Feasibility and Acceptability of a Wrist-worn Transdermal Alcohol Biosensor to Collect Data in the Field

Jennifer E. Merrill
Brown University
jennifer_merrill@brown.edu

Rachel L. Gunn
Brown University
rachel_gunn@brown.edu

Anne M. Neary
Brown University
anne_neary@brown.edu

Timothy Souza
Brown University
timothy_souza@brown.edu

Nancy P. Barnett
Brown University
nancy_barnett@brown.edu

Abstract

Transdermal alcohol monitoring allows for continuous, non-invasive, objective, and remote measurement of alcohol consumption. We evaluated feasibility and acceptability of participant use of the BACtrack Skyn biosensor bracelet in daily life. Heavy drinkers (n=20) wore the Skyn and self-reported drinking behavior for 7 days, followed by an individual interview. Recruitment and retention benchmarks were met, supporting feasibility. Participants provided both positive and constructive feedback on the Skyn during interviews, and usability of the bracelet was deemed "good". Most missing data were inconsequential (<5 mins), with data available 85% of the time participants were asked to wear it. Missing data was largely expected and due to bracelet removal during bathing or charging. Overall, results indicate promise in our ability to integrate this tool into research and/or clinical practice, passively and objectively monitoring alcohol use in participants and/or patients with minimal burden.

1. Introduction

Despite enormous progress in treatment of alcohol misuse, considerable public health burden associated with alcohol use in the US remains. Effectively intervening on alcohol misuse requires a multifaceted approach responsive to dynamic patterns of alcohol use in the real world. Self-reports of drinking behavior, even those collected in real-time [1-3], are subject to measurement error (e.g., accurate estimation of drink size) and other mistakes such as forgetting or being too intoxicated to report; these concerns are compounded at high levels of use [4]. Recommendations for substance use research and treatment are to use self-report in combination with objective measurements (e.g., using blood or urine biomarkers [5]). However, alcohol is quickly metabolized so these methods can only detect alcohol within hours or days after it is consumed. Current direct (EtG, PEth, breath) and indirect (GGT)

biomarkers are not adequate as they do not have a long window of detection, and (with the exception of breathalyzer) cannot provide information about the timing or volume of alcohol consumed. As such, objective measures are needed to provide accurate representations of alcohol use in real time, and reduce reliance on fallible self-report.

Recently developed biosensors provide continuous estimates of transdermal alcohol concentration (TAC) based on the concentration of alcohol in perspiration on the skin [6, 7]. This approach – transdermal alcohol monitoring – is the only current method for continuously monitoring alcohol use. The SCRAM Continuous Alcohol Monitor (SCRAM CAM; Alcohol Monitoring Systems, Littleton, CO) bracelet has been extensively validated over the past decade [8-10]. Until recently, the SCRAM CAM was the only device available for purchase or research, but it has limited use in voluntary populations due to its appearance (similar to monitoring ankle bracelets used in corrections). In 2015, NIAAA issued the Sensor Challenge in the interest of developing alternative devices to objectively measure alcohol. The winner, the BACtrack Skyn (Fig. 1), is now available for research use.

The Skyn provides a novel opportunity to remotely monitor alcohol use. Notably, *continuous* objective data may contribute new insights about alcohol use patterns and associated risks by providing more granular information about patterns of alcohol consumption than previously available. These devices have potential for a large

Figure 1. Skyn bracelet

impact on the field for both research and clinical purposes, but their value will be compromised if initial attempts to evaluate them are not scientifically robust.

Research to date on the Skyn has shown initial validity of the device in detecting alcohol consumption in both the laboratory and field. A laboratory study (N =

HĭCSS

30) found strong correlations between breath alcohol concentration (BrAC) and Skyn TAC readings [11]. Two studies (N=10, N=12) testing the Skyn in the field found evidence for initial acceptability and detection of self-reported drinking [12]. In a final larger study (N = 73) testing the Skyn in the laboratory and comparing readings to repeated BrAC readings, a strong correlation (r=.907) was observed between BrAC and estimated blood alcohol concentration (eBAC) readings derived from TAC using machine learning models [13]. Importantly, these models also outperformed similar models using SCRAM data, indicating that the greater granularity of these new devices likely provide more accurate depiction of real-world drinking. However, these initial tests used the early Skyn prototypes and had fairly high device failure rates (e.g., 16% reported by Fairbairn et al., 2020) [13, 14]. These studies attributed failures to a combination of device malfunction, user issues, and battery failure, noting the that the earlier prototypes were delicate and hand-assembled. Newer Skyn models need additional testing in the field.

In sum, the Skyn has the potential to provide accurate and noninvasive tracking of alcohol use to supplement or replace retrospective or real-time reports without the stigmatizing appearance of the SCRAM. However, few studies document the performance of the device and its accompanying software using naturalistic drinking episodes, and additional work is needed to characterize participant reactions to wearing the device.

1.1. The Present Study

We conducted preliminary work with a new alcohol biosensor - the BACtrack Skyn. Field work allowed us to evaluate the sensor in naturalistic drinking situations. We addressed participant reactions to the device and evaluated performance of the sensor and its accompanying software. Specific aims were to (1) evaluate feasibility and acceptability of the Skyn, and (2) characterize device functionality and missing data. This work was designed to contribute to scientific knowledge and methodology by affording "proof of concept" for Skyn feasibility and acceptability.

