A Wearable Alcohol Biosensor: Exploring the Accuracy of Transdermal Drinking Detection

Catharine E. Fairbairn, Ph.D.[a]*, Jiaxu Han, M.A.[a], Eddie P. Caumiant, M.S.[a], Aaron S. Benjamin, Ph.D.[a] and Nigel Bosch, Ph.D.[b,c]

[a] Department of Psychology, University of Illinois Urbana-Champaign; Champaign, IL 61820, USA

[b] School of Information Sciences, University of Illinois Urbana-Champaign; Champaign, IL 61820, USA

[c] Department of Educational Psychology, University of Illinois Urbana-Champaign; Champaign, IL 61820, USA

* Correspondence concerning this article should be addressed to Catharine Fairbairn, Ph.D., Department of Psychology, 603 East Daniel St., Champaign, IL 61820. Electronic mail: cfairbai@illinois.edu.

Abstract

Background: Trace amounts of consumed alcohol are detectable within sweat and insensible perspiration. However, the relationship between ingested and transdermally emitted alcohol is complex, varying across environmental conditions and involving a degree of lag. As such, the feasibility of real-time drinking detection across diverse environments has been unclear. In the current research we revisit sensor performance using new tools, exploring the accuracy of a new generation of rapid-sampling transdermal biosensor for contemporaneous drinking detection across diverse environments via machine learning.

Methods: Regular drinkers (N = 100) attended three laboratory sessions involving the experimental manipulation of alcohol dose, rate of consumption, and environmental dosing conditions. Participants further supplied breath alcohol concentration (*BAC)* readings in the field over 14 days. Participants wore compact wrist sensors capable of rapid sampling (20 sec intervals). Transdermal sensor data was translated into alcohol use estimates using machine learning, integrating only transdermal data collected prior to the point of *BAC* assessment.

Results: A total of 5.39 million transdermal readings (28,615 hours) and 12,699 *BAC* readings were collected for this research. Models indicated strong transdermal sensor accuracy for real-time drinking detection across both laboratory and field contexts (*AUROC*, 0.966, *95% CI*, 0.956-0.972; Sensitivity, 89.8%; Specificity, 90.6%). Models aimed at differentiating high-risk (≥0.08%) drinking levels yielded intermediate (*AUROC*, 0.738; *95% CI*, 0.698-0.777; only drinking episodes) to strong (*AUROC*, 0.941, *95% CI*, 0.929-0.954; all data) accuracy levels.

Conclusions: Results indicate a range of useful future applications for transdermal alcohol sensors including long-term health tracking, medical monitoring, and just-in-time relapse prevention.

Keywords: Alcohol, substance use, wearables, biosensor, transdermal

1. Introduction

The potential public health impact of a wearable alcohol monitor is large and growing (Fairbairn and Bosch, 2021). Deaths from alcohol have climbed steadily across the past two decades, surging an estimated 29% in the years since COVID-19 (Esser et al., 2024; Spencer et al., 2022). Simultaneously, adolescent and young adult drinkers show reduced binge drinking frequency compared to older cohorts (Substance Abuse and Mental Health Services Administration, 2022) and, with the increasing ubiquity of mocktails and dry-bars (Bowdring et al., 2024; World Health Organization, 2023), these populations demonstrate heightened awareness of health consequences of alcohol use (Gallup Inc, 2023). Across age cohorts, engagement with digital health monitoring technologies has burgeoned (Pew Research Center, 2024). The health impact of alcohol looms large, while unprecedented potential exists for intervention through technology-mediated care (Esser et al., 2024; Steinhubl et al., 2013).

Monitoring of active alcohol intake forms the backbone of alcohol use intervention (Epstein and McCrady, 1998; Miller, 1978; Miller and Rollnick, 2012), but the identification of effective monitoring methods has represented a formidable challenge (Swift, 2003). Alcohol exerts direct pharmacological effects on attentional, memory, and motivational resources required for self-monitoring (Fairbairn et al., 2021; Weissenborn and Duka, 2003; White, 2003), and knowledge of societal stigma surrounding alcohol use can bias users' reports (Davis et al., 2010). For individuals seeking to abstain or moderate alcohol use, motivation to engage in active alcohol use monitoring naturally fluctuates over time and across divergent contexts (DiClemente and Prochaska, 1998; Marlatt, 1996; Miller and Rollnick, 2012). Wearable sensors have been identified as representing a solution to specific

challenges of alcohol self-monitoring (Barnett, 2015; NIAAA, 2015; Swift, 1993). While traditional monitoring methods require motivated action at the time of each alcohol use assessment, wearable sensors permit the disjuncture of the decision to engage in active monitoring from the drinking context itself, limiting demands on the drinker at times when cognitive and motivational resources are likely to be constrained (Barnett, 2015). Wearable sensors have the potential to provide passive and continuous monitoring of alcohol use, requiring minimal active engagement on the part of the drinker (Swift, 1993). These monitors might have far-ranging application, including the provision of real-time support among individuals in recovery (Barnett, 2015; Nahum-Shani et al., 2017), health monitoring by providers treating patients in medical contexts (Burton and Sheron, 2018), and prevention-focused daily health tracking among broad populations of drinkers (Takacs et al., 2014).

Transdermal sensors currently represent the most widely-researched technology for wearable alcohol use assessment (van Egmond et al., 2020; Yu et al., 2022). Approximately 1% of consumed alcohol is emitted through the skin in the form of sweat and insensible perspiration (Nyman and Palmlöv, 1936; Swift and Swette, 1992). Thus, it is possible to assess the concentration of alcohol emitted via water vapor from the stratum corneum using a device that integrates basic fuel-cell technology and rests on the skin's surface (Swift, 1993).

Transdermal Alcohol Concentration (*TAC)* reflects the combined effects of passive alcohol diffusion within blood via skin capillaries (insensible perspiration) as well as active perspiration from sweat gland secretions (sensible perspiration), fluids that exhibit differential alcohol concentration and lag times to alcohol excretion (Anderson and Hlastala, 2006). As such, output from transdermal alcohol sensors is complex, and based on prior research, it has been unclear whether contemporaneous transdermal detection of alcohol

consumption might be achieved with acceptable accuracy and temporal specificity for widespread application (Anderson and Hlastala, 2006; Luczak and Ramchandani, 2019; Marques and McKnight, 2009). More specifically, due to lag times between ingested and transdermally emitted alcohol, the capability of transdermal devices for real-time or near real-time detection of alcohol consumption is currently unknown (Anderson and Hlastala, 2006; Marques and McKnight, 2009). Further, a variety of environmental factors can confound the relationship between ingested and transdermally emitted alcohol, including perspiration rate, variability in skin-sensor distance, and ambient alcohol, and the accuracy of new-generation transdermal sensors outside invariant/sterile lab settings is currently unclear (Fairbairn and Bosch, 2021; Luczak and Ramchandani, 2019). Therefore, perspiration-based alcohol sensors have been siloed for niche application as retroactive abstinence monitors within the criminal justice system (Alcohol Monitoring Services, 2018), where real-time data is non-essential and moderate accuracy can support applications that maximize specificity at a cost to sensitivity (van Egmond et al., 2020). In contrast, broader populations of drinkers as yet lack a wearable alcohol biosensor.

