

Accepted Manuscript

Mobile Alcohol Biosensors and Pharmacotherapy Development Research

Walter Roberts, PhD, Sherry A. McKee, PhD

PII: S0741-8329(18)30058-2

DOI: [10.1016/j.alcohol.2018.07.012](https://doi.org/10.1016/j.alcohol.2018.07.012)

Reference: ALC 6832

To appear in: *Alcohol*

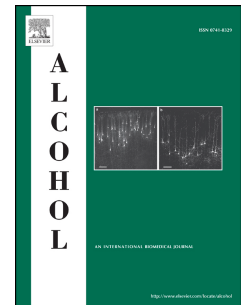
Received Date: 5 March 2018

Revised Date: 27 July 2018

Accepted Date: 30 July 2018

Please cite this article as: Roberts W. & McKee S.A., Mobile Alcohol Biosensors and Pharmacotherapy Development Research, *Alcohol* (2018), doi: 10.1016/j.alcohol.2018.07.012.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Word Count: 7,196

Tables: 3

Figures: 4

Mobile Alcohol Biosensors and Pharmacotherapy Development Research

Walter Roberts, PhD^{1,2}, Sherry A. McKee, PhD¹

¹Yale School of Medicine, Department of Psychiatry, New Haven CT

²Corresponding author: Telephone: (203) 737-3529, Fax (203) 737-4243, address: 2 Church Street South Suite 109, Yale University School of Medicine, New Haven, CT 06519, e-mail: walter.roberts@yale.edu

Funding: Supported by NIH grants T32 DA007238 (WR), R43AA026492 (SAM subcontract PI), R01AA022285 (SAM).

Disclosures: SM has ownership interest in Lumme, Inc. WR has no conflict to report.

Abstract

Recent advances in biosensor technology herald a major shift in how alcohol use will be tracked in humans. Wearable biosensors can passively and continuously monitor wearers' alcohol consumption in real time. An important application of these biosensors is to improve how medication for alcohol use disorder (AUD) is tested in clinical research. Both laboratory-based screening paradigms and clinical trials have methodological problems that impact their efficiency and predictive validity. Medication screening using laboratory-based methods is a resource intensive assessment of a single episode of behavior in a non-representative setting. Clinical trials rely on participant self-report to document medication-induced changes in drinking behavior. This review describes how mobile biosensors can be leveraged to improve AUD medication development research. We first review the current state of alcohol biosensor technology with a focus on strengths and limitations of the devices. We describe how multiple biosensors can be combined to create a far more detailed record of drinking compared to single biosensor platforms. We then discuss each phase of the medication development pipeline in turn (i.e., Phases 1 – 4) and describe how mobile biosensors can be incorporated in standard medication testing paradigms to improve efficiency and predictive validity. We conclude with discussion of challenges associated with using currently available biosensors for medication testing and recommendations for researchers wishing to incorporate alcohol biosensors into their own research.

Keywords: mobile biosensors; medication development; alcohol use disorder; transdermal alcohol sensors; pharmacotherapy

Introduction

Harmful alcohol use is a major public health concern. Alcohol contributes to approximately 88,000 deaths annually within the United States (Centers for Disease Control, 2013) and the annual economic burden of alcohol is estimated at \$249 billion (Sacks, Gonzales, Bouchery, Tomedi, & Brewer, 2015). In response to this major source of disease, considerable resources have been committed to developing treatments to help people reduce their alcohol consumption. Pharmacotherapies are a central component of alcohol use disorder (AUD) treatment. Currently, three medications are approved by the Food and Drug Administration (FDA) to treat AUD (naltrexone [oral and injectable], disulfiram, and acamprosate). Although these medications are effective for some patients, many patients treated with these medications will ultimately return to heavy drinking (Jonas et al., 2014). The National Institute of Alcohol Abuse and Alcoholism (NIAAA) has prioritized developing novel pharmacotherapies for AUD (Litten, Falk, Ryan, & Fertig, 2016). Identifying efficacious new medications will improve treatment outcomes, particularly in groups of patients for whom currently available medications are less effective (Garbutt et al., 2014).

Developing medications to treat addiction is technically challenging and expensive. Despite the substantial resources committed to AUD medication development research, the field has been slow to deliver effective pharmacotherapies (Litten, et al., 2016). There are many barriers to AUD medication development research, such as high participant attrition (Hallgren & Witkiewitz, 2013), lack of clarity regarding ideal laboratory screening models and clinical endpoints (Falk et al., 2014), and small treatment effect sizes (Jonas et al., 2014). These barriers contribute to the poor efficacy of promising pharmacotherapies when they are tested in humans (Litten, Wilford, Falk, Ryan, & Fertig, 2016). NIAAA has acknowledged these issues stymieing medication development and developed a strategic plan for improving the efficiency and predictive validity of this research. Two major components of this plan are: 1) identify translational paradigms to screen candidate medications for preliminary efficacy and safety in

humans, and 2) improve the methodological quality of pivotal phase 3 clinical trials (Litten et al., 2012; Litten et al., 2016a).

One emerging technology with potential to improve AUD medication development research is mobile health biosensors. Mobile biosensors are wearable devices that can passively detect fluctuations in a person's biological state, such as changes in blood pressure, body position, or body fluid composition. These devices are already being used clinically for remote health monitoring in other medical disciplines (Bruen, Delaney, Florea, & Diamond, 2017; Kumar et al., 2013; Steinhubl, Muse, & Topol, 2013), and there is growing interest in using mobile biosensors to track outcomes in clinical trials (Herrington, Goldsack, & Landray, 2018). Indeed, mobile biosensors can continuously monitor biological signals for symptom tracking in the management of chronic disease, such as seizure frequency in people with epilepsy, cardiac arrhythmias in people with cardiovascular pathology, and electrodermal activity in illicit drug users (Carreiro et al., 2015; Fensli, Gunnarson, & Gundersen, 2005; Poh et al., 2012; see Pantelopoulos & Bourbakis, 2010 for a review of mobile biosensor in healthcare). These devices have gained popularity due to their ability to passively and objectively monitor clinically important biological processes in real time.

We believe that biosensors can be integrated into AUD medication research to further NIAAA's strategic plan (Litten, Falk, et al., 2016). In this review, we argue that mobile biosensors have major potential to improve the quality and efficiency of AUD medication development research. First, we provide a review of recent developments in alcohol biosensor technology and describe the data collection capabilities of these devices. We describe our efforts to improve on these devices by combining multiple biosensors into a single mobile platform. Second, we review the goals and methodologies used at each stage of the medication development pipeline and describe how standard paradigms can be enhanced (or even replaced) using mobile biosensors. We conclude the review by describing some of the difficulties

associated with using mobile biosensors in medication development research and making suggestions for managing these challenges

Using Mobile Biosensors to Detecting Alcohol Consumption

Many measurable biosignals are produced during or after drinking as alcohol is consumed, metabolized, and excreted. Biosensors are devices that can detect these signals. Although alcohol consumption results in many biosignals that can be detected with appropriate sensors, most of the recent research in alcohol biosensor technology has focused on transdermal alcohol detection. Approximately 1% of ingested alcohol is excreted in perspiration, and the concentration of alcohol in sweat bears an approximately linear association with blood alcohol concentration (BAC; Brown, 1985). Transdermal alcohol sensors (TAS) can continuously monitor sweat for alcohol content using a similar procedure as fuel cell breathalyzers. A sensor is held against the skin surface and alcohol is oxidized; the oxidation current is recorded and used as a direct measure of alcohol concentration (Swift, 2003). Samples can be taken continuously at pre-specified intervals (e.g., every 10 seconds).

Two models of TAS are currently available to researchers. The Secure Continuous Remote Alcohol Monitor (SCRAM; Alcohol Monitoring Systems) is an ankle-worn TAS, and the Wrist Transdermal Alcohol Sensor (WristAS; Giner, Inc) is a similar device that is worn on the wrist (see Marques & McKnight, 2007 for a complete description of these devices). Data are temporarily stored onboard the device and transferred wirelessly via a specialized modem or through a wired connection to a user's computer. Both devices are equipped with temperature and skin resistance/conductance sensors that can detect when the device is removed. The SCRAM includes a tamper-proof strap so that the device cannot easily be removed. Both devices have been used successfully in laboratory-based research (Leffingwell et al., 2013). The SCRAM has been successfully used in clinical trials of contingency management for heavy drinking (Barnett et al., 2017; Dougherty et al., 2014) and in forensic settings to monitor DUI offenders (Flango & Cheesman, 2009; Mathias et al., 2018).

Both the SCRAM and WristAS measure the wearer's transdermal alcohol concentration (TAC). TAC readings are taken continuously and stored on the device or immediately transmitted to a connected device for storage. Measurements are stored in a database with a timestamp, TAC, skin temperature, and, in models that include the necessary sensors, skin conductance and resistance. Because readings are taken continuously, TAS can take many readings each day (e.g., 86,000+), producing a large amount of data. These raw observations are combined and transformed according to the needs of the study protocol. Researchers may be interested in, for example, number of days abstinent (Barnett et al., 2017) or distinguishing moderate from heavy drinking episodes (Dougherty et al., 2012). Pharmacokinetic variables (i.e., peak BAC, rate of BAC rise) can be inferred by charting TAC over the course of a drinking episode, which can be used to estimate BAC (Dumett et al., 2008).

The "next generation" of TAS are currently being developed. These devices resemble other commercially available health monitoring "smart" devices (e.g., smartwatches, wrist-worn activity trackers). One such device is the BACTrack Skyn (BACTrack, Inc.), a wrist-worn TAS that is currently in development but nearing completion. The BACTrack Skyn will likely be the first next generation TAS released to the public (see Figure 1). We have experience with the BACTrack Skyn (other companies are also developing new TAS), so we describe the next generation of TAS with reference to this device as an example. These next generation devices will be smaller and less conspicuous than the SCRAM and WristAS. The Skyn has improved customizability and settings can be altered according to user preference (e.g., the user can specify the frequency with which readings are taken). They also will include improved wireless connectivity and be able to easily transfer data to other devices (e.g., smartphones) using Bluetooth, such that TAS data can be transmitted and displayed in real time. We are not aware of any of new TAS that is being developed with tamper deterrents, although such modifications could be included in later versions. BACTrack, Inc has provided our research group and others with early access to prototype units, and we have successfully conducted laboratory-based

testing using the devices. In our preliminary testing, participants could easily wear and operate the device with minimal physical or social discomfort. For clinical and research populations, such next generation devices will likely be better suited for long-term alcohol use monitoring.