2. Methods

2.1. Participants

Twenty participants were recruited from the community using inclusion criteria: (a) age 21-55, (b) body mass index (BMI) 18.5-30 to avoid lower detection of alcohol use in individuals with high BMI [8] (c) ownership and use of an iPhone version 8.0 or higher with a data plan, as Skyn software is only compatible with iOS, (d) drinking at least 2x/week in the past month, (e) at least four past-month heavy drinking episodes (4+/5+ drinks for women/men), and (g) at least one episode in the past month where eBAC reached .10 g/dl. The alcohol use criteria were to ensure that there would be some episodes of heavy drinking during data collection. Exclusion criteria included: (a) being in or seeking treatment for alcohol or substance use, (b) an Alcohol Withdrawal Symptom Checklist (AWSC [15]) score above 23, (c) current use of drugs other than marijuana, (d) living outside a 10-mile radius of the study site as materials needed to be delivered by research staff, and (e) current or planned pregnancy. Laboratory-based alcohol administration was planned as a part of this study, which required some of the above eligibility criteria (e.g., withdrawal). This laboratory work was not completed due to COVID-19 restrictions.

2.2. Procedures

Procedures were approved by Brown University's Institutional Review Board and conducted remotely during the COVID-19 pandemic. The study was advertised using flyers, internet sites (e.g., Facebook), and the university's morning email. Interested individuals were directed to complete an online screening survey. Zoom videoconference meetings were held with those who passed initial eligibility screening, to confirm drinking level eligibility using a Timeline Follow-Back (TLFB) interview [16] and to provide information about the study. For those eligible and interested, informed consent was completed using Qualtrics survey data collection software.

2.2.1. Remote Orientation Session. The day of the orientation session, we dropped off a Skyn bracelet, charger, and a folder of informational handouts at a time and place of the participant's convenience. During a remote Zoom session, all study procedures and expectations for participation were explained using a PowerPoint Presentation. Baseline measures were collected via Qualtrics. Immediately following the orientation session, participants began a 1-week field trial of wearing the Skyn and reporting their alcohol use in app-based surveys. Researchers monitored the data from both the self-reports and the Skyn sensor daily and communicated with participants to ensure procedures were followed and to troubleshoot technical issues.

2.2.2. BACtrack Skyn. The Skyn bracelet resembles an activity bracelet (e.g., Fitbit) and is worn on the wrist continuously. It takes a reading from the perspiration on the surface of the skin every 20 seconds and stores it for later retrieval, which occurs by syncing the bracelet to the accompanying smartphone application. Participants were guided to attach the Skyn to their nondominant wrist and to wear the Skyn at all times except when

bathing or while it was charging. Participants were instructed on how to charge the device, which required powering it off. They were asked to charge the device once a day during typical non-drinking hours, which were identified for each person to help them plan. They were instructed to avoid (or self-report) contact with alcohol-based products (e.g., lotions, perfumes), given known interference with transdermal alcohol sensor readings. Participants were guided to install the Skyn data collection app and provided a unique username and password to log into the app. They were instructed on how to sync the bracelet by opening the app at least once every 24 hours. Participants were informed that Skyn readings (available within the app, though participants were not shown where to find them) were meaningless, and instructed not to use them to attempt to determine intoxication levels and/or ability to drive safely.

2.2.3. Ecological momentary assessment (EMA). All EMA assessments were collected via the mobile application PiLR (http://pilrhealth.com), developed by MEI Research, Inc. This app is fully customizable for conducting time-based and event-contingent surveys, and supports logic branching to ask follow-up questions when certain events are noted. Participants were trained in standard drink sizes, and how to calculate the number of standard drinks in their typical drinks. They were guided to install the app and to log in using provided credentials, and were instructed on how to use the EMA app to record their drinking. During the orientation session they practiced completing EMA reports.

Participants were asked to provide four types of reports: (1) *Event-contingent drink start surveys* were available for participants to make an entry immediately upon starting each drink. These reports recorded the time of drink start, type of drink, and the number of standard drinks it contained. At all but the first drink, participants were asked the end time of the prior drink and the running total of completed drinks that day. Following each drink report, participants were reminded to log any additional drinks consumed. (2) *Event-contingent done drinking reports* were available for participants to report the time they finished drinking for the day or night, and the total number of drinks they consumed. If a participant completed a done drinking report but then continued drinking, they would complete an additional drink start survey and then another done drinking report when they were truly finished drinking. (3) *Event-contingent environmental alcohol reports* measured the time that hand-sanitizer or another alcohol-based product came in contact with the participant's skin. (4) *Morning reports* measured alcohol use from the previous day to provide additional detail, including drinks not reported in the event-contingent surveys. Participants reported the total number of standard drinks, and the time of the first and

last sip of alcohol the prior day. Participants were asked how many times they removed the Skyn, and if they had removed the Skyn, the start time, end time, and reason for each removal (bathing/showering, charging, other). Morning reports also reminded participants to charge the Skyn and sync it to the app daily, and the app included an "information" page with reminders (Fig. 2).

