Ecological Momentary Assessment. *Annual Review of Clinical Psychology*, *19*, 107–131.
19. Skinner, A., Adams, S., Stone, C. J., Munafo, M. R., Timpson, N. J., Tilling, K., & Lawlor, D. A. (2020, March 10). Assessing the feasibility of using smartwatch-based Ecological Momentary Assessment techniques for determining longitudinal patterns and contexts of alcohol use in UK adults. https://doi.org/10.17605/OSF.IO/4ZA26
20. Boyd, A., Golding, J., Macleod, J., Lawlor, D. A., Fraser, A., Henderson, J., Molloy, L., Ness, A., Ring, S., & Davey Smith, G. (2013). Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology*, *42*(1), 111–127. https://doi.org/10.1093/ije/dys064.
21. Fraser, A., Macdonald-Wallis, C., Tilling, K., Boyd, A., Golding, J., Davey Smith, G., Henderson, J., Macleod, J., Molloy, L., Ness, A., Ring, S., Nelson, S. M., & Lawlor, D. A. (2013). Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *International Journal of Epidemiology*, *42*(1), 97–110. https://doi.org/10.1093/ije/dys066.
22. Northstone, K., Lewcock, M., Groom, A., Boyd, A., Macleod, J., Timpson, N., & Wells, N. (2019). The Avon Longitudinal Study of Parents and Children (ALSPAC): an update on the enrolled sample of index children in 2019. *Wellcome Open Research*, *4*, 51. https://doi.org/10.12688/wellcomeopenres.15132.1.
23. Rueger, S. Y., Trela, C. J., Palmeri, M., & King, A. C. (2012). Self-administered web-based timeline followback procedure for drinking and smoking behaviors in young adults. *Journal of Studies on Alcohol and Drugs*, *73*(5), 829–833.
24. Fiellin, D. A., McGinnis, K. A., Maisto, S. A., Justice, A. C., & Bryant, K. (2013). Measuring alcohol consumption using Timeline Followback in non-treatment-seeking medical clinic patients with and without HIV infection: 7-, 14-, or 30-day recall. Journal of Studies on Alcohol and Drugs,74(3), 500-4.
25. Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, *42*(2), 377–381. https://doi.org/10.1016/j.jbi.2008.08.010.
26. Piasecki T. M. (2019). Assessment of Alcohol Use in the Natural Environment. *Alcoholism, Clinical and Experimental Research*, *43*(4), 564–577.
27. Howard, A. L., & Lamb, M. (2024). Compliance Trends in a 14-Week Ecological Momentary Assessment Study of Undergraduate Alcohol Drinkers. *Assessment*, *31*(2), 277–290.
28. Thern, E., & Landberg, J. (2021). Understanding the differential effect of alcohol consumption on the relation between socio-economic position and alcohol-related health problems: results from the Stockholm Public Health Cohort. *Addiction, 116(4):799-808.*
29. Stone, A. A., & Shiffman, S. (2002). Capturing momentary, self-report data: a proposal for reporting guidelines. *Annals of Behavioral Medicine : a publication of the Society of Behavioral Medicine*, *24*(3), 236–243.
30. Liao, Y., Skelton, K., Dunton, G., & Bruening, M. (2016). A Systematic Review of Methods and Procedures Used in Ecological Momentary Assessments of Diet and Physical Activity Research in Youth: An Adapted STROBE Checklist for Reporting EMA Studies (CREMAS). *Journal of Medical Internet Research*, *18*(6), e151.