Limitations of TAS

TAC is not a perfect indicator of BAC. Although these two indicators of systemic alcohol levels are moderately correlated under most circumstances, the pharmacokinetics of alcohol are such that excretion of alcohol through the skin occurs more slowly and with less consistency than it is distributed through the blood (Anderson & Hlastala, 2006). Two major discrepancies between BAC and TAC are documented in the literature.

First, there is a well-documented time delay between alcohol consumption and peak alcohol level as measured by TAS, beyond what would be expected due to normal delay in absorption. A range of time delays have been reported with most ranging from 1 to 2 hours (e.g., Sakai et al., 2006), although upward estimates of 4.5 hours delay have been reported (Marques & McKnight, 2007). The nature of this delay is reviewed elsewhere in depth (Karns-Wright et al., 2016), but it is worth noting that the delay is in part physiologically based. Rates of alcohol distribution through the epidermis and stratum corneum are highly variable (Anderson, 2006). Unfortunately, this variability in delay timeframe makes it difficult to translate TAC from BAC, particularly without knowledge of individual drinkers' physiological characteristics (e.g., stratum corneum thickness). This delay period also limits real-time alcohol use detection, which limits several potential applications of TAS that we describe later.

The second systematic limitation is that some TAS devices (e.g., SCRAM) show poor sensitivity for detecting low level drinking episodes. Roache et al. (2015) conducted a study in which participants consumed 1-5 beers (1 beer each 24 minutes) under controlled conditions while wearing a SCRAM. They found that SCRAM only detected 45.9% of low level drinking episodes (i.e., 1-3 beers) when using the manufacturer recommended detection cutoff. However, other studies have found that adapting more liberal detection criteria (i.e., lowering

the TAC cut score for indicating a drinking episode) can improve detection sensitivity (Barnett et al., 2014). Researchers should consider the relative strengths limitations of different devices and detection parameters and select those most appropriate for their needs. Alternatively, as we propose below, additional mobile biosensor devices could be used in tandem with TAS to supplement these limitations and improve overall detection accuracy.

Other Alcohol Mobile Biosensors

Although transdermal alcohol detection has been most frequent strategy for mobile alcohol use monitoring, there are other approaches to continuously monitoring alcohol use via biosensor. NIAAA is currently funding Small Business Innovation Research (SBIR) grants to fund biotech companies with novel strategies to passively monitor for alcohol use. Some devices utilize innovative strategies to improve transdermal alcohol detection (e.g., pilocarpine induced sweating at the sampling site; Kim et al., 2016), whereas other groups are developing devices that monitor other metabolic or elimination pathways to detect consumption (see Campbell, Kim, & Wang, 2018 for a recent review). Many of these devices are in early stages of development: NIAAA is currently funding at least 8 small businesses and research groups working to develop these biosensors (<https://projectreporter.nih.gov/>). These innovative approaches include, for example, sampling interstitial fluid using microneedles (Mohan, Windmiller, Mishra, & Wang, 2017) and colorimetric analysis of saliva using smartphones (Jung et al., 2015). Another approach is to prompt user-generated input when drinking episodes occur (e.g., beverage type and number of drinks reported on smartphones) to adjust TAC-based BAC estimates (Luczak et al., 2018).

Our research group is partnering with Lumme, Inc-- an NIAAA-funded small biotech company—to develop a combined biosensor platform that measures multiple biosignals to improve drinking detection accuracy. This system combines TAS (i.e., BACTrack Skyn) with a commercially available smartwatch loaded with proprietary software developed by Lumme. This software utilizes data from the gyroscope and accelerometer sensors contained in most

commercially available smartwatches. Using a classification algorithm developed via machine learning, the program can detect the distinctive arm movements that occur during a drinking episode (Parate, Chiu, Chadowitz, Ganesan, & Kalogerakis, 2014 used this approach to remotely detect cigarette smoking and eating behavior). Drinking gestures are initially detected by the movement-based sensors and later confirmed as an alcohol drinking episode when transdermal alcohol is detected. An advantage of this approach is that drinking can be detected immediately when it occurs, even before alcohol is distributed into the blood.

Beyond its use in improving alcohol detection, the combined TAS/Lumme system can provide a richer description of a drinking session relative to TAS alone. As the TAS collects detailed data on alcohol pharmacokinetics, the gyroscope/accelerometer concurrently measures drinking topography. Importantly, measures of drinking gestures can capture number of sips and sip length, and combined with knowledge of alcohol content by volume, these variables can provide reasonably accurate estimated blood alcohol levels. Additionally, topographical measures (e.g., number of sips, sip length, inter-sip interval, inter-drink interval, start/end of drinking session) are themselves important and relevant for medication development. In studies of tobacco use disorder, smoking topography often serves as a secondary endpoint during phase 2 testing (Brandon et al., 2011; Roche, Bujarski, Hartwell, Green, & Ray, 2015). Topographical variables are used less frequently in alcohol research, although such topographical measures are associated with self-reported craving and AUD diagnosis (Rankin, Hodgson, & Stockwell, 1979; Sobell, Schaefer, & Mills, 1972). We have used the Lumme system to measure topographical variables in cigarette use, and we are currently collecting similar outcome variables to describe drinking episodes.

The Lumme software serves as a data hub so that multiple data streams can be combined to improve detection accuracy. The program records and consolidates data from other sensor devices present on the smartwatch or linked smartphone, including global position system (GPS), available Bluetooth and Wifi connections, activity (e.g., walking, driving), heartrate, and

skin temperature. These data streams (e.g., GPS location may confirm that the wearer is at a bar, numerous Bluetooth/Wifi connections suggests that the wearer is in a social situation) can be incorporated into the machine learning algorithm to enhance drinking detection (Bae et al., 2018; Parate et al., 2014). These data streams also can be used independently to provide a detailed record of the subject's environment or physical/psychological state when drinking occurs. With sufficient data input, the Lumme system can even be trained to predict drug self-administration several minutes before it occurs (McKee et al., unpublished data). We have developed predictive models that incorporate time of day, location, activity, and movement data to predict tobacco use, and a similar approach could be used to anticipate alcohol consumption before it occurs.

In sum, mobile alcohol biosensors are a promising technology for researchers who wish to passively monitor drinking in real time. Recent developments have significantly improved these devices; integration with other “smart” devices will continue to improve data collection capacity. Although these devices will have a major impact on the alcohol field in general, detection of drinking is particularly important for AUD pharmacotherapy development research.

Improving AUD Pharmacotherapy Research Using Mobile Biosensors

The AUD medication development pipeline is composed of a semi-sequential series of studies in which drugs' effects on health and alcohol use behavior are systematically evaluated. The goal of the medication development pipeline is to identify safe and effective pharmacotherapies and move these drugs towards widespread clinical use quickly and cost effectively. Although preclinical research plays a major role in this process, this article will focus on human research. Figure 2 depicts the AUD pharmacotherapy human research pipeline. This section will review phases of the AUD pharmacotherapy development pipeline and describe how mobile biosensors can be incorporated into standard paradigms.

Phase 1: Safety and Tolerability Testing

Phase 1 human studies are primarily concerned with establishing that new compounds are safe and tolerable in humans and identifying appropriate dosing protocols (Dunbar et al., 2006; Meyer, Straughn, Lo, Schary, & Whitney, 1984). Participants in Phase 1 studies are typically healthy volunteers and the research occurs in a controlled laboratory with medical oversight. With few exceptions, volunteers will not receive alcohol with a candidate medication on board during phase 1 testing. Even in cases where alcohol administration occurs (e.g., alcohol/drug interaction screening), volunteers are monitored directly for safety reasons. Alcohol pharmacokinetic data can be monitored using direct assessment strategies (e.g., blood draw, breath sampling) that can provide a stronger estimate of BAC than mobile biosensors. For these reasons, it is unlikely that mobile biosensors will play a major role in phase 1 testing.

Phase 2: Human Laboratory Studies and Proof-of-Concept Clinical Trials

The goals of phase 2 testing are to refine the use of a candidate medication (e.g., identify optimal dosing patterns) and to establish preliminary evidence for efficacy in humans. Phase 2 research plays an important role in the strategic allocation of resources by predicting whether a medication will work in subsequent confirmatory clinical trials (Roache, 2010). Mistakes are costly. Type I errors (i.e., medication is deemed effective when no effect exists) may lead investigators to expend resources testing a medication that will ultimately fail in Phase 3 testing. More concerning are Type II errors (i.e., no medication effect is detected when a true effect exists), which may cause researchers to give up on a medication with therapeutic potential. Given the gate-keeping role of Phase 2 testing in pharmacotherapy development research, it is critical that these studies provide valid predictions of medications' clinical potential (Litten et al., 2016a).

Phase 2 research includes both human laboratory and preliminary "proof-of-concept" clinical trials. These methodologies serve complimentary purposes in establishing medication efficacy. Human laboratory studies demonstrate that a medication can reduce alcohol use outcomes under controlled experimental conditions. Preliminary clinical trials carry this signal

forward and evaluate the effects of medication on naturalistic alcohol use among treatment-seeking patients. In this section, we will describe how mobile biosensors can improve on both human laboratory studies and preliminary clinical trials.

Human laboratory studies. Phase 2 laboratory studies are often the first to evaluate a medication's therapeutic potential in humans (Ray, Hutchison, & Tartter, 2010; Yardley & Ray, 2017). Laboratory studies are well-suited for this purpose. The laboratory setting affords researchers maximum control over the testing environment so the impact of nuisance variables can be minimized. Outcomes are directly observed so measurement error is minimized. As such, these studies are highly sensitive to medication effects on alcohol use outcomes. There is currently substantial interest in developing procedures that maximize the predictive validity of this research (Litten et al., 2016a).

Human laboratory studies evaluate the effects of medication on behavioral models of real world alcohol use outcomes (Plebani et al., 2012). A list of frequently used paradigms and corresponding medication targets are described in Table 1. As seen in this table, there are multiple paradigms that are used to model core alcohol use processes. The most direct method is to directly measure alcohol consumption during a self-paced drinking task. Another strategy is to evaluate alcohol reactivity using fixed-dose alcohol administration paradigms. Alcohol reactivity and *ad libitum* consumption amount are primary outcome variables. Human laboratory studies also can uncover the behavioral mechanisms by which medications reduce drinking. For example, targeted assessments of subjective alcohol effects during fixed-dose alcohol administration can determine whether medications reduce the stimulating, reinforcing effects of alcohol (King, Volpicelli, Frazer, & O'Brien, 1997).