Figure 2. PILR app info page

2.2.4. Remote Exit Interview. After the week of field-based data collection, the bracelet was collected and participants attended a Zoom interview session. The participants were guided in how to uninstall the apps and we reviewed the data provided by the daily self-report and the Skyn sensor with the participant. We identified and investigated days when self-report and/or Skyn data were missing. We anticipated some missing data because the bracelet must be turned off to be charged.

However, participants were asked about times when data were missing and they had not reported that they had removed the bracelet, or if the length of time the data were missing was more than 3 hours. This information was critical to determine whether missing data were due to technology failure or participant noncompliance. A qualitative interview (approximately 30-45 mins) following a semi-structured agenda focused on sensor user experience, feasibility, and acceptability (Aim 1). We asked about the physical comfort of the device and psychological comfort with being monitored. We queried about barriers to compliance (encountered and perceived), suggestions for protocol improvements, and adequacy of the study compensation and length, to inform future research protocols. We also assessed perceived reactivity by asking participants whether they thought their behavior changed due to the monitoring. Sessions were recorded and captured in interviewer notes, and debrief summaries were discussed among the research team. Participants were compensated $25 for attending the baseline/orientation session, $35 for the week of field data collection ($5 per day), a $15 bonus for 100% compliance with morning reports, and $25 for the follow-up session. Total possible compensation was $100, provided via electronic gift cards to Amazon.com.

2.3. Baseline and Follow-up Measures.

2.3.1. Demographics. Demographic information (sex, race, ethnicity) was collected at baseline.
2.3.2. Alcohol Use. Past 30-day alcohol use was measured with 5 items to assess: (a) number of drinking days, (b) number of heavy drinking days (4+/5+ drinks for women/men), (c) typical drinks per drinking day, (d) maximum drinks on a drinking day, and (3) times drunk. Participants also completed the Alcohol Use Disorder Identification Test (AUDIT), a 10-item self-report assessment of past year drinking behavior [17].
2.3.3. System Usability Scale (SUS). At follow-up, we administered the SUS, a 10-item measure of the ease of and interest in using a product/system [18]. We modified the scale to refer to the BACtrack Skyn; item examples included "I thought the bracelet was easy to use" and "I felt very confident using the bracelet." Item 6 ("I thought there was too much inconsistency in this system") was not included because it was not applicable to the bracelet. Each item is rated from 1 (strongly disagree) to 5 (strongly agree). Raw SUS ratings were converted to a sum score ranging from 0-100 with an adjustment for the removal of one item [19]. The SUS had a high internal consistency in this sample (α = .84).

2.4. Data Analytic Plan

To address **Aim 1** (feasibility and acceptability of the Skyn), we descriptively examined (a) willingness to participate, using 80% of eligible participants enrolling as a benchmark; (b) retention, using 90% of those enrolled completing participation as a benchmark; (c) SUS scores; and (d) qualitative interview responses. We documented reasons for not participating, attrition, and noncompliance. To address **Aim 2** (device functionality and missing data), we documented any apparent problems with Skyn readings and examined (a) total number of missing blocks of data when the Skyn was worn in the field, (b) time length of missing data in these blocks, (c) number of missing blocks of data per person, and (d) reports of reasons for missing data.

3. Results

3.1. Feasibility and Acceptability

Twenty-three (100%) of the participants who were eligible following the remote TLFB consented to participate. One was withdrawn by the researchers due to repeatedly rescheduling the orientation session, one changed their mind about their interest in the study after consenting, and one did not provide a reason. Twenty (87%) were fully enrolled and participated (Table 1), achieving our first feasibility benchmark, and 100% of those enrolled completed the study, achieving our second benchmark.
3.1.1. System Usability. A curved grading scale (CGS) can be used to interpret SUS scores [20]. Table 2 describes the breakdown of the Sauro-Lewis CGS and the frequency of those grades in this sample. The bracelet's SUS score (M = 72.5, SD = 15.7) can be translated to a C+ grade, indicating "good" perceived usability [18]. Three participants (two women, one man) scored in the D/F range. We reviewed these cases and indeed they did reflect the more challenging issues we had with the Skyn connectivity.

3.2. Qualitative Interview Data.

3.2.1. Positive feedback on Skyn. Of the 20 participants, 10 expressed at least one specific positive sentiment about the Skyn (e.g., easy to put on and take off, light and comfortable, easy to wear continuously) or the Skyn app (ease of logging into the app with the thumbprint iOS security feature). Five participants liked the monitoring aspect of the bracelet and the ability to see their data in real-time, despite having been told during orientation that the readings were not meaningful for them. Participants reported liking aspects of the

overall protocol, including developing more awareness of what and when they were drinking (n=7) and the ease of participation (n=5). Five participants thought the overall experience was "cool" and/or interesting, and almost all (n=19) indicated they would recommend a similar study to a friend.

3.2.2. Negative feedback on Skyn. There were several notable dislikes regarding the Skyn bracelet and app. The most frequent (n=14) surrounded the comfort of the bracelet and an inability to keep it tight on one's wrist. Participants indicated that the bracelet often came loose, particularly while sleeping for some. One noted "You told me to wear it tight...it loosens quickly, and I have to constantly tighten it up every two seconds."