Yet the limitations of transdermal assessment more broadly have been challenging to separate from the limitations associated with the technology and analytic tools available to previous generations of researchers (Fairbairn and Bosch, 2021; Yu et al., 2022). Prior research has predominantly examined output from transdermal bracelets featuring a bulky ankle-worn design and a relatively sparse sampling interval (30-minutes; Yu et al., 2022). This work has been further characterized by small datasets (average $N = 17$, see Yu et al., 2022), subjective reports of drinking as "ground truth," and laboratory-based testing conditions (see Yu et al., 2022 for a review). Within this work, considerable delays between

ingested and transdermally diffused alcohol have been evident (e.g., Marques and McKnight, 2009), interference posed by environmental factors has been difficult to analytically remove (Fairbairn and Bosch, 2021; Gunn et al., 2023; Luczak and Ramchandani, 2019), and even retrospective accuracy metrics have often emerged as moderate in magnitude (Ash et al., 2022; Kianersi et al., 2023; Richards et al., 2023; Croff et al., 2021; although see Didier et al., 2023).

Here we report results of a large-scale validation study of a wearable alcohol biosensor integrating real-time objective alcohol use assessment and variable environmental testing conditions. Recent decades have seen advances in both hardware and computational tools with the potential to advance transdermal alcohol measurement, including the introduction of a new generation of compact, rapid-sampling alcohol biosensor (NIAAA, 2015; Wang et al., 2021), and analytic approaches capable of analysis and forecasting based on complex time-series trends (Fairbairn et al., 2020; Fairbairn and Bosch, 2021). In the current study we assess the accuracy of a transdermal wrist sensor against objective alcohol use data collected among participants examined across variable environmental conditions. Transdermal data is translated into "real time" alcohol use estimates using machine learning algorithms aimed at addressing lags between ingested and transdermally emitted alcohol. More specifically, aims of the current project were as follows: 1) To assess transdermal alcohol sensor accuracy in predicting drinking vs. sobriety in real-time across variable environmental conditions; 2) To assess moderators of transdermal drinking detection accuracy, including person-level and within-person factors theorized to impact transdermal sensor output; and 3) To assess transdermal sensor accuracy for detecting binge or "high-risk" (≥0.08% *BAC*) drinking (National Institute on Alcohol Abuse and Alcoholism., 2004).

2. Materials and Methods

2.1. Participants

Participants were recruited via social media advertisements and posted notices in the local community. Individuals were excluded if they were especially light or infrequent drinkers (≤ 1 drinking day/week), reported a history of medical conditions for which alcohol consumption would be contraindicated, reported taking medications or other drugs with the potential to interact with alcohol, or indicated a history of adverse reactions to the type or amount of beverage administered in the study (see National Advisory Council on Alcohol Abuse and Alcoholism, 1989). Individuals actively seeking treatment for alcohol problems, those with a history of severe psychiatric illness, non-English speakers, individuals with medical conditions contraindicating moderate aerobic exertion, and women who reported being pregnant or trying to become pregnant were also excluded. Participants consisted of 100 regular drinkers aged 21+ (M_{age} = 24.20; *SD* = 4.36; *Range*=21-46). Participants were 50% female, and 41% White. Participants reported drinking an average of 9.39 (*SD* = 5.06) and binge drinking 4.18 (*SD* = 3.87) days/month (see Table 1).

2.2. Procedure

2.2.1. Study Overview and Design: Study procedures took place over the course of 14 days and employed a hybrid laboratory-ambulatory design integrating observations drawn from both precisely controlled (laboratory) as well as ecologically valid (ambulatory) consumption contexts. Participants attended three experimental laboratory visits, held at 1-week intervals over the course of the study (Figure 1). Laboratory visits also served as ambulatory orientation, check-in, and close-out visits. A detailed description of study methods is provided in Supplemental Materials.

2.2.2. Laboratory Procedures: The laboratory study arm permitted the acquisition of precise data surrounding quantity and timing of alcohol ingestion while also facilitating systematic manipulation of environmental factors theorized to impact transdermal sensor readings, so ensuring adequate variability along these factors in the broader dataset (Anderson and Hlastala, 2006; Luczak and Ramchandani, 2019). Breath alcohol concentration (*BAC*) readings were taken at baseline and at 10-minute intervals following beverage administration. All participants received low (.03%), moderate (.06%), and high (.09% target *BAC*) alcohol doses across the three laboratory sessions (Figure 2). To simulate environmental factors with the potential to confound *TAC* readings, laboratory sessions integrated an aerobic bicycle workout to induce perspiration (mild/8-minutes; moderate/20-minutes), timed arm/wrist movement to induce variability in skin-sensor distance (two 90-second bursts), and exposure to alcohol-containing products (e.g., hand sanitizer; See Figures 1-2 and Supplemental Methods). Finally, to simulate variable consumption patterns, approximately equal numbers of participants were randomized to consume study beverages at a relatively rapid (.01% *BAC*/3-minutes) or slow (.01%/6-minutes) pace (see Table 1). The transdermal sensor employed in this study was the BACtrack Skyn, a compact, wrist-worn device (37g; 4.7x2.5x0.6cm) that links with a smartphone via Bluetooth and features passive *TAC*, motion, and temperature sensing (20-second sampling interval; Fairbairn and Bosch, 2021; NIAAA, 2015). Devices employed in this research were originally shipped from the manufacturer 9/2021-5/2023. See also Supplemental Methods and Figure S1.

2.2.3. Ambulatory Procedures: Participation involved a 14-day intensive assessment period aimed at capturing transdermal device performance in real-world drinking environments. Participants were instructed to wear the transdermal device throughout this period except for

times when they were bathing or charging the device. During ambulatory assessment, participants provided prompted *BAC* readings via smartphone-connected breathalyzers (Ariss et al., 2023) in response to random smartphone prompts 4-6 times/day, as well as at 30-minute intervals during drinking episodes. To ensure high-quality ambulatory *BAC* readings, participants underwent laboratory training in mouth alcohol effects and received mid-study feedback and, ultimately, compensation commensurate with the validity and completeness of their ambulatory data. At the final study visit, after the completion of ambulatory assessment, participants reported on transdermal device acceptability, including their level of social and physical comfort in wearing the bracelet (1= extremely comfortable; 9=extremely uncomfortable) and their willingness to wear the bracelet beyond the end of the study (yes/no; see Supplemental Methods). Participants received $200 for attending all three laboratory visits and wearing the transdermal device throughout ambulatory assessment, and an additional $100 for responding to at least 70% of all ambulatory prompts.