Clearly human laboratory paradigms provide important contributions to AUD medication testing; however, there are well-recognized limitations to these techniques (Roache, 2010). From an efficiency standpoint, significant financial and human resources are required to coordinate laboratory sessions involving alcohol administration. During most laboratory

screening studies, participants are observed during one or two drinking episodes, which provides an incomplete assessment of the underlying trait (Mischel, Shoda, & Mendoza-Denton, 2002). Another major limitation to this research is that drinking in the laboratory is an imperfect proxy for naturally occurring alcohol use. Human laboratory studies occur in a controlled environment that may limit the external validity of the research (Anderson, Lindsay, & Bushman, 1999; Mitchell, 2012).

An outpatient-based alternative to laboratory-based screening paradigms using mobile biosensors. As described above, laboratory screening studies show if and how medications can alter alcohol use behavior under tightly controlled experimental conditions. Although laboratory drinking behavior is correlated with retrospectively reported alcohol use (Jones et al., 2016), the extent to which changes in laboratory alcohol use generalize to day-to-day drinking is an open question.

Laboratory studies take place in the laboratory out of necessity; researchers historically lack the ability to monitor naturalistic alcohol use outcomes at the level of detail that can be achieved by directly observing participants as they drink. However, mobile biosensors may facilitate direct monitoring of naturalistic alcohol use. Our group is currently developing a biosensor-based medication screening paradigm that we plan to use to evaluate the effects of candidate medications on drinking outside of the laboratory. This paradigm leverages biosensors to collect data at a level of detail that was previously practical only in the laboratory. A brief description of the paradigm is as follows (see also Figure 3 for a temporal schematic of the paradigm). Non-treatment seeking heavy drinkers with AUD are screened and trained to use the biosensor equipment. After a baseline monitoring period, they are assigned to a medication condition (active medication or placebo). Their alcohol use is tracked over the course of the medication maintenance phase. When drinking is detected, the system initiates an assessment protocol and key alcohol use outcomes are assessed using a combination of both active and passive assessment strategies. Table 1 outlines data collection procedures and describes

analogous human laboratory outcomes and Table 2 describes strengths and weaknesses of this paradigm compared to laboratory-based methods.

A primary advantage of this biosensor-based paradigm is that self-paced drinking during naturally occurring episodes are tracked objectively at a high level of detail that parallels direct observation of drinking behavior in the human laboratory. The paradigm will provide a detailed record of medication effects on objective drinking outcomes, such as peak BAC, rate of BAC rise, and drinking topography. From a predictive validity standpoint, medication effects on naturally occurring drinking will likely bear a closer relationship to medication effects during subsequent clinical trials. Although drinking outcomes may be more “noisy” due to variability in drinking contexts, data on drinking contexts (e.g., location, activity, time of day, social network) can be collected using secondary data streams. For example, GPS location and wireless network availability can be used as an indicator drinking location. These data can be used to statistically control for drinking context. The Lumme software can use contextual data to develop predictive models that can detect when participants encounter “high-risk” drinking circumstances (McKee et al., unpublished data). For example, GPS data might be used to identify physical locations at which participants often engage in heavy drinking; an assessment could be prompted when they return to that location even before drinking begins.

Similarly, the detailed record of drinking context will allow researchers to isolate specific situations to characterize how medication effects drinking in high-risk contexts. This technique mirrors laboratory studies in which participants are exposed to alcohol use triggers, such as stress or alcohol cue exposure, and then given access to alcohol (Fox et al., 2012; Thomas, Bacon, Randall, Brady, & See, 2011; Roberts et al., 2017). As a medication screening tool, researchers can investigate whether medications reduce drinking under these high-risk circumstances, based on the hypothesized mechanism of action. For example, stress-drinking paradigms can be used to test the ability of sympatholytic drugs to block stress-induced drinking

(Fox et al., 2012). These effects, while clinically important, may not be observable when participants are not in a high-stress state.

In the future and contingent on improvements in mobile detection technology allowing real-time BAC estimation, an advantage of this paradigm will be that the platform can conduct active prompted assessments of medication targets (e.g., subjective alcohol reactivity, mood, cognitive functioning; King et al., 1997; Ray & Hutchison, 2007; Tiplady, Oshinowo, Thomson, & Drummond, 2009) at key points during naturally occurring drinking episodes. These prompts, which are delivered on participants' smartphones, can be linked to key points in drinking episodes. For example, a prompted assessment could be triggered when a participant first reaches BAC = 80 mg% during a drinking episode. This strategy closely parallels ecological momentary assessment in substance use research (Miranda et al., 2016; Shiffman, 2009), except that assessment prompts are tied to an event rather than being conducted at random or initiated by participants. Table 1 outlines several prompted assessment targets that could be used to probe medication mechanisms; this table also compares the proposed assessment targets to measures frequently included in human laboratory studies. Figure 4 charts the course of a hypothetical natural drinking episode over time as a function of BAC, illustrating how standardized assessments could be delivered when participants reach specific BACs.

Another major advantage of this approach is that medication effects can be assessed over the course of multiple drinking episodes, whereas laboratory studies typically examine medication effects on a single episode of behavior. There is a mathematical advantage to recording multiple drinking episodes: the average of multiple measurement is a more reliable estimate of the underlying trait than a single measurement (Jacob, Tennenbaum, & Krahn, 1987). Further, recording the effect of a medication on multiple drinking episodes allows for the examination of time varying medication effects on drinking. A medication may become more or less effective over time as tolerance or sensitization develops (Marley, Shimosato, Gewigg Thorndike, Goldberg, & Schindler, 1995). Measuring medication effects on alcohol use outcomes

over multiple drinking episodes also improve the efficiency of preliminary testing by facilitating the examination of multiple medication mechanisms over the course of the medication maintenance period. For example, assessments can be tied to both periods of drinking and periods of abstinence. This more comprehensive series of assessments will provide greater insight into the multiple mechanisms by which medication might alter drinking. Given the flexibility of the biosensors and assessment prompts, hypothesis-driven assessment targets could be used to probe a litany of potential medication mechanisms (e.g., mood during periods of abstinence; acute withdrawal symptoms following a night of heavy drinking), all within a single research design.

This paradigm can be used to screen medications in populations that cannot receive alcohol experimentally. Administering alcohol to youth, elderly people, and adults in treatment requires special consideration of risk/benefit ratios, and is prohibited in pregnant women (NIAAA, 2014). The mobile-biosensor based paradigm avoids this issue because researchers are not providing participants with alcohol; data are collected during naturally occurring drinking episodes. As such, researchers using this paradigm will have greater flexibility to evaluate medication effects in populations that cannot participate in laboratory alcohol administration research.

Limitations of this paradigm. Despite the potential of this paradigm to improve on Phase 2 medication screening research, there are several limitations to current alcohol biosensor technology that should be considered. As such, this paradigm should be considered aspirational in nature; additional development of mobile alcohol biosensors will be required to realize aspects of the paradigm. The time delay between initial alcohol consumption and transdermal diffusion may limit researchers' ability to conduct real-time assessments of alcohol effects using TAS. If alcohol consumption is not detected until several hours following drinking initiation, then prompted assessments would not be initiated until the middle or end of a drinking episode. Considering the importance of alcohol reactivity during the ascending limb of the BAC curve

(Newlin & Thomson, 1990), this shortcoming could limit the utility of such assessments for tracking medication effects. For this future application, it will be critically important that researchers improve real-time drinking detection and BAC estimation, either by combining TAS with additional biosensors or developing and using sensors that can perform these functions in real time.

There are several laboratory-based methodologies that cannot be used in the above described paradigm. Increasingly, researchers are using techniques that rely on equipment that cannot be easily mobilized (e.g., positron emission topography, fMRI, eye-tracking cameras; Lukas et al., 2013; Weerts et al., 2008; Roberts et al., 2012) to evaluate neurobiological and cognitive drug effects in the laboratory. Likewise, the proposed paradigm cannot utilize intravenous alcohol administration, which have numerous methodological advantages over oral administration (Zimmermann, O'Connor, & Ramchandani, 2011). Adverse event monitoring may prove challenging, as fewer laboratory visits will be necessary over the monitoring period. However, researchers can screen for treatment emergent events using prompted assessments delivered via smartphone and intervene if serious adverse events are reported. Regarding the use of mobile biosensors to examine medication by situation interactions (e.g., drinking in social contexts vs. alone), researchers cannot standardize the situations in which participants are drinking. These contexts will manifest at different intensities among participants, and this variability will need to be recorded and addressed statistically.

In sum, mobile biosensors have the potential to move early phase 2 medication screening research out of the laboratory and into the real world. Although limitations exist with currently available technology, given the pace of technological development, it is likely that real-time detection will be achieved soon. Laboratory purists may argue that something important is lost in this transition because real-life drinking will never be as clean as laboratory models of alcohol self-administration. It is our view, however, that outpatient-based assessment using biosensors will provide a more rigorous test of potential medication effects on drinking, precisely because

these real-world conditions are closer to those experienced by patients completing outpatient treatment.

Phase 2 proof-of-concept clinical trials. When results of a human laboratory study support medication efficacy, the next step in the pipeline is clinical trial testing. Phase 2 proof-of-concept clinical trials are important because they are typically the first to evaluate medication efficacy when it is used as it would be in the clinic. These trials serve an intermediary role between human laboratory studies and large-scale pivotal Phase 3 testing by translating results from the laboratory to clinical outcomes. In the gold-standard design for medication testing (i.e., randomized clinical trials), patients are randomized to receive active medication or placebo and their drinking is tracked over a treatment period. Investigators identify *a priori* a primary endpoint that determines whether the treatment was effective. Primary endpoints are typically based on patients' self-reported alcohol use.

A major limitation of these clinical trials is that self-report of alcohol use provides a relatively impoverished description of past drinking. There are well-documented shortcomings of retrospective self-report as a measure of past consumption (Del Boca & Noll, 2000), both in terms of accuracy and data quality. Self-report is subject to response biases, such as recall inaccuracy and social desirability, that may lead patients to over- or under-report their alcohol use. In some clinical trials, biomarkers (e.g., gamma-glutamyltransferase, ethyl glucuronide) are used to verify the accuracy of self-report data. However, these biomarkers are relatively insensitive to drinking amount and timeframe, which limit their potential as primary outcome variables (Litten, Bradley, & Moss, 2010). Primary analyses are invariably based on self-reported drinking (Witkiewitz, Finney, Harris, Kivlahan, & Kranzler, 2015). Pharmacokinetic and behavioral variables such as peak BAC and rate of drinking can be estimated indirectly (Carey & Hustad, 2002; Miller, Leckman, Delaney, & Tinkcom, 1992); however, these estimates can be noisy and subject to reporting bias. These problems with currently used assessment

strategies limits our ability to accurately measure treatment effect sizes during these clinical trials.