Several participants noted difficulty with protocol requirements related to the Skyn. In particular, some found the charging protocol difficult (n=11), including remembering to charge the bracelet daily, to turn it back on after charging, and/or to put it back on their wrist after charging. Some also disliked that the bracelet was not waterproof (n=6) and some disliked the appearance of the bracelet, including the Skyn logo (n=6). Only three participants described difficulty with the app functionality, including problems syncing, app crashes, and getting logged out. Only two participants indicated difficulty with the requirement to sync the bracelet to the app daily. Two reported difficulty wearing the bracelet continuously, particularly during physical activities (e.g., tennis), and one found it problematic to avoid alcohol-based lotions/perfumes.

Table 1. Sample descriptives (n=20)

	n (%)
Female Sex[a]	14 (70%)
Hispanic/Latino ethnicity	3 (15%)
Race (check all that apply)	
White	15 (75%)
Asian	6 (30%)
Black/African American	1 (5%)
American Indian/Native Alaskan	1 (5%)
	M (SD), range
Age	24.30 (6.54), range 21 – 49
Drinking days in past 30	13.80 (4.65), range 7 – 25
Drinks per typical drinking day (past 30 days)	3.85 (2.08), range 1 – 10
Maximum drinks in a day (past 30 days)	9.15 (2.72), range 6 – 15
Times drunk in past 30	6.60 (4.83), range 1 – 20
Number of heavy drinking days in past 30 (4+women/5+ men)	7.20 (3.21), range 3 – 15
AUDIT total score	11.25 (5.33), range 4 – 25

Note. [a]One individual of male sex reported non-binary/genderqueer/gender non-conforming for gender identity; in all other cases gender identity was the same as sex, [b] 30% scored <8, 60% scored 8-15, and 10% scored 20+ on AUDIT.

Table 2. System Usability Scale scores frequency using the Sauro-Lewis CGS

	Sauro-Lewis CGS		Frequency	
Grade	Score Range	Percentile	N	%
A+	84.1–100	96–100	4	20
A	80.8–84.0	90–95	3	15
A−	78.9–80.7	85–89	1	5
B+	77.2–78.8	80–84	0	0
B	74.1–77.1	70–79	2	10
B−	72.6–74.0	65–69	0	0
C+	71.1–72.5	60–64	1	5
C	65.0–71.0	41–59	5	25
C−	62.7–64.9	35–40	1	5
D	51.7–62.6	15–34	2	10
F	0.0–51.6	0–14	1	5

3.2.3. Concerns about Skyn prior to participating. Some participants indicated that, prior to participation, they had some concerns about using the Skyn. One concern was that others would notice the bracelet and ask questions about it, requiring them to disclose participation in a research study (n=2). Others were concerned about following instructions (n=3); losing the bracelet (n=1); that it would be bulky, inconvenient, or uncomfortable (n=3); and about data privacy (n=1).

3.2.4. Participant suggestions. Participants had several suggestions to address some of the reported concerns above, and more generally how to prepare participants for similar research. These included giving participants ideas about how to respond if someone asks about the bracelet in daily life (n=2), changing the Skyn label (n=1), providing informative videos on how and when to charge the bracelet (n=1), and making participants aware that the bracelet often becomes loose (n=1). Additional useful suggestions from participants were to have more communication from the research team, such as a confirmation on the first day of data collection that data had been received and a mid-week check-in (n=2), more incentives throughout the study (n=1), reminder notifications for key tasks related to the Skyn (e.g., turn on after charging, n=1), and distributing training materials for future reference (n=1). Finally, three participants explicitly suggested a different type of band (e.g., canvas) would be preferable.

3.2.5. Feedback on compensation. The majority of participants thought the compensation they received for a week of wearing the Skyn and self-reporting their drinking behavior was acceptable to good (n=17). Two participants found the compensation acceptable, but thought it should be increased to reflect the amount of effort. Only one participant was unhappy, and this was specific to their failure to earn the payment bonus due to what the participant perceived to be app difficulties.

3.2.6. Feedback on study length. Participants differed in how long they would be willing to participate in a study that required both wearing the Skyn and self-reporting their drinking behavior. Three reported that they would not want to participate for more than one week, seven indicated they would participate for up to 2 weeks, two for a few weeks, five for about a month, one for a semester, and two "indefinitely."

3.2.7. Perception of assessment reactivity. Fourteen participants indicated that they did not change their drinking in any way because it was being monitored. Two participants thought they drank more, and this was motivated by a desire to provide the study with data. Two participants thought they drank less; one did not want to have to log drinks consumed and the other became more mindful of their drinking during the study.

3.3. Device and App Functionality

We encountered a technical issue with one bracelet provided to a participant, in which baseline TAC readings (i.e., when no alcohol was consumed) consistently fell below zero. We discontinued use of this participant's bracelet at the end of their data collection. As noted above, only three participants described difficulty with the app functionality (e.g., problems syncing, app crashes).

3.4. Missing Data

Skyn readings are expected to be taken every 20 seconds, as long as the bracelet is turned on (whether or not it is being worn). Notably, participants were instructed to remove the bracelet when it might get wet (e.g., showering, swimming). Additionally, we required that participants charge the bracelet daily, and turn the bracelet off at that time. As such, we would expect approximately 7 blocks of missing data per person.