2.3. Data Processing and Analysis

All data required for the replication of results in this article, together with code permitting researchers to use the transdermal translation algorithms developed here with their own data sets, are provided here: https://osf.io/bdthf/?view_only=3554401322674c3ab33cecff2b7c27c9. Time-series features extracted from transdermal sensor output (e.g., *TAC* rise rate, quantiles) were entered into Extra-Trees machine learning algorithms (Geurts et al., 2006). To produce a model that might be applied for *real-time BAC* estimation, models included only transdermal sensor data *preceding* (not following) *BAC* readings (Fairbairn et al., 2020). We used 5-fold, participant-independent cross-validation to ensure that predictions were not over-fit to specific data points or participants. Model performance was evaluated on testing sets using area under the

receiver operating characteristic curve (*AUROC*) calculated using nonparametric models accounting for the clustering of observations within participants. Moderators of device accuracy were examined at the within- (device wear time, device ship date) and the between- (sex, age, race, drinking history) participant level, selected as factors theorized to impact transdermal sensor output (Ash et al., 2022; Fairbairn and Kang, 2020; Luczak and Ramchandani, 2019). Diagnostic thresholds were determined using the maximum Youden index, and sensitivity and specificity values were calculated using bootstrapping. Drinking episode start time was estimated at the timestamp of the first positive *BAC* reading. Finally, for models differentiating no (0.00%), low-risk (>0.00%, <0.08%), and high-risk (≥0.08%) drinking levels,[1] predictions were jointly estimated in a single model, and resulting test statistics were averaged for models predicting each risk category to yield an omnibus *AUROC* value. See Supplemental Materials for the complete data analysis plan and details of data processing.

3. Results

3.1. Descriptives, Compliance, and Acceptability

3.1.1. BAC and Alcohol Use Descriptives: A total of 12,699 unique *BAC* readings (6,349 from field contexts; 8,054 >0.00%) were collected for the purposes of the current study. Regarding ambulatory data, participants on average provided breathalyzer readings in response to 69% of prompts. Each participant provided an average of 63.49 (*SD* = 31.28) *BAC* readings across the 14-day assessment interval, engaging in 7.41 drinking episodes lasting on average 148.2 minutes (*SD* = 126.7) in duration. Peak *BAC* levels in the field ranged from 0.01%-0.27%

[1] For the purposes of the current research, a high-risk drinking episode is operationalized as one where BAC meets or exceeds NIAAA's binge threshold of .08% (National Institute on Alcohol Abuse and Alcoholism., 2004). BAC levels that do not meet these criteria are designated low-risk events. We adopt these labels for ease of reference and to reflect research documenting adverse drinking consequences linked with drinking episodes that meet or exceed the binge drinking threshold (e.g., Jones et al., 2018). Note, however, that episodes that are truly "low" vs "high" risk in terms of their potential for negative effects will inevitably depend on context and individual.

(*M* = 0.095%, *SD* = 0.054). Within the laboratory study arm, achieved peak *BAC*s on low, moderate, and high-dose sessions were 0.032% (*SD* = 0.011), 0.057% (*SD* = 0.012), and 0.084% (*SD* = 0.015) respectively. Additional *BAC* descriptives are provided in Figure 3.

3.1.2. Transdermal Compliance and User Acceptability: A total of 5.39 million transdermal readings (28,615 hours of *TAC* data) were collected across the course of the study. Regarding compliance, application of wear-detection thresholds (<26° Celsius) indicated that participants wore transdermal devices for the majority (*Mdn* = 92.79%; *SD* = 15.54) of the ambulatory assessment period (*Mdn* = 12.99 days estimated wear time/participant). Participants rated transdermal devices as moderately physically comfortable (*M* = 4.72, *SD* = 1.91) and highly socially comfortable (*M* = 2.84, *SD* = 2.08). Of participants, 80.4% indicated they would be willing to wear the transdermal device beyond the end of the study (see Supplemental Measures).

3.1.3. Transdermal Devices and Device Failures: A total of *N* = 38 transdermal devices were employed across all study participants. During the course of the study, devices were returned to the manufacturer and replaced in response to anomalies observed in the data (*n* = 5) and/or larger hardware updates from the manufacturer (*n* = 20). Median usage time for each transdermal device employed in the study was 666.83 hours (*SD* = 392.51; *Range* 130.57-1707.73). Gaps in transdermal recording occurred attributable to both protocol consistent and non-consistent causes, including when devices were turned off for charging (*freq*: ~4-6 days) as well as due to device malfunction and user non-compliance (e.g., devices not charged as instructed). Across 14 days of assessment, the median total duration of recording gaps per participant was 19.31 hours (*SD* = 52.61; *Range* 0-311), with gaps lasting a median of 3.19 hours (*SD* = 26.71; *Range* 0-121).

3.2. Transdermal Sensor Accuracy

3.2.1. Accuracy in Detecting Drinking vs. Sobriety: Analyses assessed accuracy for transdermal alcohol monitors in predicting drinking (>0.00% *BAC*) vs. sobriety (0.00% *BAC*) in real time. We first explored transdermal sensor accuracy in analyses including all data collected across study contexts, including laboratory and field settings. The accuracy of transdermal sensors in distinguishing episodes of alcohol consumption emerged as excellent (*AUROC*, 0.966, *95% CI*, 0.956-0.972; Figure 4A). Analyses yielded strong sensitivity and specificity, indicating the algorithm capable of correctly detecting 89.8% (*95% CI*, 88.6%-90.8%) of true drinking and 90.6% (*95% CI*, 89.5%-91.7%) of true sober moments. Regarding the time course of detection, approximately 70% of drinking episodes were correctly identified for readings provided 0-30 minutes after the first positive *BAC* value, 92.2% for minutes 30-120, and 93.6% for minutes 120+. When data from field (i.e., real world) contexts was examined independent of laboratory data, accuracy metrics were also strong (*AUROC*, 0.941, *95% CI*, 0.927-0.955; Sensitivity 84.2%, *95% CI*, 83.1%-87.8%; Specificity 88.8%, *95% CI*, 86.3%-89.8%; Figure 4B).

3.2.2. Moderators of Detection Accuracy: We next explored moderators of transdermal sensor accuracy in detecting drinking vs. sobriety for data collected across study contexts (Figure 4D-I). Accuracy decreased with transdermal device usage as days of wear progressed within participants, with *AUROC* values decreasing from 0.976 (day 0-5) to 0.961 (days 5-10) to 0.944 (days 10+; see Figure 5). No significant differences emerged with respect to participant sex, race, age, drinking patterns, and transdermal device age (i.e., ship date; see Figure 5).

3.2.3. Accuracy in Determining Drinking Risk Category: Models next examined drinking according to three risk categories: high-risk drinking, low-risk drinking, and sobriety. Results yielded an averaged omnibus *AUROC* value of 0.957. Models differentiating high-risk drinking

from a combined comparison group of low-risk drinking and sobriety indicated excellent accuracy (*AUROC*, 0.941, *95% CI*, 0.929-0.954; Figure 4C), correctly identifying 89.3% (*95% CI*, 85.4%-92.8%) of true instances of high-risk drinking, and 86.4% (*95% CI*, 82.2%-88.1%) of instances of sobriety/low risk drinking. Models differentiating between low- and high-risk drinking levels while *excluding* 0.00% *BAC* values yielded moderate accuracy (*AUROC*, 0.738, *95% CI*, 0.698-0.777; Sensitivity 58.8%, *95% CI*, 50.6%-76.9%; Specificity 77.3%, *95% CI*, 58.4%-84.1%).