Mobile biosensors in proof-of-concept Phase 2 clinical trials. Mobile biosensors have the potential to supplement, or even replace, self-reported drinking as a tool for tracking alcohol use during proof of concept clinical trials. Biological verification of abstinence, typically assessed through urine toxicology, is standard practice in treatment trials for other substances of abuse (e.g., opioids, stimulants, cannabis; Fiellin et al., 2014; Jayaram-Lindstrom, Hammarberg, Beck, & Franck, 2008; McRae-Clark et al., 2015). Mobile biosensor platforms will provide an objective, biologically verified record of abstinence.

The biosensor system also will allow researchers to collect a robust set of secondary clinical outcomes. As seen in Table 3, there are multiple biosensor-based secondary endpoints that could be evaluated. Because data is collected and delivered continuously, many of the same prompted assessments described in the previous section could also be used during these clinical trials if drinking is detected. Secondary self-report outcomes, such as craving, can be measured when patients are outside of the clinic, improving the ecological validity of these measurements. This continuous monitoring also will facilitate data collection during critical events, such as during the first relapse following a period of abstinence. Data that is collected in real time during episodes of relapse will offer insight into the contextual drivers of relapse, which will provide mechanistic data on the medication being tested and, perhaps more importantly, clarify why people fail to remain abstinence (Marlatt, 1996).

A smartphone-linked biosensor platform also can be used as a continuous point of contact with clinical trial patients. In most clinical trials, participants attend weekly or biweekly clinic visits over the course of a treatment period. During these clinic visits, participants typically provide biosamples, report adverse events, and report primary (i.e., past week drinking) and secondary (e.g., craving) outcomes, and receive their weekly supply of medication. Clinic visits often include behavioral therapy. Many of these activities are no longer necessary

when biosensors are included in the clinical trial, because outcome variables (i.e., alcohol use, self-reported outcomes, adverse events) are tracked passively or through prompted assessments delivered on smartphones. Computerized behavioral interventions can be delivered through the smartphone platform. Indeed, there has been significant progress recently in adapting behavioral therapies for computerized delivery, including for substance use disorders (Carroll et al., 2008; Kay-Lambkin, Baker, Lewin, & Carr, 2009). By interacting with patients primarily through a biosensor-linked platform that includes a smartphone user interface, fewer clinic visits will be necessary. This shift will reduce burden on both staff and patients and may increase retention. Another potential application of continuous remote contact with patients is that smartphones can be used to deliver medication prompts. Medication taking can be reported and confirmed immediately. This system may improve the poor rates of medication compliance among participants in AUD pharmacotherapy clinical trials (Chick et al., 2000).

Despite the considerable benefits of mobile biosensors to enhance alcohol use detection during clinical trials, additional research is needed to confirm the utility and acceptability of these devices. Although preliminary studies have supported the feasibility of long-term use of mobile biosensors among heavy drinkers in treatment studies (Barnett et al., 2017; Dougherty et al., 2014), it will be important to establish that patients are willing and able to continuously wear newer biosensors (e.g., BACtrack Skyn) that can be removed and must be periodically charged. Another priority will be to identify how treatment response is best defined using mobile biosensor data. Endpoints based on estimated BAC cutoffs— rather than number of drinks—may be more sensitive indicators of treatment response (Litten et al., 2016). It will be important for researchers to strategically select devices and detection criteria that best suit the needs of their protocol. For example, biosensors that show strong sensitivity to detect low-level drinking episodes would be best suited for clinical trials confirmation of complete abstinence is prioritized, whereas biosensors that use more conservative detection parameters may be more appropriate when detecting heavy drinking is the primary goal.

Phase 3: Special Considerations for Pivotal Clinical Trials

Phase 3 clinical trials use similar methodologies as Phase 2 proof-of-concept except that they are scaled up in size and cost. As such, many of the benefits of using biosensors in proof of concept studies also apply to Phase 3 clinical trials. There are, however, additional points to consider that apply specifically to the use of biosensors in Phase 3 pivotal trials. These clinical trials are necessary to obtain regulatory approval for a medication, so specific guidelines must be followed in the design and reporting of results (Litten et al., 2016a). Pivotal trials include an active treatment phase (up to 6 months) during which participants are maintained on medication and attend regular clinical visits. Following the active treatment phase, follow-up assessments are conducted to examine the durability of treatment effects. The FDA and other regulatory bodies offer strict guidelines on acceptable primary endpoints (FDA, 2015).

Currently, approved endpoints are based on self-report of alcohol use, generally collected using daily drinking diary or timeline follow-back (Sobell & Sobell, 1992). In the United States, the current guidelines define acceptable endpoints as the number of participants with no heavy drinking days or percent of participants who achieve abstinence (FDA, 2015). These endpoints were determined based on secondary analysis of large scale clinical trials that tested which outcomes were most sensitive to reductions in drinking consequences (Litten et al., 2016a). The analyses guiding these decisions were based on self-reported drinking data. When mobile biosensors are incorporated into pivotal clinical trials, additional research will be required to determine an ideal endpoint with maximum sensitivity to treatment effects and high predictive validity.

Mobile biosensor platforms may streamline multisite clinical trials by automatically transferring outcome data to a central server location. Centralized data storage can increase efficiency by reducing data management burden. Likewise, the remote nature and reduced participant burden afforded by biosensor-based assessment may allow quicker recruitment of a more diverse sample of participants.

Phase 4: Postmarketing surveillance and dissemination

Phase 4 testing includes systematic assessment of use and efficacy of approved medications in real-world settings and confirmatory clinical trials that occur after a medication is approved to treat AUD. Less emphasis has been placed on Phase 4 research in the AUD treatment development field given the low rates of treatment utilization among people with AUD (Mark, Kassed, Vandivort-Warren, Levit, & Kranzler, 2009). However, additional research in this area could have large public health benefit. Mobile alcohol biosensors may have a future role in routine clinical care, particularly among people being treated for alcohol-related health problems (Barnett, 2015). Because data collected on these biosensors is digitized, treatment outcomes could be easily aggregated and analyzed to track drug efficacy as it is used in clinical settings (Kumar et al., 2013). As a consumer product, mobile alcohol biosensors have the potential to raise awareness of harmful alcohol consumption and facilitate communication between patients and healthcare providers in much the same manner as other home health monitoring devices (Yarows, Julius, & Pickering, 2000).

Challenges of Using Mobile Biosensors in Medication Development Research

We have discussed challenges associated with using mobile biosensors within each phase of medication development research in their respective section. However, there are several issues that cut across each phase that should be considered. Below we describe these issues make recommendations for addressing them.

Equipment Management

Like many advanced technologies used in research, mobile biosensors can be technically challenging for both researchers and participants. Issues such as hardware or software failures, lost wireless connections, and limited battery life can result in lost data. Our group has successfully mitigated some of these issues by piloting our experiments with expert users before beginning participant recruitment. Problems can be anticipated and addressed so that data loss is minimized. It is also important to create an ongoing dialogue between research staff and other

investigators so that technical issues can be addressed as they emerge. Participant feedback is equally helpful for fine-tuning protocols to improve mobile biosensor ease-of-use and reduce technical complications. Conversely, it is helpful to adequately train participants on proper use of the equipment. Such training can help ensure that participants know how to care for and maintain the biosensors and avoid actions (e.g., submersion in water) that may damage the devices. For example, we educate participants about the importance of charging the devices overnight to avoid power loss during the day.

Compliance

It is critical that research participants continuously wear the biosensors. The SCRAM TAS comes equipped with a tamper-proof lock to ensure continuous use, but other devices (i.e., BACTrack Skyn, WrisTAS, smartwatches) can be easily removed. Temperature and skin conductance/resistance sensors can alert researchers when the device is removed or turned off, but alcohol use data cannot be collected while the device is removed. Poor compliance in clinical trials is a recognized problem and researchers have developed strategies for improving protocol adherence with varying degrees of success (Haynes, McKibbin, & Kanani, 1996). For example, providing participants with adequate training on using the device and reinforcing compliance will likely improve adherence. Preliminary outpatient studies suggest that participants tolerated wearing SCRAM devices up to three months, although some participant concern, and even dropout, associated with physical discomfort and social embarrassment was noted (Alessi, Barnett, & Petry, 2017; Barnett et al., 2017; Dougherty et al., 2014). However, participants in one trial reported that many of these issues could be remediated by reducing the size of the SCRAM device (Alessi et al., 2017). The newer generation of devices are significantly less bulky and conspicuous, although the lack of anti-tamper features might reduce the high rates of compliance observed in pilot work using the SCRAM (Barnett et al., 2017). Additional research will be necessary to confirm the feasibility and acceptability of newer devices (Herrington et al., 2018).

Data Management and Analysis

Mobile biosensors provide rich datasets that can track nuanced physiological and psychological processes; however, there are significant challenges to managing and analyzing this high-density data, particularly when combining data streams from multiple biosensors (Parate et al., 2014). A single biosensor can produce upwards of 500,000 data points per subject per day. Experts in mobile health have outlined appropriate machine learning techniques for making classification decisions based on complex data streams (Kumar et al., 2013). Developing appropriate classification and detection algorithms will be necessary to infer variables of interest (e.g., drinking gestures, binge episodes) from raw biosensor data. Projects involving biosensors should include team members with expertise in machine learning and pattern recognition to support data storage and processing. A related issue is that real-time data collection requires ongoing data transfer from participants' storage devices to a central server, particularly when research protocols require event-linked assessments. Special expertise is required to develop and maintain the computer software and hardware required to facilitate the flow of this high-density data to a centralized storage location.

Confidentiality

Researchers are required to take steps to guarantee the privacy of participants' personal health information. Data that is collected and transferred through wireless networks is at risk for information breach. There are procedures that can be used to minimize risk of data breach, including encrypting data using Advanced Encryption Standard and transferring data using Transport Layer Security protocols (see Martinez-Perez, de la Torre-Diez, & Lopez-Coronado, 2015 for a technical review of data security recommendations in mobile health). Participants should be informed of potential of confidentiality breaches and steps that investigators have taken to minimize risks. Researchers should also consult with experts in information technology security to ensure that data safeguards are up to date.