To characterize missing data, we looked specifically at the time span beginning at 9pm the day participants were oriented through 9pm the day before they returned their bracelet. Across participants, the average portion of this week during which we received Skyn data was 85% (range 46%-96%), with data for at least 90% of this time from 11 of 20 participants. Across the total 3360 hours of data collection (n=24 hours x 7 days x 20 participants), there were 206 blocks of time greater than 20 seconds where a reading was not transmitted. Among these, the median was 17 minutes. However, 88 of those blocks (42.7%) were less than one minute, and another 5 (2.4%) were between 1 and 5 minutes; these were not deemed to represent meaningful missingness, given the short duration and availability of proximal readings. The distribution of the remaining missing time blocks is shown in Figure 3. As can be seen, the modal amount of time with missing data was between 1 and 2 hours (32 blocks), and very few (5 blocks) were missing for 12+ hours. The average number of blocks of missing data (5+ minutes) per person was 5.7 (SD=2.4), with a range from 1-10.

Next, we examined the extent to which the 113 missing blocks of data (>5 mins) corresponded to self-reports of bracelet removal (including times when the bracelet was still powered on/reading even though not on wrist). The time of the missing data (when the bracelet was powered off/not reading) was matched with the time of self-reported bracelet removal to determine the participant's reason. See Figures 4 and 5 respectively for a graphical view of a short block of missing TAC data during a non-drinking period, and complete TAC data during a drinking event.
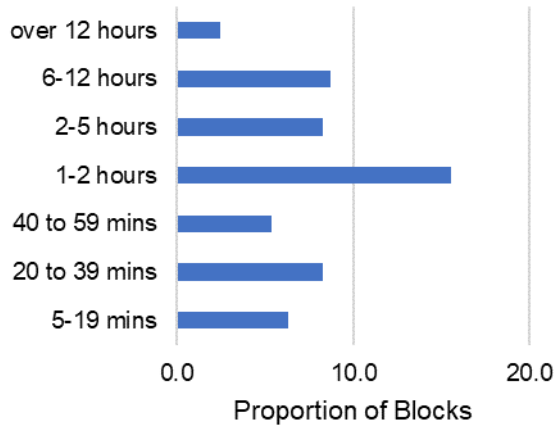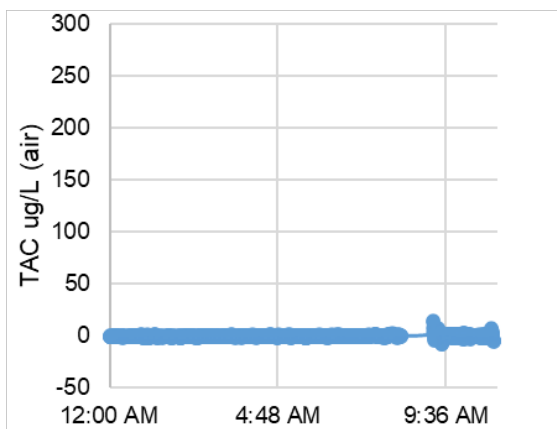
Figure 3. Blocks of missing Skyn data (>5 mins)

Figure 4. Example of TAC data with a missing time block during a non-drinking
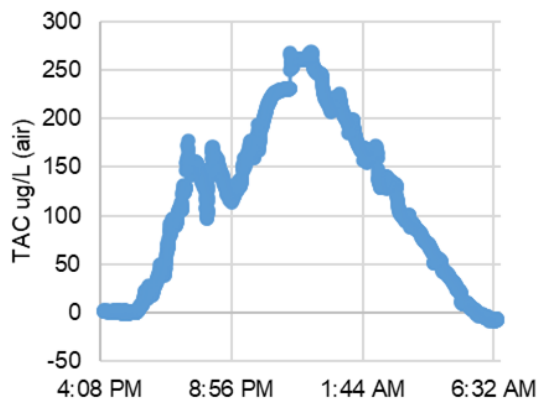

Figure 5. Example of TAC data with no missing time blocks during a drinking event

In total, 81% of the missing data blocks were accompanied by a self-report of bracelet removal. Of the 113 missing blocks, most were due to charging, and 19% did not overlap with a self-report of bracelet removal (Figure 6). The average span of missing data blocks during self-reported charging was 2 hours 24 mins (SD = 3 hours 50 mins), and the average span of missing data blocks during bathing was 1 hour 36 mins (SD = 2 hours 30 mins). Blocks of missing data that were not explained by participants were the longest on average (13 hours 35 mins, SD = 14 hours 7 mins).
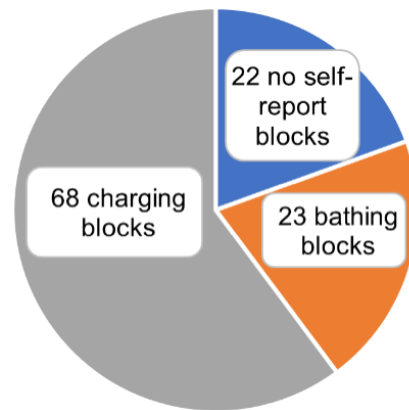

Figure 6. Categorization of missing Skyn data (≥ 5 mins)

Finally, we examined *all* self-reported reasons for bracelet removal (regardless of whether data were missing, such as in the case that the bracelet was removed but still powered on/reading), including the morning report from day 8 (corresponding to day 7 of Skyn data collection). Here, there were a total of 149 responses. Of these, 82 (55%) were reported as due to charging, 54 (36%) due to bathing/showering, and 13 (9%) due to "other" (e.g., washing dishes, tennis). Two of the "other" responses indicated problems with the bracelet (i.e., "fell off during sleep" and "bracelet having issues with charging"). Although there were 149 self-reported bracelet removals, only 91 (61%) of the missing data blocks (5+ minutes) were explained via self-report. This indicates that the remaining 58 self-reported bracelet removals corresponded either to missing data blocks less than 5 minutes, or to times when the bracelet was off one's wrist but still powered on. Of note, self-reported removals for "other" reasons did not match times with any of the blocks of missing Skyn data, supporting the latter explanation.