4. Discussion

Wearable sensors are unique in that they provide objective alcohol use data while placing minimal demands on the drinker, so lifting engagement burden at times when cognitive and motivational resources are likely to be scarce. The current research indicates high accuracy for contemporaneous alcohol use detection via wearable alcohol biosensor across variable environmental conditions. Rates of detecting true consumption and differentiating from non-consumption ranged from 84-91%. Models isolating drinking risk levels specifically within drinking episodes (>0.00% *BAC*) produced intermediate accuracy levels for distinguishing low- from high-risk drinking. Accuracy was strong for models differentiating high-risk drinking (≥0.08% *BAC*) from a combined comparison group of non-drinking and low-risk drinking. The new-generation sensors employed in this research were rated as acceptable by participants for longer-term wear.

Output yielded by perspiration-based alcohol sensors is complex (Anderson and Hlastala, 2006). On the basis of prior work, it has been unclear the extent to which transdermal estimates of drinking could be achieved with adequate accuracy and temporal specificity to permit widespread application (Fairbairn and Kang, 2020; Yu et al., 2022).

Here we revisit transdermal sensor accuracy using machine learning methods applied to dense time-series data from new-generation, rapid-sampling sensors. Time-series represent a key tool for addressing problems of prediction lags, containing information on over-time trends and so potentially indicating not only where a signal currently *is* but also where it is *going* (e.g., time-series forecasting techniques; Box et al., 1994; Fairbairn et al., 2020). Further, as environmental confounds can yield over-time patterns that diverge from those linked to true alcohol ingestion, time-series can further be leveraged to parse *TAC* signal from noise. For example, a sudden spike in *TAC* is more likely indicative of environmental confound than is a gradual rise (Fairbairn and Bosch, 2021). Yet in examining individual, decontextualized *TAC* readings such over-time information is lost. Here we translate transdermal data using machine learning algorithms capable of capturing complex, non-linear patterns in time-series data (Geurts et al., 2006). Algorithms are trained using a large sample of precise *BAC* readings collected across diverse environmental conditions, producing models that display strong accuracy levels for transdermal sensors in contemporaneous alcohol use detection.

In the future, as larger transdermal datasets accrue, application of data-hungry methods such as transformer neural networks might facilitate transdermal measurement of still more detailed drinking metrices, including precise *BAC* values, while also providing increased ability to discriminate low- from high-risk drinking. Even in the absence of such fine-grained measures, however, a range of potential applications exist for a compact sensor capable of passive, objective drinking and sobriety detection, especially in light of links between sensor output and alcohol-related consequences (Russell et al., 2022). For example, a preventative health tracker capable of keeping a long-term record of drinking and high-risk drinking days could increase

health awareness among broad populations of drinkers (Takacs et al., 2014). A medical monitor might serve to provide an objective record of consumption for healthcare providers in treating patients with conditions requiring the cessation or moderation of alcohol use (e.g., cardiovascular disorders, diabetes; Burton and Sheron, 2018; Howard et al., 2004). Finally, a sensor with the potential for contemporaneous drinking detection might provide a range of supports for individuals with alcohol use disorders, including by triggering real-time intervention in response to relapse (Nahum-Shani et al., 2017).

Gaps in transdermal device recording tended to be brief and were generally consistent with protocol-compliant behaviors on the part of participants (e.g., device powered off for charging). These results stand in contrast to findings for early new-generation sensor prototypes, where sensor failure was relatively frequent (Fairbairn and Kang, 2019). Potential for sensor degradation exists for fuel-cell based transdermal devices due to humidity buildup within the gas chamber between the sensor and skin (Ash et al., 2022; van Egmond et al., 2020). Evidence of some sensor degradation over the course of continuous wear was evident in the current sample. For the 2-week time interval captured here, however, accuracy rates were consistently strong. It is notable that, in the current study, device wear time (i.e., days of wear by a given study participant), but not overall device age (i.e., device ship date), moderated sensor performance. One possibility is that the sensor aeration and cleaning afforded by periods of non-wear between participants served to interrupt humidity buildup and refresh sensors, and that such breaks integrated into future research and applications might improve sensor performance. Related, devices in this study were replaced in response to data anomalies, presenting circumstances potentially favorable to device performance. The extent to which such degradation would emerge as problematic over longer time periods under typical day-to-day use conditions requires further

investigation, as does the extent to which performance degradation is permanent vs. might rebound following a period of rest, and whether or not technological advances in sensor manufacturing can eliminate these losses.

Regarding the time course of drinking detection, sensors identified most (70%) cases of alcohol consumption for data points within 30 minutes of the time of first positive breathalyzer reading, increasing to 92% by minutes 30-120. In the current investigation, transdermal device accuracy was examined only at times of active *BAC* assessment. Additional research is needed to chart the precise time course from the point of drinking initiation to first transdermal alcohol use detection. Such research would be important prior to the initiation of more time-sensitive transdermal device applications wherein unanticipated detection delays might result in harm to the drinker or others (e.g., real-time driving advisability feedback; Fairbairn and Bosch, 2021).

TAC dynamics have been theorized to vary substantially across individuals due to variability in the physical properties of the stratum corneum, including differences in skin thickness, hydration, and permeability (Anderson and Hlastala, 2006; Luczak and Rosen, 2014). In the current study no significant differences in transdermal detection accuracy emerged according to sex, race, age, or prior drinking history. Nonetheless, although the sample for the current study was large in comparison to many prior studies (Yu et al., 2022), sample size requirements for between-subject moderator analyses are high. Additionally, although no individual-level differences reached significance, yet some non-significant group effects (e.g., race) might warrant examination in larger samples. Future research might employ oversampling techniques to increase power for detecting variability across individuals and ensure the development of transdermal alcohol detection algorithms generalizable across populations.

Additional limitations and future directions should be noted. The current study comprises, to our knowledge, the largest database of objectively-assessed high-risk drinking episodes (>.08%) captured in transdermal sensor research to date (see Yu et al., 2022). But high-risk episodes were still relatively scarce in this sample (n = 618; Figure 4). Future research should revisit the question of low- vs. high-risk drinking differentiation in larger datasets. Related, in the current study we operationalized drinking risk level according to NIAAA's binge drinking threshold (National Institute on Alcohol Abuse and Alcoholism., 2004), yet alternative thresholds exist for the determination of high- vs. low-risk drinking. Future research might explore the accuracy of transdermal sensors for detecting drinking episodes both above and below the .08% threshold. Further, although initial data was presented, a thorough examination of user acceptability was beyond the scope of the current study (Ash et al., 2022; Rosenberg et al., 2023). Additional analyses, including qualitative examinations and models exploring individual differences, are needed to understand user experience with these sensors. Finally, although our dataset explored the accuracy of transdermal sensors in the context of a dataset that integrated diverse contextual conditions, an examination of context-level moderators was beyond the scope of this initial investigation. Future research is needed to examine context-level moderators of transdermal sensor accuracy, including controlled testing environments, physical exertion, skin-sensor distance, and ambient alcohol.

In sum, for nearly a century, researchers have known that ingested alcohol can be detected within perspiration (Nyman and Palmlöv, 1936). Alcohol's journey to the skin's surface is complex, dependent on activity of the digestive, circulatory, endocrine, and integumentary systems, spurring questions surrounding the viability of contemporaneous perspiration-based drinking detection (Anderson and Hlastala, 2006). Additional research is needed to explore the

potential of perspiration-based sensors for detecting more fine-grained drinking metrics, including precise BAC and/or standard drink estimation. Results of the current study nonetheless lay the foundation for wearable alcohol sensors with applications across medical, intervention, prevention, and research domains.