Conclusions

Mobile biosensors for alcohol use detection have great potential to improve the efficiency and validity of medication development research. The approaches proposed here can help realize strategic goals identified by NIAAA (Litten, Falk, et al., 2016), primarily by augmenting data collection in Phase 2 mechanistic and proof-of-concept research and providing an objective endpoint in Phase 2 and 3 clinical trials. Future research goals include establishing the feasibility and acceptability of using newer TAS devices in medication testing research, integrating alcohol use sensors with other mobile health monitoring devices, and developing standardized protocols using mobile biosensors in medication testing paradigms.

References

- Alessi, S. M., Barnett, N. P., & Petry, N. M. (2017). Experiences with SCRAMx alcohol monitoring technology in 100 alcohol treatment outpatients. *Drug and Alcohol Dependence*, 178, 417-424.
- Anderson, C. A., Lindsay, A. J., & Bushman, B. J. (1999). Research in the psychological laboratory: Truth or triviality? *Current Directions in Psychological Science*, 8, 3-9.
- Anderson, J. C., & Hlastala, M. P. (2006). The kinetics of transdermal ethanol exchange. *Journal of Applied Physiology*, 100, 649-655.
- Bae, S., Chung, T., Ferreira, D., Dey, A. K., & Suffoletto, B. (2018). Mobile phone sensors and supervised machine learning to identify alcohol use events in young adults: Implications for just-in-time adaptive interventions. *Addictive Behaviors*, 83, 42-47.
- Barnett, N. P., Meade, E. B., & Glynn, T. R. (2014). Predictors of detection of alcohol use episodes using a transdermal alcohol sensor. *Experimental and Clinical Psychopharmacology*, 22, 86-96.
- Barnett, N. P. (2015). Alcohol sensors and their potential for improving clinical care. *Addiction*, 110, 1-3.
- Barnett, N. P., Celio, M. A., Tidey, J. W., Murphy, J. G., Colby, S. M., & Swift, R. M. (2017). A preliminary randomized controlled trial of contingency management for alcohol use reduction using a transdermal alcohol sensor. *Addiction*, 112, 1025-1035.
- Brandon, T. H., Drobos, D. J., Unrod, M., Heckman, B. W., Oliver, J. A., Roetzheim, R. C., et al. (2011). Varenicline effects on craving, cue reactivity, and smoking reward. *Psychopharmacology (Berl)*, 218, 391-403.
- Brown, D. J. (1985). A method for determining the excretion of volatile substances through skin. *Methods and Findings in Experimental and Clinical Pharmacology*, 7, 269-274.
- Bruen, D., Delaney, C., Florea, L., & Diamond, D. (2017). Glucose sensing for diabetes monitoring: Recent developments. *Sensors*, 17, e1866.

Campbell, A. S., Kim, J., & Wang, J. (2018). Wearable electrochemical alcohol biosensors.

Current Opinion in Electrochemistry. Advanced online publication.

Carey, K. B., & Hustad, J. T. (2002). Are retrospectively reconstructed blood alcohol concentrations accurate? Preliminary results from a field study. *Journal of Studies on Alcohol*, 63, 762-766.

Carreiro, S., Smelson, D., Ranney, M., Horvath, K. J., Picard, R. W., Boudreaux, E. D., et al. (2015). Real-time mobile detection of drug use with wearable biosensors: A pilot study. *Journal of Medical Toxicology*, 11, 73-79.

Carroll, K. M., Ball, S. A., Martino, S., Nich, C., Babuscio, T. A., Nuro, K. F., et al. (2008). Computer-assisted delivery of cognitive-behavioral therapy for addiction: A randomized trial of CBT4CBT. *American Journal of Psychiatry*, 165, 881-888.

Chick, J., Anton, R., Chечinski, K., Croop, R., Drummond, D. C., Farmer, R., et al. (2000). A multicentre, randomized, double-blind, placebo-controlled trial of naltrexone in the treatment of alcohol dependence or abuse. *Alcohol and Alcoholism*, 35, 587-593.

Del Boca, F. K., & Noll, J. A. (2000). Truth or consequences: The validity of self-report data in health services research on addictions. *Addiction*, 95, 347-360.

Dougherty, D. M., Charles, N. E., Acheson, A., John, S., Furr, R. M., & Hill-Kapturczak, N. (2012). Comparing the detection of transdermal and breath alcohol concentrations during periods of alcohol consumption ranging from moderate drinking to binge drinking. *Experimental and Clinical Psychopharmacology*, 20, 373-381.

Dougherty, D. M., Hill-Kapturczak, N., Liang, Y., Karns, T. E., Cates, S. E., Lake, S. L., et al. (2014). Use of continuous transdermal alcohol monitoring during a contingency management procedure to reduce excessive alcohol use. *Drug and Alcohol Dependence*, 142, 301-306.

Dumett, M., Rosen, G., Sabat, J., Shaman, A., Tempelman, L., Wang, C., et al. (2008). Deconvolving an estimate of breath measured blood alcohol concentration from

- biosensor collected transdermal ethanol data. *Applied Mathematics and Computation*, 196, 724-743.
- Dunbar, J. L., Turncliff, R. Z., Dong, Q., Silverman, B. L., Ehrich, E. W., & Lasseter, K. C. (2006). Single- and multiple-dose pharmacokinetics of long-acting injectable naltrexone. *Alcoholism: Clinical & Experimental Research*, 30, 480-490.
- Falk, D. E., Litten, R. Z., Anton, R. F., Kranzler, H. R., Johnson, B. A., & Active, W. (2014). Cumulative proportion of responders analysis (CPRA) as a tool to assess treatment outcome in alcohol clinical trials. *Journal of Studies on Alcohol and Drugs*, 75, 335-346.
- Fensli, R., Gunnarson, E., & Gundersen, T. (2005). A wearable ECG-recording system for continuous arrhythmia monitoring in a wireless tele-home-care situation. *18th IEEE Symposium on Computer-Based Medical Systems, Proceedings*, 407-412.
- Fiellin, D. A., Schottenfeld, R. S., Cutter, C. J., Moore, B. A., Barry, D. T., & O'Connor, P. G. (2014). Primary care-based buprenorphine taper vs maintenance therapy for prescription opioid dependence: a randomized clinical trial. *JAMA Internal Medicine*, 174, 1947-1954.
- Flango, V. E., & Cheesman, F. L. (2009). Effectiveness of the SCRAM alcohol monitoring device: A preliminary test. *Drug Court Review*, 6, 109-134.
- Food and Drug Administration (2015). Alcoholism: Developing drugs for treatment. Retrieved from <https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm433618.pdf>
- Fox, H. C., Anderson, G. M., Tuit, K., Hansen, J., Kimmerling, A., Siedlarz, K. M., et al. (2012). Prazosin effects on stress- and cue-induced craving and stress response in alcohol-dependent individuals: Preliminary findings. *Alcoholism: Clinical & Experimental Research*, 36, 351-360.

- Garbutt, J. C., Greenblatt, A. M., West, S. L., Morgan, L. C., Kampov-Polevoy, A., Jordan, H. S., et al. (2014). Clinical and biological moderators of response to naltrexone in alcohol dependence: A systematic review of the evidence. *Addiction*, *109*, 1274-1284.
- Godwin, M., Ruhland, L., Casson, I., MacDonald, S., Delva, D., Birtwhistle, R., et al. (2003). Pragmatic controlled clinical trials in primary care: the struggle between external and internal validity. *BMC Medical Research Methodology*, *3*, 28.
- Hallgren, K. A., & Witkiewitz, K. (2013). Missing data in alcohol clinical trials: a comparison of methods. *Alcoholism: Clinical & Experimental Research*, *37*, 2152-2160.
- Haynes, R. B., McKibbon, K. A., & Kanani, R. (1996). Systematic review of randomised trials of interventions to assist patients to follow prescriptions for medications. *The Lancet*, *348*, 383-386.
- Herrington, W. G., Goldsack, J. C., & Landray, M. J. (2018). Increasing the use of mobile technology-derived endpoints in clinical trials. *Clinical Trials*, *15*, 313-315.
- Higley, A. E., Crane, N. A., Spadoni, A. D., Quello, S. B., Goodell, V., & Mason, B. J. (2011). Craving in response to stress induction in a human laboratory paradigm predicts treatment outcome in alcohol-dependent individuals. *Psychopharmacology*, *218*, 121-129.
- Jacob, T., Tennenbaum, D. L., & Krahn, G. (1987). Factors influencing the reliability and validity of observation data. In T. Jacob (Ed.), *Family Interaction and Psychopathology: Theories, Methods and Findings* (pp. 297-328). Boston, MA: Springer US.
- Jayaram-Lindstrom, N., Hammarberg, A., Beck, O., & Franck, J. (2008). Naltrexone for the treatment of amphetamine dependence: A randomized, placebo-controlled trial. *American Journal of Psychiatry*, *165*, 1442-1448.
- Jonas, D. E., Amick, H. R., Feltner, C., Bobashev, G., Thomas, K., Wines, R., et al. (2014). Pharmacotherapy for adults with alcohol use disorders in outpatient settings: A systematic review and meta-analysis. *JAMA*, *311*, 1889-1900.

- Jones, A., Button, E., Rose, A. K., Robinson, E., Christiansen, P., Di Lemma, L., et al. (2016). The ad-libitum alcohol 'taste test': Secondary analyses of potential confounds and construct validity. *Psychopharmacology*, 233, 917-924.
- Jung, K., Jim, J., Awofeso, O., Kim, H., Regnier, F., & Bae, E. (2015). Smartphone-based colorimetric analysis for detection of saliva alcohol concentrations. *Applied Optics*, 54, 9183-9189.
- Karns-Wright, T. E., Roache, J. D., Hill-Kapturczak, N., Liang, Y., Mullen, J., & Dougherty, D. M. (2016). Time delays in transdermal alcohol concentrations relative to breath alcohol concentrations. *Alcohol and Alcoholism*, 52, 35-41.
- Kay-Lambkin, F. J., Baker, A. L., Lewin, T. J., & Carr, V. J. (2009). Computer-based psychological treatment for comorbid depression and problematic alcohol and/or cannabis use: A randomized controlled trial of clinical efficacy. *Addiction*, 104, 378-388.
- Kim, J., Jeerapan, I., Imani, S., Cho, T. N., Bhandodkar, A., Cinti, S., et al. (2016). Noninvasive alcohol monitoring using a wearable tattoo-based iontophoretic-biosensing system. *ACS Sensors*, 1, 1011-1019.
- King, A. C., Volpicelli, J. R., Frazer, A., & O'Brien, C. P. (1997). Effect of naltrexone on subjective alcohol response in subjects at high and low risk for future alcohol dependence. *Psychopharmacology*, 129, 15-22.
- Kumar, S., Nilsen, W. J., Abernethy, A., Atienza, A., Patrick, K., Pavel, M., et al. (2013). Mobile health technology evaluation: The mHealth evidence workshop. *American Journal of Preventive Medicine*, 45, 228-236.
- Leffingwell, T. R., Cooney, N. J., Murphy, J. G., Luczak, S., Rosen, G., Dougherty, D. M., et al. (2013). Continuous objective monitoring of alcohol use: Twenty-first century measurement using transdermal sensors. *Alcoholism: Clinical & Experimental Research*, 37, 16-22.