Notably, two participants had particularly large amounts of missing data (61 and 90, out of 168 possible hours), most of which occurred at the end of their data collection period. These participants reported minor connectivity issues when syncing, so we asked them to

reset their Skyn app. However, these participants were noncompliant with this request and unresponsive to communication. It was determined via self-report that that these participants were also noncompliant with our instructions to charge and sync the bracelet daily.

4. Discussion

The present study is among the first to examine feasibility and acceptability of the use of the BACtrack Skyn alcohol sensor bracelet in daily life in a sample of heavy drinkers. Overall, results indicate promise in our ability to integrate this tool in research. With respect to feasibility, close to 90% of eligible participants enrolled, and 100% of those who enrolled completed the study. System usability of the bracelet was "good," and qualitative data generally indicated that the bracelet was acceptable. Regarding device functionality, missing data were relatively minimal. In most cases, blocks of missing data were either very short (less than 5 minutes) or expected based on instructions that participants should remove and shut off the bracelet daily for charging. Furthermore, there were no times at which we could verify that missing data was due to malfunction of the bracelet, and we also had only one time when we could verify the bracelet was not functioning as expected (below zero baseline TAC).

The most common complaint regarding the Skyn was that it was uncomfortable and difficult to keep tight around one's wrist (a requirement for data collection). This input has been provided to BACtrack, in hopes that different bands will be available for future Skyn models. Other things that will enhance user experience and/or acceptability would include removal of the Skyn logo, and a version that is waterproof. Participants also had some trouble with the charging protocol, which required turning the bracelet off, and remembering to charge in the absence of a battery life indicator. The Skyn's usability would be enhanced for future work by a low battery indicator within the app or on the bracelet itself. Despite these noted limitations, most participants thought the bracelet and associated app were relatively easy to use. Our findings are somewhat contrary to those reported by others. Specifically, Fairbairn et al. [13] noted significant failure rates, whereas we only had one bracelet that showed data that was out of the expected range. It is possible that some of our missing data was due to bracelet failure; since we were conducting field research we cannot confirm whether the bracelet was turned on and charged when the data were missing (this would indicate that the bracelet should be collecting data). However, there was never a time when a bracelet clearly stopped collecting data. With the exception of the two participants described earlier who stopped following the study protocol toward the end of the week,

most missing data blocks were followed by collected data, indicating that the data were missing due to battery depletion, not bracelet failure. It is likely that the newer bracelet models are more stable in their functionality.

A second finding that appears contrary to prior work is that we collected more negative feedback from our participants about the bracelet and app than the other study that reported on acceptability of the Skyn. Wang et al. [12] reported on a study with 12 participants who wore the Skyn for 2 weeks, and reported highly positive reactions from participants. Though the same measures were not used by our study, our impressions are that our participants had a less positive experience, as more of our participants had negative feedback than positive. However, 19 of 20 of our participants indicated they would recommend the study to others, and most said they would participate longer, indicating participation was not highly aversive. Additionally, participants' negative feedback may have been influenced in part by the additional burden of repeated daily self-reports in the current study. We also would note that some of the negative feedback from participants was similar across the studies, including that some participants had difficulty syncing their bracelets with the app, could not tell when the bracelet needed to be charged, and did not like that it was not waterproof.

4.1. Recommendations: Skyn Use in Research

Our experience in this pilot, and much of the qualitative data collected from our participants, allow us to make several recommendations for other researchers interested in using the Skyn. First, we conducted a comprehensive orientation session that included providing participants with detailed instructions on how to use the Skyn within the study. In spite of thorough training with slides and Zoom-based demonstrations, we learned that some participants required follow-up contact to ensure they understood instructions completely. It is possible that additional contact was required because we had conducted our orientations on Zoom; having no in-person contact made it difficult to tell whether participants were understanding and following instructions. Indeed, in our follow-up interviews, some participants expressed a desire for more communication from the research team, reminder notifications to charge, and training materials for future reference. Therefore, while the Zoom-based orientation allowed us to complete the research during COVID restrictions, and can be used effectively and efficiently in future studies, we still recommend adjustments to the orientation and to initial follow ups to ensure participants are following instructions (e.g., more frequent check-ins by phone or text during the data collection period). An additional recommendation is that researchers instruct their participants to charge the

bracelet while bathing, as to minimize the number of blocks of missing data per day.