Acknowledgements

Our thanks to the students and staff of the Alcohol Research Laboratory at the University of Illinois at Urbana-Champaign. Thanks also to Nancy Barnett and Vijay Ramchandani for support with funding acquisition and feedback on the parent project. Finally, thanks to Laura Gurrieri, Talia Ariss, Rosalie Ruhlmann, Hieu Nguyen, and Yang Lu for aid in the conduct of this research.

References

Alcohol Monitoring Services, 2018. About SCRAM systems.

Anderson, J.C., Hlastala, M.P., 2006. The kinetics of transdermal ethanol exchange. Journal of Applied Physiology 100, 649–655.

Ariss, T., Fairbairn, C.E., Bosch, N., 2023. Examining new-generation transdermal alcohol biosensor performance across laboratory and field contexts. Alcoholism: Clinical & Experimental Research.

Ash, G.I., Gueorguieva, R., Barnett, N.P., Wang, W., Robledo, D.S., DeMartini, K.S., Pittman, B., Redeker, N.S., O'Malley, S.S., Fucito, L.M., 2022. Sensitivity, specificity, and tolerability of the BACTrack Skyn compared to other alcohol monitoring approaches among young adults in a field-based setting. Alcoholism: Clinical and Experimental Research 46, 783–796. https://doi.org/10.1111/acer.14804

Barnett, N.P., 2015. Alcohol sensors and their potential for improving clinical care. Addiction 110, 1–3.

Bowdring, M.A., McCarthy, D.M., Fairbairn, C.E., Prochaska, J.J., 2024. Non-alcoholic beverage consumption among US adults who consume alcohol. Addiction 119, 1080–1089. https://doi.org/10.1111/add.16452

Box, G.E.P., Jenkins, G.M., Reinsel, G.C., 1994. Time series analysis: Forecasting and control. Prentice-Hall, Englewood Cliffs, N.J.

Burton, R., Sheron, N., 2018. No level of alcohol consumption improves health. The Lancet 392, 987–988.

Croff, J.M., Hartwell, M.L., Chiaf, A.L., Crockett, E.K., Washburn, I.J., 2021. Feasibility and reliability of continuously monitoring alcohol use among female adolescents and young adults. Drug and Alcohol Review 40, 1143–1154. https://doi.org/10.1111/dar.13045

Davis, C.G., Thake, J., Vilhena, N., 2010. Social desirability biases in self-reported alcohol consumption and harms. Addictive Behaviors 35, 302–311.

DiClemente, C.C., Prochaska, J.O., 1998. Toward a comprehensive, transtheoretical model of change: Stages of change and addictive behaviors., in: Miller, W.R., Heather, N. (Eds.), Treating Addictive Behaviors. Plenum Press, New York, NY, pp. 3–24.

Didier, N.A., King, A.C., Polley, E.C., Fridberg, D.J., 2023. Signal processing and machine learning with transdermal alcohol concentration to predict natural environment alcohol consumption. Experimental and Clinical Psychopharmacology 32, 245–254. https://doi.org/10.1037/pha0000683

Epstein, E.E., McCrady, B.S., 1998. Behavioral couples treatment of alcohol and drug use disorders: Current status and innovations. Clinical Psychology Review 18, 689–711. https://doi.org/10.1016/S0272-7358(98)00025-7

Esser, M.B., Sherk, A., Liu, Y., Naimi, T.S., 2024. Deaths from excessive alcohol use—United States, 2016–2021. Centers for Disease Control and Prevention Morbidity and Mortality Weekly Report 73, 154–161. https://doi.org/10.15585/mmwr. mm7308a1

Fairbairn, C.E., Bosch, N., 2021. A new generation of transdermal alcohol biosensing technology: Practical applications, machine learning analytics, and questions for future research. Addiction 116, 2912–2920.

Fairbairn, C.E., Kang, D., 2020. Transdermal alcohol monitors: Research, applications, and future directions, in: Frings, D., Albery, I. (Eds.), The Handbook of Alcohol Use and Abuse. Elsevier.

Fairbairn, C.E., Kang, D., 2019. Temporal dynamics of transdermal alcohol concentration measured via new-generation wrist-worn biosensor. Alcoholism: Clinical and Experimental Research 43, 2060–2069. https://doi.org/10.1111/acer.14172

Fairbairn, C.E., Kang, D., Bosch, N., 2020. Using machine learning for real-time BAC estimation from a new-generation transdermal biosensor in the laboratory. Drug and Alcohol Dependence 216, 108205. https://doi.org/10.1016/j.drugalcdep.2020.108205

Fairbairn, C.E., Kang, D., Federmeier, K.D., 2021. Alcohol and neural dynamics: A meta-analysis of acute alcohol effects on event-related brain potentials. Biological Psychiatry 89, 990–1000.

Gallup Inc, 2023. Gallup Consumption Habits Survey.

Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. Machine Learning 63, 3–42.

Gunn, R.L., Merrill, J.E., Haines, A.M., Fernandez, M.E., Souza, T., Berey, B.L., Leeman, R.F., Wang, Y., Barnett, N.P., 2023. Use of the BACtrack Skyn alcohol biosensor: Practical applications for data collection and analysis. Addiction 118, 1586–1595. https://doi.org/10.1111/add.16207

Howard, A.A., Arnsten, J.H., Gourevitch, M.N., 2004. Effect of alcohol consumption on diabetes mellitus: A systematic review. Annals of Internal Medicine 140, 211–219.

Jones, S.A., Lueras, J.M., Nagel, B.J., 2018. Effects of binge drinking on the developing brain: studies in humans. Alcohol research: current reviews 39, 87.

Kianersi, S., Ludema, C., Agley, J., Ahn, Y., Parker, M., Ideker, S., Rosenberg, M., 2023. Development and validation of a model for measuring alcohol consumption from transdermal alcohol content data among college students. Addiction 118, 2014–2025. https://doi.org/10.1111/add.16228

Luczak, S.E., Ramchandani, V.A., 2019. Special issue on alcohol biosensors: Development, use, and state of the field. Alcohol 81, 161–165.

Luczak, S.E., Rosen, I.G., 2014. Estimating BrAC from transdermal alcohol concentration data using the BrAC estimator software program. Alcoholism: Clinical and Experimental Research 38, 2243–2252. https://doi.org/10.1111/acer.12478

Marlatt, G.A., 1996. Taxonomy of high-risk situations for alcohol relapse: Evolution and development of a cognitive-behavioral model. Addiction 91, 37–50. https://doi.org/10.1046/j.1360-0443.91.12s1.15.x

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

Miller, W.R., 1978. Behavioral treatment of problem drinkers: A comparative outcome study of three controlled drinking therapies. Journal of Consulting and Clinical Psychology 46, 74.

Miller, W.R., Rollnick, S., 2012. Motivational interviewing: Helping people change. Guilford Press, New York.