- Litten, R. Z., Bradley, A. M., & Moss, H. B. (2010). Alcohol biomarkers in applied settings: recent advances and future research opportunities. *Alcoholism: Clinical & Experimental Research*, 34, 955-967.
- Litten, R. Z., Egli, M., Heilig, M., Cui, C., Fertig, J. B., Ryan, M. L., et al. (2012). Medications development to treat alcohol dependence: A vision for the next decade. *Addiction Biology*, 17, 513-527.
- Litten, R. Z., Falk, D. E., Ryan, M. L., & Fertig, J. B. (2016a). Discovery, development, and adoption of medications to treat alcohol use disorder: Goals for the phases of medications development. *Alcoholism: Clinical & Experimental Research*, 40, 1368-1379.
- Litten, R. Z., Ryan, M. L., Falk, D. E., Reilly, M., Fertig, J. B., & Koob, G. F. (2015). Heterogeneity of alcohol use disorder: Understanding mechanisms to advance personalized treatment. *Alcoholism: Clinical & Experimental Research*, 39, 579-584.
- Litten, R. Z., Ryan, M. L., Fertig, J. B., Falk, D. E., Johnson, B., Dunn, K. E., et al. (2013). A double-blind, placebo-controlled trial assessing the efficacy of varenicline tartrate for alcohol dependence. *Journal of Addiction Medicine*, 7, 277-286.
- Litten, R. Z., Wilford, B. B., Falk, D. E., Ryan, M. L., & Fertig, J. B. (2016b). Potential medications for the treatment of alcohol use disorder: An evaluation of clinical efficacy and safety. *Substance Abuse*, 37, 286-298.
- Luczak, S. E., Hawkins, A. L., Dai, Z., Wichmann, R., Wang, C., & Rosen, I. G. (2018). Obtaining continuous BrAC/BAC estimates in the field: A hybrid system integrating transdermal alcohol biosensor, *Intellidrink* smartphone app, and *BrAC Estimator* software tools. *Addictive Behaviors*, 83, 48-55.
- Lukas, S. E., Lowen, S. B., Lindsey, K. P., Conn, N., Tartarini, W., Rodolico, J., et al. (2013). Extended-release naltrexone (XR-NTX) attenuates brain responses to alcohol cues in alcohol-dependent volunteers: A BOLD fMRI study. *Neuroimage*, 78, 176-185.

- Mark, T. L., Kassed, C. A., Vandivort-Warren, R., Levit, K. R., & Kranzler, H. R. (2009). Alcohol and opioid dependence medications: Prescription trends, overall and by physician specialty. *Drug and Alcohol Dependence*, 99, 345-349.
- Marlatt, G. A. (1996). Taxonomy of high-risk situations for alcohol relapse: Evolution and development of a cognitive-behavioral model. *Addiction*, 91, 37-49.
- Marley, R. J., Shimosato, K., Gewiss, M., Thorndike, E., Goldberg, S. R., & Schindler, C. W. (1995). Long-term sensitization to the behavioral effects of naltrexone is associated with regionally specific changes in the number of μ and δ opioid receptors in rat brain. *Life Sciences*, 56, 767-774.
- Marques, P. R., & McKnight, A. S. (2007). *Evaluating transdermal alcohol measuring devices*: United States. National Highway Traffic Safety Administration.
- Martinez-Perez, B., de la Torre-Diez, I., & Lopez-Coronado, M. (2015). Privacy and security in mobile health apps: A review and recommendations. *Journal of Medical Systems*, 39, 181.
- Mason, B. J., Light, J. M., Williams, L. D., & Drobes, D. J. (2009). Proof-of-concept human laboratory study for protracted abstinence in alcohol dependence: Effects of gabapentin. *Addiction Biology*, 14, 73-83.
- Mathias, C. W., Hill-Kapturczak, N., Karns-Wright, T. E., Mullen, J., Roache, J. D., Fell, J. C., et al. (2018). Translating transdermal alcohol monitoring procedures for contingency management among adults recently arrested for DWI. *Addictive Behaviors*, 83, 56-63.
- McKee, S. A., Harrison, E. L. R., O'Malley, S. S., Krishnan-Sarin, S., Shi, J., Tetrault, J. M., et al. (2009). Varenicline reduces alcohol self-administration in heavy-drinking smokers. *Biological Psychiatry*, 66, 185-190.
- McRae-Clark, A. L., Baker, N. L., Gray, K. M., Killeen, T. K., Wagner, A. M., Brady, K. T., et al. (2015). Bupirone treatment of cannabis dependence: A randomized, placebo-controlled trial. *Drug and Alcohol Dependence*, 156, 29-37.

- Meyer, M. C., Straughn, A. B., Lo, M. W., Schary, W. L., & Whitney, C. C. (1984). Bioequivalence, dose-proportionality, and pharmacokinetics of naltrexone after oral administration. *Journal of Clinical Psychiatry*, 45, 15-19.
- Miller, W. R., Leckman, A. L., Delaney, H. D., & Tinkcom, M. (1992). Long-term follow-up of behavioral self-control training. *Journal of Studies on Alcohol*, 53, 249-261.
- Miranda, R., Jr., MacKillop, J., Treloar, H., Blanchard, A., Tidey, J. W., Swift, R. M., et al. (2016). Biobehavioral mechanisms of topiramate's effects on alcohol use: An investigation pairing laboratory and ecological momentary assessments. *Addiction Biology*, 21, 171-182.
- Mischel, W., Shoda, Y., & Mendoza-Denton, R. (2002). Situation-behavior profiles as a locus of consistency in personality. *Current Directions in Psychological Science*, 11, 50-54.
- Mitchell, G. (2012). Revisiting truth or triviality: The external validity of research in the psychological laboratory. *Perspectives on Psychological Science*, 7, 109-117.
- Mohan, A. M V, Windmiller, J. R., Mishra, R. K., Wang, J. (2017). Continuous minimally-invasive alcohol monitoring using microneedle sensor arrays. *Biosensors and Bioelectronics*, 91, 574-579.
- National Institute of Alcohol Abuse and Alcoholism. (2014). *Administering alcohol in human studies*. Retrieved from: <https://www.niaaa.nih.gov/Resources/ResearchResources/job22.htm>
- Newlin, D. B., & Thomson, J. B. (1990). Alcohol challenge with sons of alcoholics: A critical review and reanalysis. *Psychological Bulletin*, 108, 383-402.
- Pantelopoulos, A., & Bourbakis, N. G. (2010). A survey on wearable sensor-based systems for health monitoring and prognosis. *IEEE Transactions on Systems Man and Cybernetics Part C: Applications and Reviews*, 40, 1-12.
- Parate, A., Chiu, M. C., Chadowitz, C., Ganesan, D., & Kalogerakis, E. (2014). RisQ: Recognizing smoking gestures with inertial sensors on a wristband. *MobiSys*, 2014, 149-161.

- Plebani, J. G., Ray, L. A., Morean, M. E., Corbin, W. R., MacKillop, J., Amlung, M., et al. (2012). Human laboratory paradigms in alcohol research. *Alcoholism: Clinical & Experimental Research*, 36, 972-983.
- Poh, M. Z., Loddenkemper, T., Reinsberger, C., Swenson, N. C., Goyal, S., Sabtala, M. C., et al. (2012). Convulsive seizure detection using a wrist-worn electrodermal activity and accelerometry biosensor. *Epilepsia*, 53, 93-97.
- Center for Disease Control and Prevention (2013). Alcohol Related Disease Impact (ARDI) application. Retrieved Jan 2, 2018, from www.cdc.gov/ARDI
- Rankin, H., Hodgson, R., & Stockwell, T. (1979). The concept of craving and its measurement. *Behavioral Research and Therapy*, 17, 389-396.
- Ray, L. A., & Hutchison, K. E. (2007). Effects of naltrexone on alcohol sensitivity and genetic moderators of medication response: A double-blind placebo-controlled study. *Archives of General Psychiatry*, 64, 1069-1077.
- Ray, L. A., Hutchison, K. E., & Tartter, M. (2010). Application of human laboratory models to pharmacotherapy development for alcohol dependence. *Current Pharmaceutical Design*, 16, 2149-2158.
- Roache, J. D. (2010). Role of the human laboratory in the development of medications for alcohol and drug dependence. In: Johnson B. (ed), *Addiction Medicine*. Springer, New York, NY.
- Roache, J. D., Karns, T. E., Hill-Kapturczak, N., Mullen, J., Liang, Y., Lamb, R. J., et al. (2015). Using transdermal alcohol monitoring to detect low-level drinking. *Alcoholism: Clinical & Experimental Research*, 39, 1120-1127.
- Roberts, W., Fillmore, M. T., & Milich, R. (2012). Drinking to distraction: Does alcohol increase attentional bias in adults with ADHD? *Experimental and Clinical Psychopharmacology*, 20, 107-117.