As is typical for event sampling studies that commonly use EMA methods, we compensated participants for wearing the bracelet and for completing EMA morning reports. We also provided a compliance bonus for completing all of the morning reports. However, we did not have a method for measuring compliance with wearing the bracelet, and we believe that having some compensation contingent on complying with expectations to charge and sync the bracelet daily would result in less missing data. This would require establishing some guidelines, including determining confidently how long it takes to charge a bracelet and indications that a bracelet is being worn, not just powered on and collecting data (i.e., using temperature readings, see future directions below).

4.2. Limitations

There are several limitations to acknowledge. First, our sample size was quite small, though characteristic of this type of pilot work and of qualitative investigations [21, 22]. Larger, more diverse samples should be used to understand participant characteristics associated with Skyn acceptability. Second, our eligibility criteria, including heavy drinking status, and the characteristics of our sample (e.g., majority female) restrict the ability to generalize our findings to other groups. Only participants with iOS devices were recruited due to required device compatibility with the Skyn app. Generalizability will be improved if future studies provide an iOS device for participants who do not own one. Third, while we had hoped to collect Skyn data in the laboratory using a controlled alcohol administration paradigm to assess correspondence between TAC and BrAC readings, in-person research was thwarted by the COVID-19 pandemic as we prepared to launch the present study. Future work pairing use of the Skyn in both the laboratory and the field will further advance our knowledge of the capabilities of this tool.

4.3. Future Directions

Essential next steps in moving toward an ability to use the Skyn for research and clinical applications are to develop methods for interpreting the data. Currently, the data provided by the sensors is best described as a "raw signal", with no interpretability nor meaningful scale. When presented in graphical form, the data forms a curve that reflects transdermal alcohol (Figure 5); the alcohol curve goes up and comes down but the *values* of the data are not interpretable (i.e., the raw signal does not have an interpretable scale similar to BAC), nor do we know the nature of the scale (i.e., whether it is linear

or not). In addition, the device manufacturer does not indicate to the user when the Skyn data reflects alcohol consumption (and when it does not). In other words, there are no defined *criteria* to use to evaluate the sensor data in order to determine whether alcohol was consumed. For example, the SCRAM system provides a determination as to whether alcohol was consumed, and these criteria (which involve peak TAC, absorption and elimination rates) have been modified successfully for research purposes, obtaining excellent specificity and sensitivity [23]. Without such criteria for the Skyn, the device will be difficult to use for field research when alcohol use is not controlled or directly measured.

Another requisite for using the Skyn for field research is the ability to tell when the Skyn has been removed by the wearer. The temperature sensor on the Skyn can be used for this purpose; when the Skyn is removed, the temperature drops. However, there are not yet temperature criteria established and tested that will confidently allow a researcher to determine when the bracelet is not being worn. This is a different problem than that of missing data described above; when the bracelet is charged and turned on, data are produced, but if the bracelet is not being worn, these data are invalid. It is critical that we be able to identify data that are likely not reflecting transdermal alcohol levels.

Transdermal bracelets detect alcohol in perspiration on the skin's surface, but we know of no research with sensors that has investigated the effect of perspiration (e.g., from exercise) on the transdermal signal. It is unclear for example, if a high degree of sweating would produce similar readings to low sweat circumstances. Laboratory-based work could manipulate sweating, and field-based work could measure it carefully through self-report or another sensor.

If validation work continues to support the use of the Skyn, it may have useful clinical implications as well. Of note, participants in our study were interested in viewing their TAC data through the mobile application. While the ability to see these data are not ideal for observational research (where the goal is to observe natural behavior), this is a feature that could be leveraged in other types of research and/or intervention. For example, once TAC-BrAC translation is possible, participants could self-monitor times when they may be approaching a risky and/or undesirable level of alcohol intoxication. Future treatment development research might also benefit from using TAC readings as a way to provide just-in-time interventions [24], such as reminders of protective behavioral strategies one could implement when TAC is first detected, or alerts when TAC is rising too quickly and/or to a dangerous level.

4.4. Conclusion

The Skyn may be an extremely valuable tool to monitor alcohol use in participants and/or patients with minimal burden. Adjustments in wearer orientation to the bracelet and app, and improvements in bracelet design and data interpretation will enhance the potential of investigations using this promising device.

Devices were leased from BACtrack, and the researchers have no conflict of interest to report.

5. References

[1] R. Miranda, P.M. Monti, L. Ray, H.R. Treloar, E.K. Reynolds, J. Ramirez, J., T. Chun, C.J. Gwaltney, A., Justus, J. Tidey, A. Blanchard, and M. Magill, "Characterizing Subjective Responses To Alcohol Among Adolescent Problem Drinkers", J Abnorm Psychol, 2014, pp. 117-29.

[2] R. Miranda, J. MacKillop, J., H. Treloar, A. Blanchard, J. Tidey, R.M. Swift, T. Chun, D.J. Rohsenow, and P.M. Monti, "Biobehavioral Mechanisms Of Topiramate's Effects On Alcohol Use: An Investigation Pairing Laboratory And Ecological Momentary Assessments", Addict Biol, 2016, pp. 171-82.

[3] G. Huntley, H. Treloar, A. Blanchard, P.M. Monti, K.B. Carey, D.J. Rohsenow, and R. Miranda, "An Event-Level Investigation Of Hangovers' Relationship To Age And Drinking", Exp Clin Psychopharmacol, 2015, pp. 314-323.