Nahum-Shani, I., Smith, S.N., Spring, B.J., Collins, L.M., Witkiewitz, K., Tewari, A., Murphy, S.A., 2017. Just-in-time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support. Annals of Behavioral Medicine 52, 446–462.

National Advisory Council on Alcohol Abuse and Alcoholism, 1989. Recommended Council Guidelines on Ethyl Alcohol Administration on Human Experimentation. National Institute on Alcohol Abuse and Alcoholism, Bethesda, MD.

National Institute on Alcohol Abuse and Alcoholism., 2004. NIAAA council approves definition of binge drinking (No. NIH Publication No. 04-5346). NIAAA Office of Research Translation and Communications., Bethesda, MD.

NIAAA, 2015. Wearable alcohol biosensor challenge [WWW Document]. URL https://www.nih.gov/news-events/news-releases/niaaa-selects-winners-its-wearable-alcohol-biosensor-challenge

Nyman, E., Palmlöv, A., 1936. The elimination of ethyl alcohol in sweat. Skandinavisches Archiv Für Physiologie 74, 155–159. https://doi.org/10.1111/j.1748-1716.1936.tb01150.x

Pew Research Center, 2024. Mobile Fact Sheet. Washington, D.C.

Richards, V.L., Barnett, N.P., Cook, R.L., Leeman, R.F., Souza, T., Case, S., Prins, C., Cook, C., Wang, Y., 2023. Correspondence between alcohol use measured by a wrist-worn alcohol biosensor and self-report via ecological momentary assessment over a 2-week period. Alcohol: Clinical and Experimental Research 47, 308–318. https://doi.org/10.1111/acer.14995

Rosenberg, M., Kianersi, S., Luetke, M., Jozkowski, K., Guerra-Reyes, L., Shih, P.C., Finn, P., Ludema, C., 2023. Wearable alcohol monitors for alcohol use data collection among college students: feasibility and acceptability. Alcohol 111, 75–83.

Russell, M.A., Turrisi, R.J., Smyth, J.M., 2022. Transdermal sensor features correlate with ecological momentary assessment drinking reports and predict alcohol-related consequences in young adults' natural settings. Alcoholism: Clinical and Experimental Research 46, 100–113.

Spencer, M.R., Curtin, S.C., Garnett, M.F., 2022. Alcohol-induced death rates in the United States, 2019–2020. Centers for Disease Control and Prevention NCHS Data Brief 448, 1–8. https://doi.org/10.15620/cdc:121795

Steinhubl, S.R., Muse, E.D., Topol, E.J., 2013. Can mobile health technologies transform health care? JAMA 310, 2395–2396.

Substance Abuse and Mental Health Services Administration, 2022. National Survey on Drug Use and Health 2022 (No. NSDUH-2022-DS0001). Center for Behavioral Health Statistics and Quality.

Swift, R.M., 2003. Direct measurement of alcohol and its metabolites. Addiction 98, 73–80.

Swift, R.M., 1993. Transdermal measurement of alcohol consumption. Addiction 88, 1037–1039.

Swift, R.M., Swette, L., 1992. Assessment of ethanol consumption with a wearable, electronic ethanol sensor/recorder, in: Litten, R.Z., Allen, J.P. (Eds.), Measuring Alcohol Consumption: Psychosocial and Biochemical Methods. Humana Press, Totowa, NJ, pp. 189–202.

Takacs, J., Pollock, C.L., Guenther, J.R., Bahar, M., Napier, C., Hunt, M.A., 2014. Validation of the Fitbit One activity monitor device during treadmill walking. Journal of Science and Medicine in Sport 17, 496–500.

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

Wang, Y., Fridberg, D.J., Shortell, D.D., Leeman, R.F., Barnett, N.P., Cook, R.L., Porges, E.C., 2021. Wrist-worn alcohol biosensors: Applications and usability in behavioral research. Alcohol 92, 25–34.

Weissenborn, R., Duka, T.A., 2003. Acute alcohol effects on cognitive function in social drinkers: Their relationship to drinking habits. Psychopharmacology 165, 306–312. https://doi.org/10.1007/s00213-002-1281-1

White, A.M., 2003. What happened? Alcohol, memory blackouts, and the brain. Alcohol Research and Health 27, 186–196.

World Health Organization, 2023. A public health perspective on zero- and low-alcohol beverages (No. Brief 10). World Health Organization, Geneva.

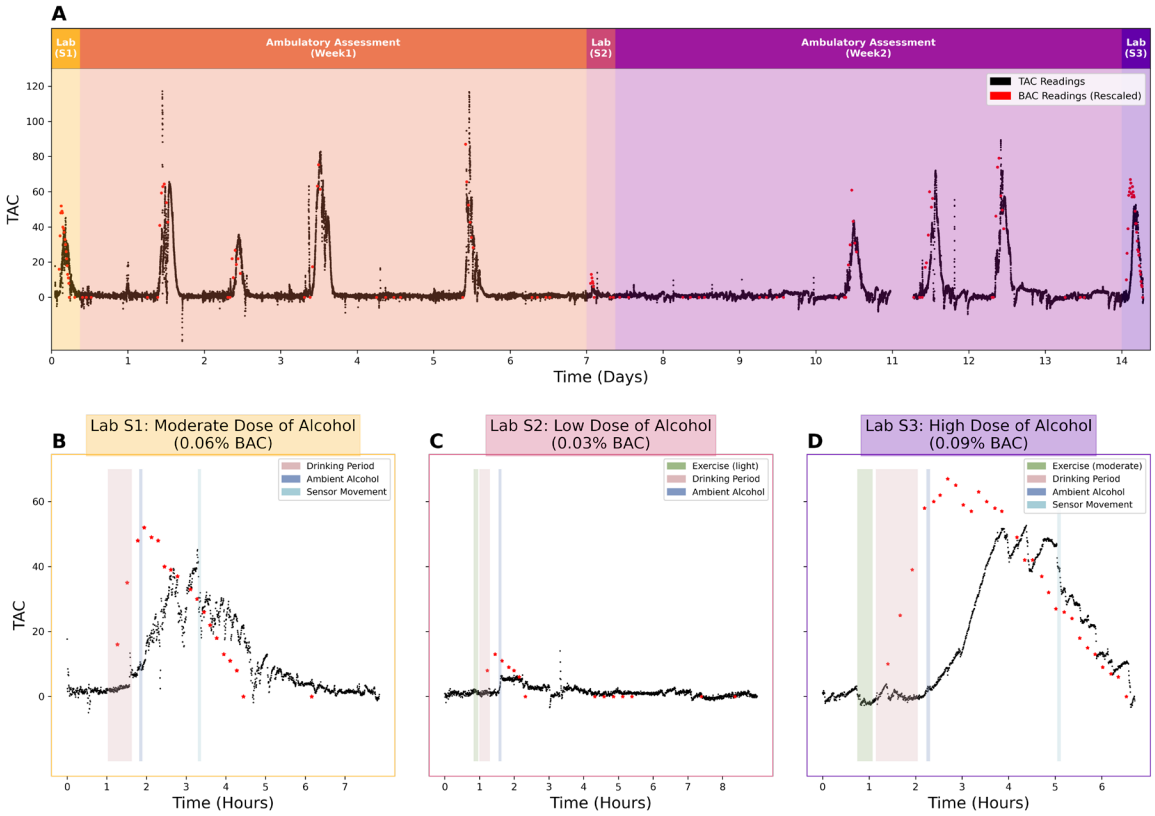
Yu, J., Fairbairn, C.E., Gurrieri, L., Caumiant, E.P., 2022. Validating transdermal alcohol biosensors: A meta-analysis of associations between BAC and transdermal alcohol sensor output. Addiction 117.