- Roberts, W., Harrison, E. L. R., & McKee, S. A. (2017). Effects of varenicline on alcohol cue reactivity in heavy drinkers. *Psychopharmacology*, 234, 2737-2745.
- Roberts, W., & McKee, S. A. (2018). Effects of varenicline on cognitive performance in heavy drinkers: Dose-response effects and associations with drinking outcomes. *Experimental and Clinical Psychopharmacology*, 26, 49-57.
- Roche, D. J., Bujarski, S., Hartwell, E., Green, R., & Ray, L. A. (2015). Combined varenicline and naltrexone treatment reduces smoking topography intensity in heavy-drinking smokers. *Pharmacology, Biochemistry, and Behavior*, 134, 92-98.
- Sacks, J. J., Gonzales, K. R., Bouchery, E. E., Tomedi, L. E., & Brewer, R. D. (2015). 2010 national and state costs of excessive alcohol consumption. *American Journal of Preventative Medicine*, 49, 73-79.
- Shiffman, S. (2009). Ecological momentary assessment (EMA) in studies of substance use. *Psychological Assessment*, 21, 486-497.
- Sobell, L. C., & Sobell, M. B. (1992). Timeline follow-back. In: Litten R.Z., Allen J.P. (eds) *Measuring Alcohol Consumption*. Humana Press, Totowa, NJ.
- Sobell, M. B., Schaefer, H. H., & Mills, K. C. (1972). Differences in baseline drinking behavior between alcoholics and normal drinkers. *Behavioral Research and Therapy*, 10, 257-267.
- Steinhubl, S. R., Muse, E. D., & Topol, E. J. (2013). Can mobile health technologies transform health care? *JAMA*, 310, 2395-2396.
- Swift, R. (2003). Direct measurement of alcohol and its metabolites. *Addiction*, 98, 73-80.
- Thomas, S. E., Bacon, A. K., Randall, P. K., Brady, K. T., & See, R. E. (2011). An acute psychosocial stressor increases drinking in non-treatment-seeking alcoholics. *Psychopharmacology*, 218, 19-28.

- Tiplady, B., Oshinowo, B., Thomson, J., & Drummond, G. B. (2009). Alcohol and cognitive function: Assessment in everyday life and laboratory settings using mobile phones. *Alcoholism: Clinical & Experimental Research*, 33, 2094-2102.
- Weerts, E. M., Kim, Y. K., Wand, G. S., Dannals, R. F., Lee, J. S., Frost, J. J., et al. (2008). Differences in delta- and mu-opioid receptor blockade measured by positron emission tomography in naltrexone-treated recently abstinent alcohol-dependent subjects. *Neuropsychopharmacology*, 33, 653-665.
- Witkiewitz, K., Finney, J. W., Harris, A. H., Kivlahan, D. R., & Kranzler, H. R. (2015). Guidelines for the reporting of treatment trials for alcohol use disorders. *Alcoholism: Clinical & Experimental Research*, 39, 1571-1581.
- Yardley, M. M., & Ray, L. A. (2017). Medications development for the treatment of alcohol use disorder: Insights into the predictive value of animal and human laboratory models. *Addiction Biology*, 22, 581-615.
- Yarows, S. A., Julius, S., & Pickering, T. G. (2000). Home blood pressure monitoring. *Archives of Internal Medicine*, 160, 1251-1257.
- Zimmermann, U. S., O'Connor, S., & Ramchandani, V. A. (2011). Modeling alcohol self-administration in the human laboratory. In: Sommer W., Spanagel R. (eds) *Behavioral Neurobiology of Alcohol Addiction. Current Topics in Behavioral Neurosciences*, vol 13. Springer: Berlin, Heidelberg.

Table Captions

Table 1. Comparison of human laboratory paradigms with mobile biosensor-based outpatient testing in non-treatment seeking drinkers

Table 2. Comparison of strengths and weaknesses of human laboratory versus mobile biosensors-based medication screening paradigms

Table 3. Recommended mobile biosensor-derived endpoints for outpatient clinical trial

Figure Captions

Figure 1. Prototype BACtrack Skyn unit linked to a smartphone. The above image is a prototype and not the final production unit. The depicted app is an alpha version designed for demonstration purposes.

Figure 2. Goals and paradigms of each stage of human testing in the AUD medication development pipeline.

Figure 3. Timeline of a Phase 2 biosensor-based outpatient medication screening paradigm in non-treatment seekers.

Figure 4. BAC during exemplary drinking episode with prompted assessments. Assessments can be delivered via smartphone to evaluate variables of interest (e.g., craving, mood). Early intoxication and binge-level intoxication assessments are analogous to BAC-linked assessments in frequently performed in Phase 2 medication development research that involves alcohol administration. This application of mobile alcohol biosensors requires real-time alcohol use detection and BAC estimation. Additional development will be required before a biosensor platform is capable of delivering these prompted assessments.

	Human Laboratory Approach		Mobile Biosensor Approach	
	Method	Primary Outcomes	Method	Primary Outcomes ^c
Medication Effect				
Alter responses to alcohol on ascending/descending limb of BAC	Fixed-dose alcohol administration Alcohol clamp	Subjective ^a /objective ^b alcohol reactivity during ascending/descending BAC curve	Alcohol reactivity assessed during naturally occurring drinking episodes	Subjective ^d /objective ^e alcohol reactivity at specific BACs (collected using smartphone assessments)
Reduce alcohol consumption	<i>Ad libitum</i> alcohol use (oral or IV)	Decision to drink Amount consumed Drinking topography ^f	Alcohol self-administration tracked during naturally occurring drinking episodes	Decision to drink Peak BAC (calculated from TAC) Drinking topography (calculated from Lumme Startwatch)
Reduce heavy drinking once drinking starts	Hybrid fixed-low dose alcohol + self-administration (oral or IV)	Subjective ^a /objective ^b reactivity to fixed dose Amount consumed Drinking topography ^f	Continued drinking tracked when low-level drinking is detected	Subjective ^d /objective ^e alcohol reactivity to fixed dose (assessed on smartphone) Peak BAC (calculated from TAC) Drinking topography ^f (calculated from Lumme Startwatch)
Reduce responses to alcohol-related cues	Cue-reactivity assessment	Subjective ^a /objective ^b reactivity	Assessment conducted when system detects presence of cues associated with prior naturalistic drinking and BAC = 0.0%	Subjective ^d /objective ^e reactivity (assessed on smartphone)

Notes. BAC estimates reported in the primary outcomes column are estimated from smartphone and TAC data. BAC-related outcomes in the primary outcomes column are aspirational in nature and will require additional development before these applications of the platform can be realized.

^aSubjective laboratory assessments include craving, mood, sedation, stimulation, subjective intoxication; ^bObjective laboratory assessments include blood alcohol levels, physiological reactivity, perceptual-motor effects, cognitive effects, hormone levels, neuroimaging); ^cPrimary outcomes are generated using data collected on transdermal alcohol sensor (transdermal alcohol concentration, temperature/skin conductance), Lumme smartwatch (Accelerometer/gyrometer coordinates, GPS coordinates, wireless network detection, heartrate), prompted smartphone assessments (response prompts, cognitive task performance);

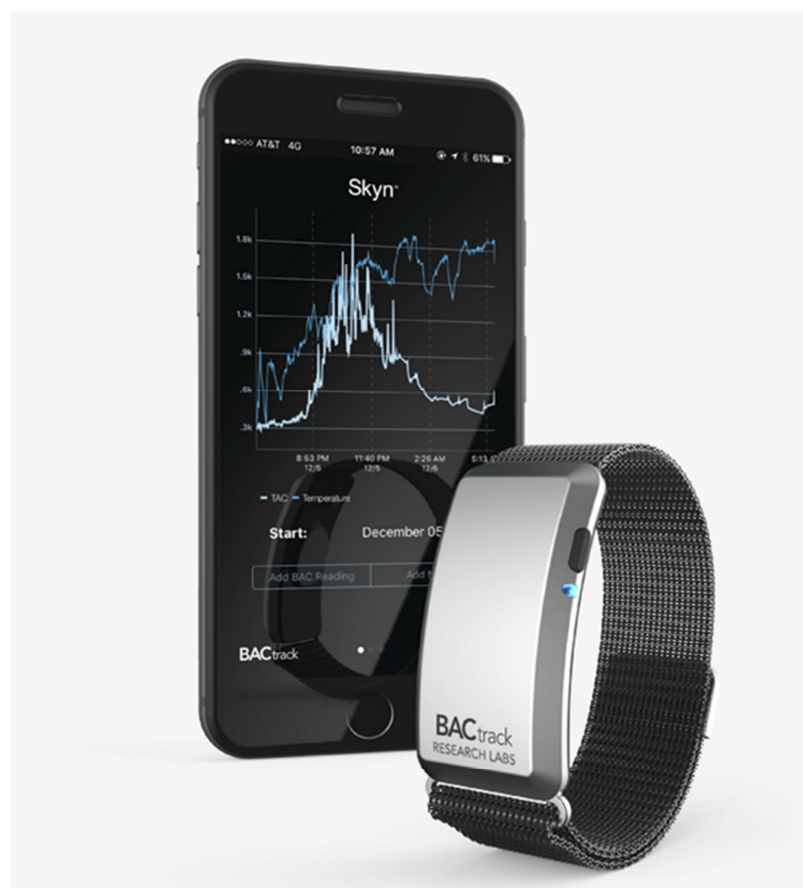
^dSubjective mobile biosensor assessments include craving, mood, sedation, stimulation, subjective intoxication; ^eObjective biosensor assessments include transdermal alcohol concentrations, perceptual-motor effects, cognitive effects, heartrate; ^fDrinking topography include start/end times of drinking, number of sips, inter-sip interval, rate of drinking; ^g Contextual information: GPS, time of day, social network (Bluetooth), and activity tied to drinking episodes.