[4] J. Northcote and M. Livingston, "Accuracy Of Self-Reported Drinking: Observational Verification Of 'Last Occasion' Drink Estimates Of Young Adults", Alcohol Alcohol, 2011, pp. 709-13.

[5] D.M. Donovan, G.E. Bigelow, G.S Brigham, K.M. Carroll, A.J. Cohen, J.G. Gardin, J.A. Hamilton, M.A. Huestis, J.R Hughes, R. Lindblad, G.A. Marlatt, K.L. Preston, J.A. Selzer, E.C. Somoza, P.G. Wakim and E.A. Wells, "Primary Outcome Indices In Illicit Drug Dependence Treatment Research: Systematic Approach To Selection And Measurement Of Drug Use End-Points In Clinical Trials", Addiction, 2012, pp. 694-708.

[6] R. Swift, "Direct Measurement Of Alcohol And Its Metabolites", Addiction, 2003, pp. 73-80.

[7] R. Swift., "Transdermal Alcohol Measurement For Estimation Of Blood Alcohol Concentration", Alcohol Clin Exp Res, 2000, pp. 422-3.

[8] N.P. Barnett, E.B. Meade, and T.R. Glynn, "Predictors Of Detection Of Alcohol Use Episodes Using A Transdermal Alcohol Sensor", Exp Clin Psychopharmacol, 2014, pp. 86-96.

[9] D.M. Dougherty, D.M., N. Hill-Kapturczak, Y. Liang, T.E. Karns, S.L. Lake, S.E. Cates, and J.D. Roache, J, "The Potential Clinical Utility of Transdermal Alcohol Monitoring Data to Estimate the Number of Alcoholic Drinks Consumed", Addict Disord Their Treat, 2015, pp. 124-130.

[10] N. Hill-Kapturczak, N., S.L. Lake, J.D. Roache, S.E. Cates, Y. Liang and D.M. Dougherty, "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, 2014, pp. 2517-2522.

[11] C.E. Fairbairn and D. Kang, "Temporal Dynamics of Transdermal Alcohol Concentration Measured via New-Generation Wrist-Worn Biosensor", Alcohol Clin Exp Res, 2019, pp. 2060-2069.

[12] Y. Wang, Y., D.J. Fridberg, D.D. Shortell, R.F. Leeman, N.P. Barnett, R.L. Cook and E.C. Porges, "Wrist-Worn Alcohol Biosensors: Applications And Usability In Behavioral Research", Alcohol, 2021, pp. 25-34.

[13] C.E. Fairbairn, D. Kang, and N. Bosch, "Using Machine Learning For Real-Time BAC Estimation From A New-Generation Transdermal Biosensor In The Laboratory", Drug Alcohol Depend, 2020, pp.108205.

[14] C.E. Fairbairn and D. Kang, "Temporal Dynamics of Transdermal Alcohol Concentration Measured via New-Generation Wrist-Worn Biosensor", Alcohol Clin Exp Res, 2019, pp. 2060-2069.

[15] B. Pittman, R. Gueorguieva, E. Krupitsky, A.A. Rudenko, B.A. Flannery, and J.H. Krystal, "Multidimensionality of the Alcohol Withdrawal Symptom Checklist: a factor analysis of the Alcohol Withdrawal Symptom Checklist and CIWA-Ar", Alcohol Clin Exp Res, 2007, pp. 612-8.

[16] L.C. Sobell, J.A. Cunninghamm, and M.B. Sobell, "Recovery From Alcohol Problems With And Without Treatment: Prevalence In Two Population Surveys", American Journal of Public Health, 1996, pp. 966-972.

[17] J.B. Saunders, O.G. Aasland, T.F. Babor, J.R. de la Fuente, and M. Grant, "Development Of The Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project On Early Detection Of Persons With Harmful Alcohol Consumption: II", Addiction, 1993, pp. 791-804.

[18] A. Bangor, A., P.T. Kortum, and J.T. Miller, "An Empirical Evaluation of the System Usability Scale", International Journal of Human–Computer Interaction, 2008, pp. 574-594.

[19] J.R. Lewis and J. Sauro, "Can I Leave This One Out? The Effect Of Dropping An Item From The SUS", J. Usability Studies, 2017, pp. 38–46.

[20] J. Sauro and J.R.. Lewis, Quantifying The User Experience: Practical Statistics For User Research, 2nd ed, Morgan Kauffmann, Cambridge, MA, 2016.

[21] B. Marshall, P. Cardon, A. Poddar and R. Fontenot, "Does Sample Size Matter in Qualitative Research?: A Review of Qualitative Interviews in is Research", Journal of Computer Information Systems, 2013, pp. 11-22.

[22] S.L. Dworkin, "Sample Size Policy for Qualitative Studies Using In-Depth Interviews", Archives of Sexual Behavior, 2012. pp. 1319-1320.

[23] N.P. Barnett, J. Tidey, J.G. Murphy, R. Swift, and S.M. Colby, "Contingency Management For Alcohol Use Reduction: A Pilot Study Using A Transdermal Alcohol Sensor", Drug and alcohol dependence, 2011, pp. 391-399.

[24] I. Nahum-Shani, S.N. Smith, B.J. Spring, L.M. Collins, K. Witkiewitz, A. Tewari, and S.A. Murphy, "Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support", Ann Behav Med, 2018, pp. 446-462.