Table 1. Participant sample characteristics and drinking rate condition

Age	
Mean (SD)	24.20 (4.36)
Sex (%)	
Female	50
Male	50
Race (%)	
Native American	6
Asian	40
Black/African American	9
White	41
Mixed race	4
Ethnicity (%)	
Hispanic or Latino	24
Not Hispanic or Latino	76
Drinking Days/30	
Mean (SD)	9.39 (5.06)
Binge Days/30	
Mean (SD)	4.18 (3.87)
Heavy Drinker (%)	
Non-Heavy Drinker	67
Heavy Drinker	33
Drink Rate (%)	
Fast (0.01% *BAC*/3 minutes)	45
Slow (0.01% *BAC*/6 minutes)	55

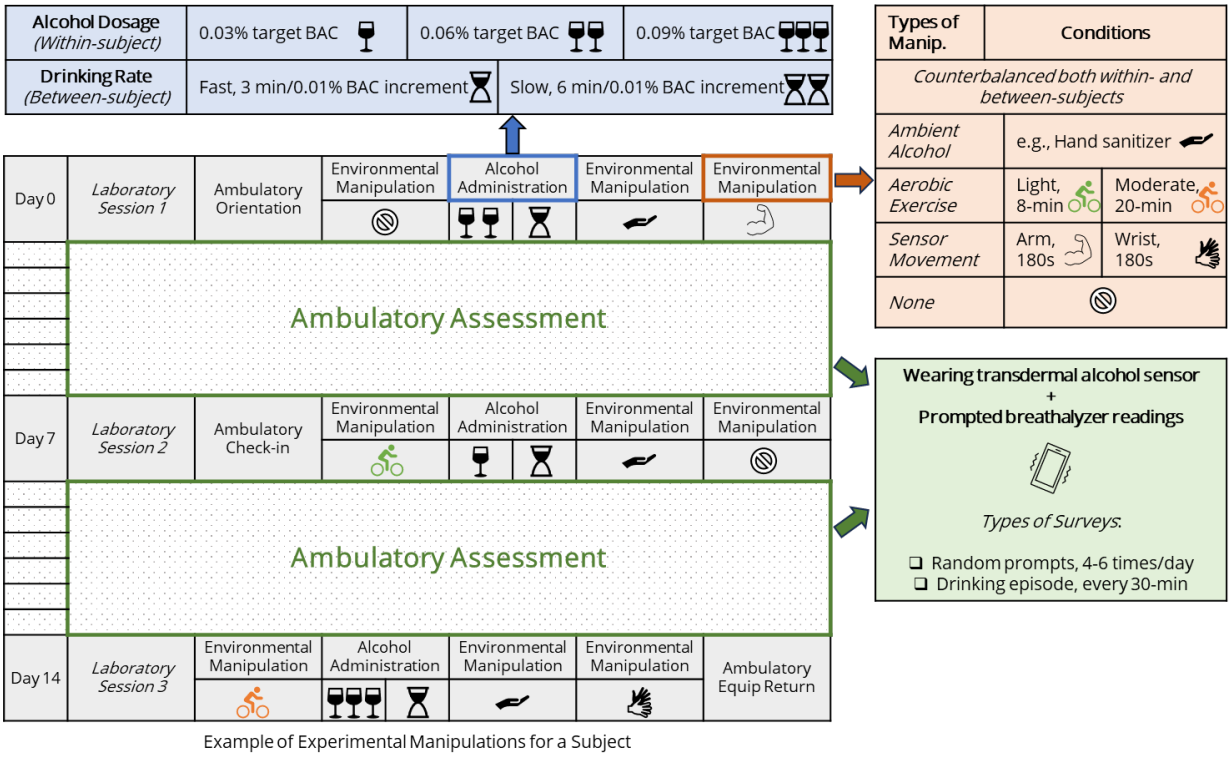
Note: Binge drinking is used to refer to 4 or more standard drinks consumed in a single sitting for females and 5 or more for males.
Heavy drinkers are defined as those who reported ≥5 binge drinking days within the past 30 days.

Figure 1. Laboratory-ambulatory study design with timeline and example data



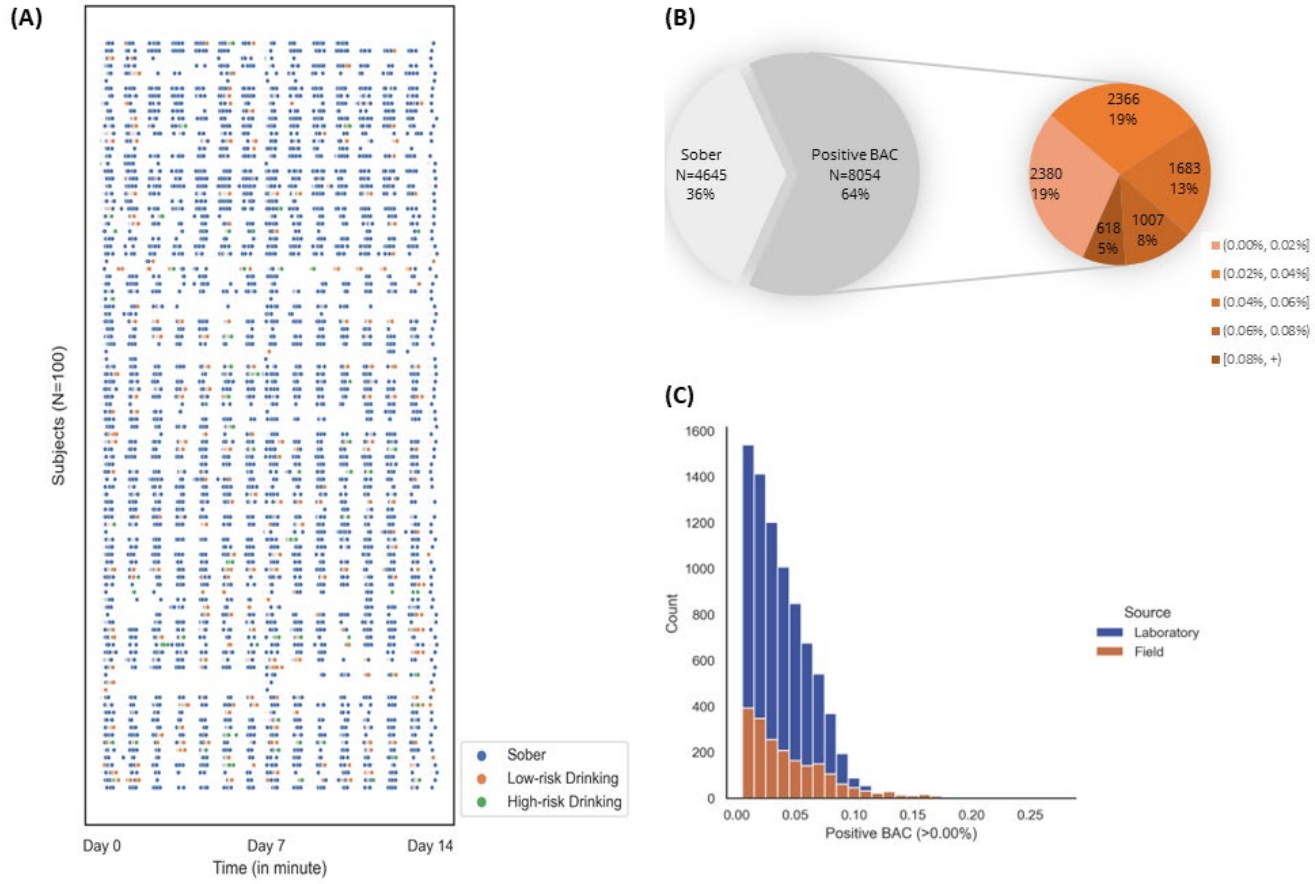
Note. *BAC* data is scaled up by a factor of 1000 for visualization purposes. (**A**) Example of raw *TAC* and *BAC* data collected from a single participant in both laboratory and field contexts. Ambulatory assessment lasted 14 days, during which time participants wore transdermal sensors and supplied *BAC* readings in real-world contexts in response to custom prompts. Experimental laboratory visits doubled as ambulatory orientation, check-in, and close-out sessions. (**B**) During laboratory session 1, participants were oriented to ambulatory study procedures and engaged in experimental alcohol-administration. (**C**) At the midpoint of the study, participants received feedback on ambulatory data provided during days 0-7 and engaged in alcohol-administration Session 2. (**D**) During the final study visit, participants engaged in the final alcohol-administration session and returned study equipment.