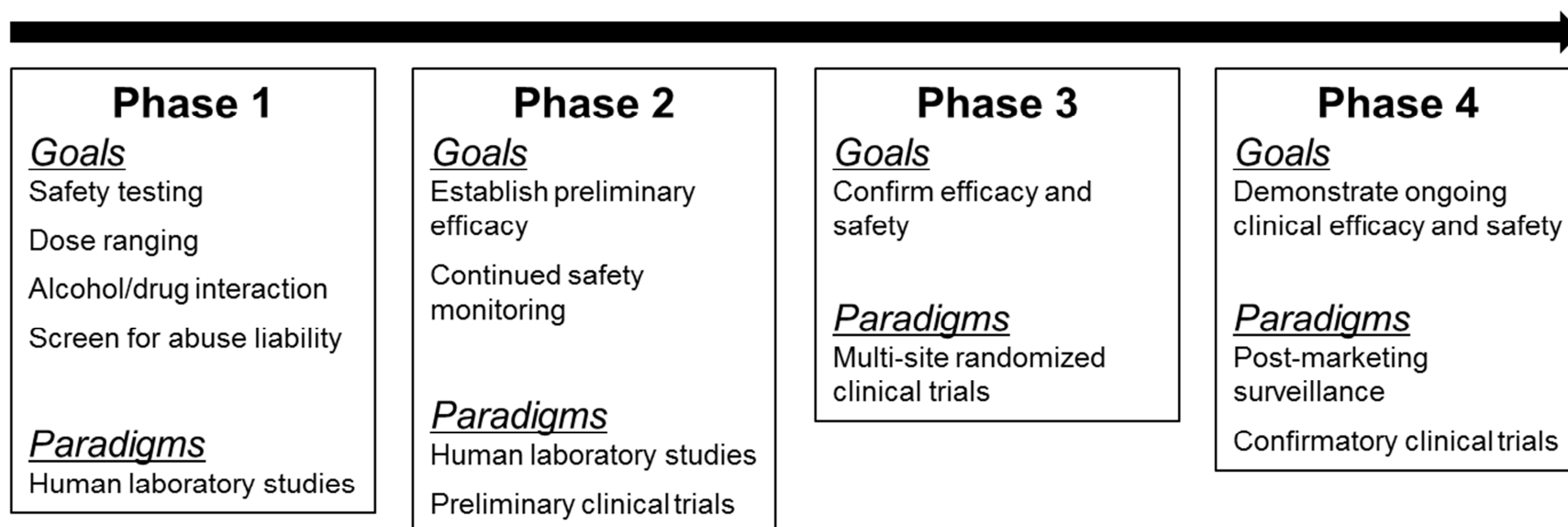
Human Laboratory Paradigms	Mobile Biosensor Paradigm
Strengths	
• High degree of experimental control • Direct observation of behavior • Flexibility of assessments being collected • Precise assessment timing	• Detailed record of behavior • Multiple episodes assessed • Contextual information linked to drinking episodes^s • Improved generalizability • Greater flexibility in participant selection
Weaknesses	
• Single episode assessment • Non-representative setting • Resource intensive • Multiple protocols to test different mechanisms • Some populations cannot receive alcohol experimentally	• Lack of experimental control • Risk of non-compliance • Some techniques (e.g., neuroimaging, IV alcohol administration) cannot be used • Lack of precision in assessment timing

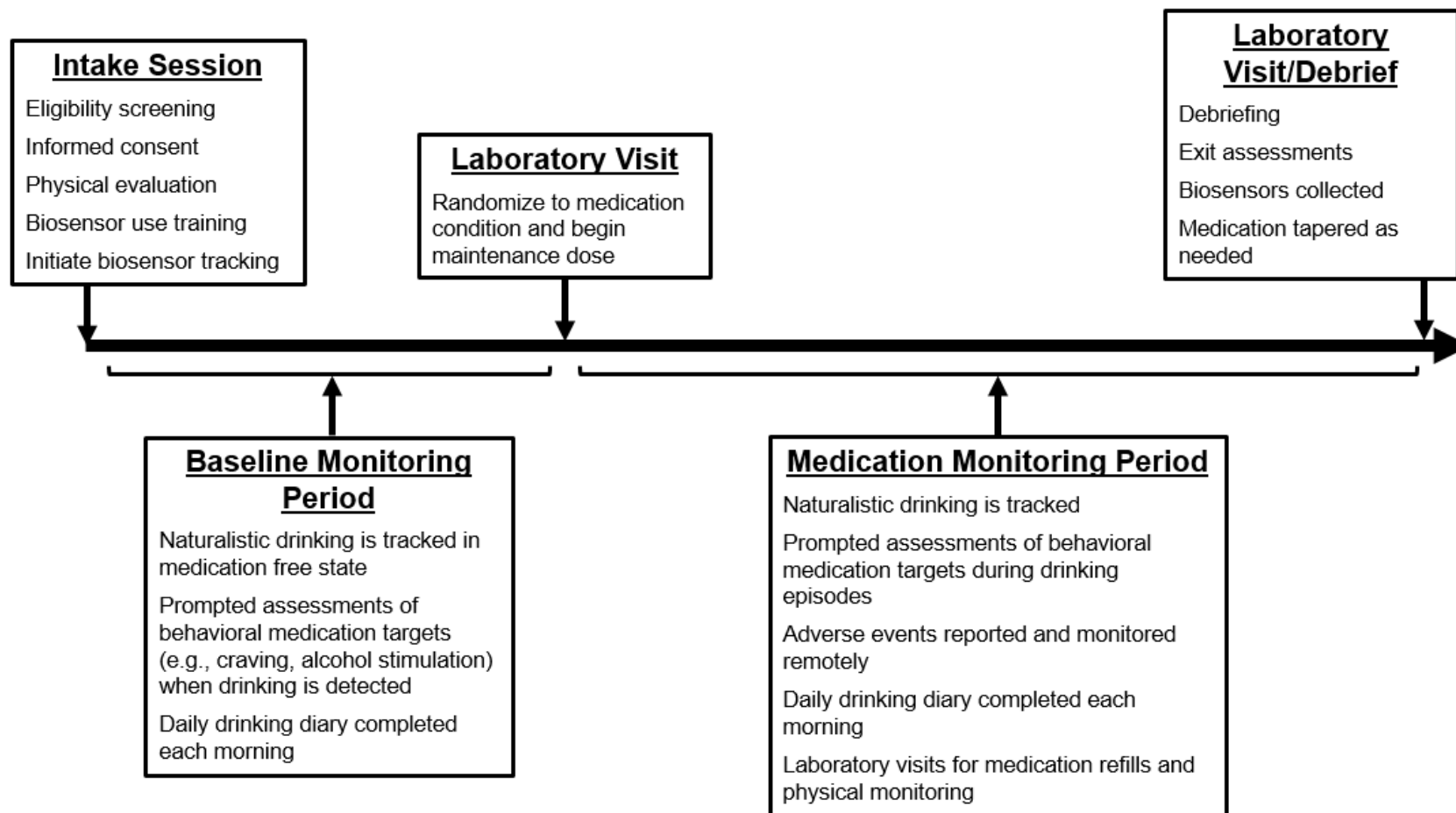
Methodology	Primary Biosensor ^a Outcomes	Secondary Biosensor ^a Data Collected	Advantages vs. Self-Reported Drinking
Outpatient Clinical Trial: Treatment seeking participants with AUD are maintained on medication during treatment period and alcohol use is monitored	FDA approved drinking outcomes (percent subjects with no drinking days/no heavy drinking days) Treatment emergent adverse events	TAS drinking outcomes	Objective/passive recording of drinking
		Drinking topography	Assessments tied to key critical events (e.g.):
		Mood	First relapse is detected
		Craving	Heavy drinking episode detected (e.g., BAC $\geq$ 80 mg%)
		Cognitive performance	Drinking session ended
		Biosensor system compliance	Following periods of extended abstinence
		Medication compliance	High-risk drinking circumstances detected
		Location	Physiological craving indicators detected
		Social circumstances	Improved contact with patients
		Subjective alcohol effects ^b	Reduced burden on patients and clinicians/research staff
		Objective alcohol effects ^c	Remote nature of data collection logistically advantageous for multiple treatment sites.

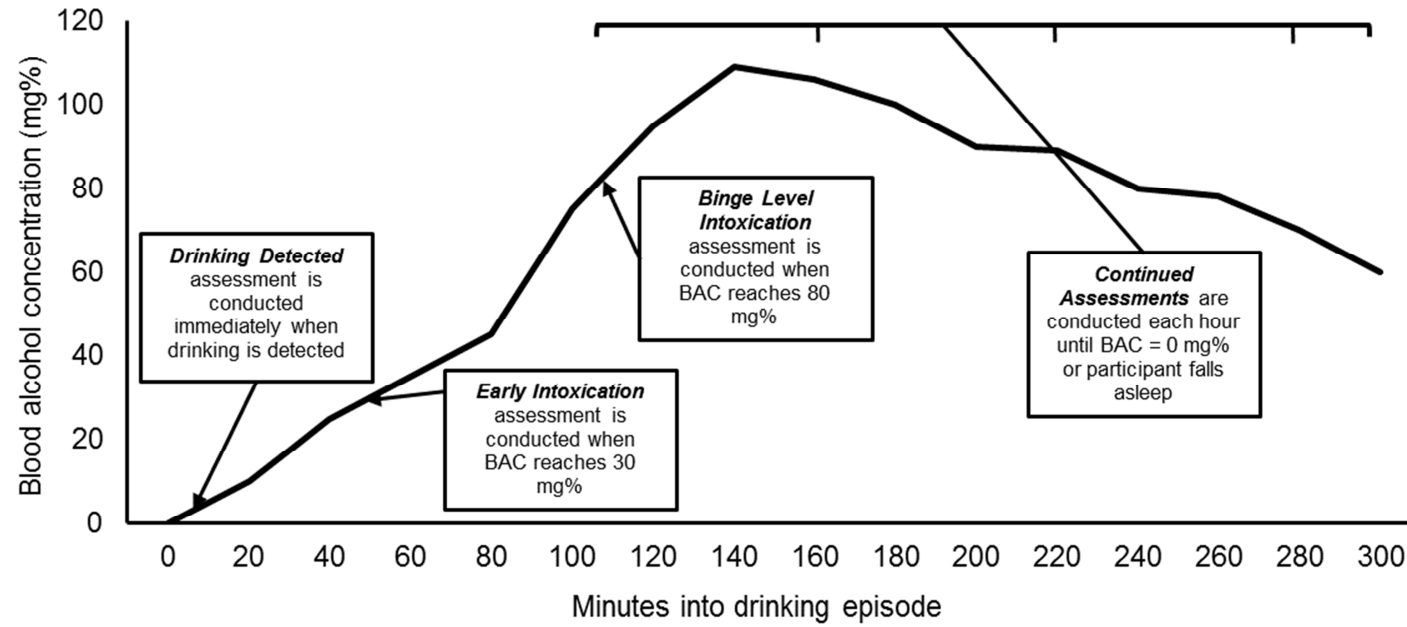
Notes. BAC estimates reported in the primary outcomes column are estimated from smartphone and TAS data. BAC-related outcomes are aspirational in nature and will require additional development before these applications of the platform can be realized.

^aBiosensors required to collect the suggested outcome variables include transdermal alcohol sensor, Lumme smartwatch, and a smartphone capable of conducting prompted assessments. ^bSubjective alcohol effects include self-reported subjective stimulation, sedation, and intoxication at key points during ascending and descending limbs of the BAC curve (see Fig 3). ^cObjective alcohol effects include changes in heart rate and cognitive functioning.









- Identifying new medications to treat harmful alcohol use is a priority
- Mobile alcohol biosensors will improve medication development research
- A mobile biosensor-based paradigm can effectively screen new medications
- Biosensor-based endpoints can improve clinical trials