Supplementary Materials

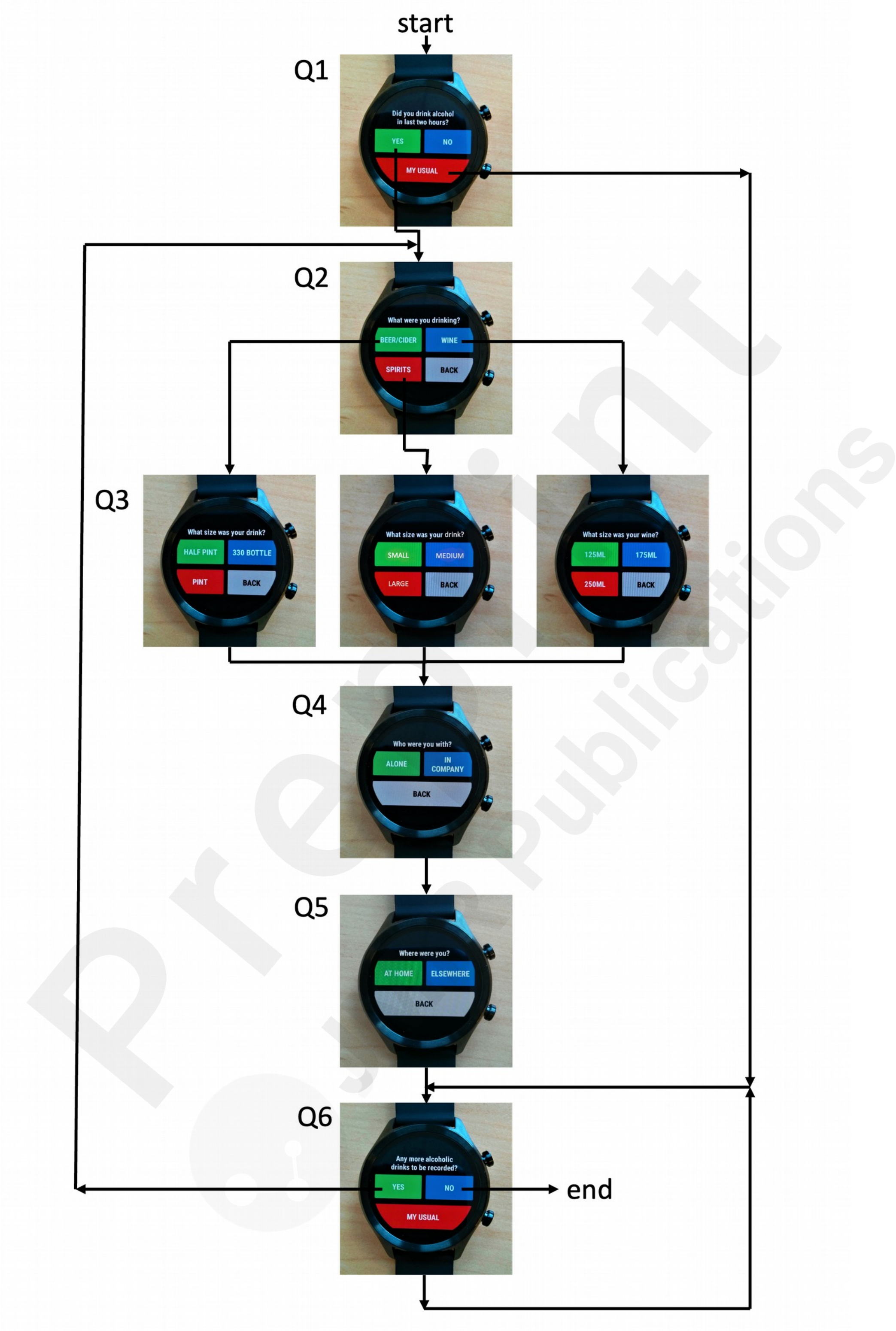

Figure S1. AlcoWatch μEMA smartwatch question formatting and sequencing

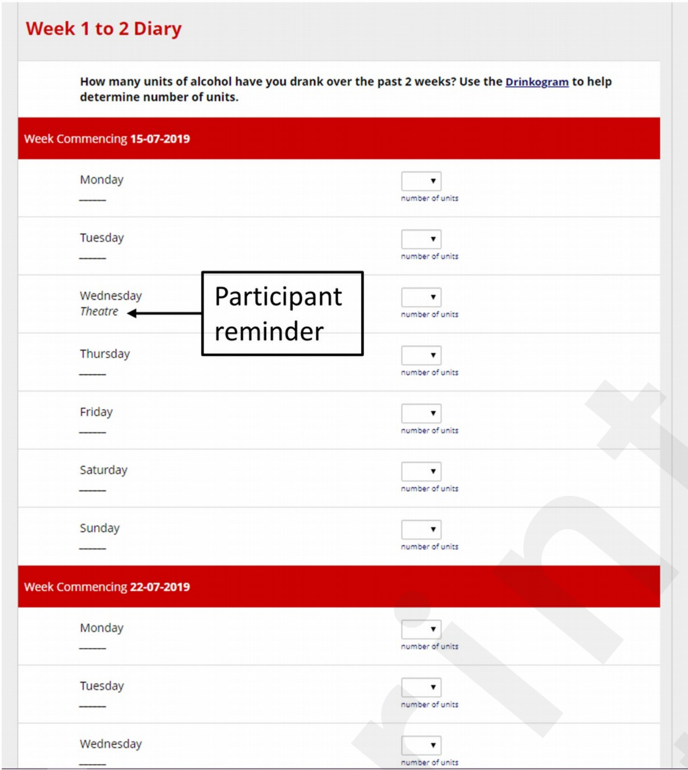

Figure S2. Timeline Follow Back diary format

Table 2. High and low SEP participants' likes and dislikes of the online TLFB method.

	Low SEP	High SEP
TLFB liked	n=19	n=13
Quick and easy	11 (58%)	5 (38%)
The event prompts	1 (5%)	2 (15%)
TLFB disliked	n=19	n=13
Difficult to remember drinks consumed	6 (32%)	5 (38%)
Difficult to remember to complete	0 (0%)	4 (30%)
Limited window in which to complete	6 (32%)	3 (23%)
Would have liked shortcuts	0 (0%)	2 (15%)
Wanted more info on units in drinks	2 (10%)	0 (0%)
Technical issues	1 (5%)	1 (8%)