Figure 2. Schematic of study design, manipulations, and conditions



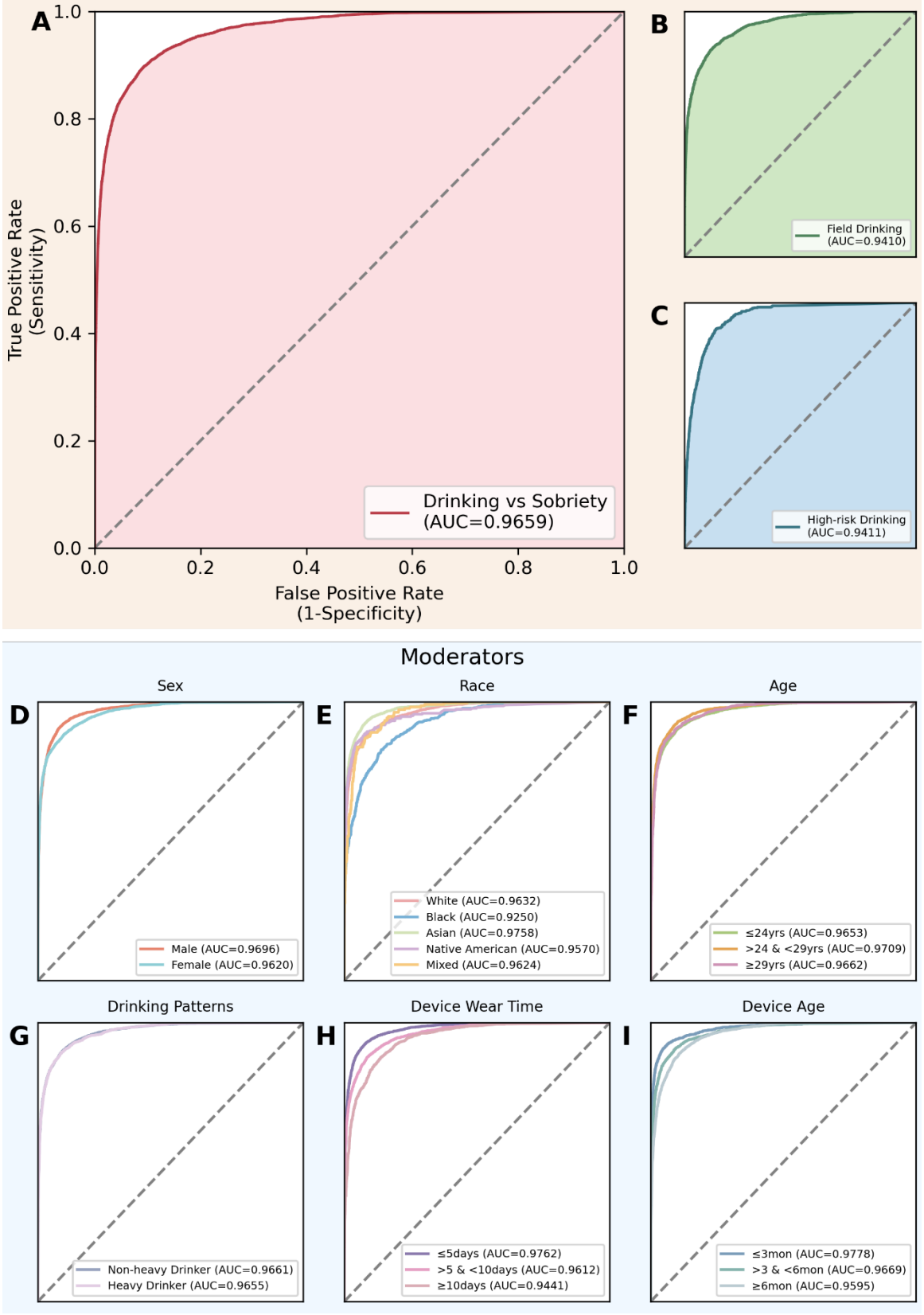
Note. All participants consumed three doses of alcohol over three laboratory sessions, targeting peak *BAC* levels of 0.03%, 0.06%, and 0.09%. Rate of consumption was operationalized as a 0.01% increment increase in target peak *BAC* being associated with a corresponding 3-minute (fast) or 6-minute (slow) increment increase in the duration of the beverage administration period. The order and type of environmental manipulation were randomized both across and between participants to yield a balanced distribution. Products employed for ambient alcohol manipulations included hand sanitizer gel, alcohol-containing hand lotion, (simulated) spilled alcoholic drink, perfume, alcohol-based cleaning product, and hand sanitizer spray. Between laboratory sessions, participants provided prompted breathalyzer readings in real-world contexts. More details on laboratory and ambulatory procedures are provided in Supplemental Methods.

Figure 3. Descriptives for breathalyzer data employed in model training



Note. *BAC* readings employed for machine learning model training displayed above. **(A)** The full dataset comprises 12,699 breathalyzer readings provided by 100 participants over 14 days. Each row represents one subject, with breathalyzer readings represented by dots. For illustration purposes, in the case of participants whose final study sessions were delayed ($N = 5$), date-stamps for these visits are rescaled to display on Day 14. **(B)** and **(C)** represent the distribution of *BAC* readings provided by subjects and the distributions of context for *BAC* assessment.

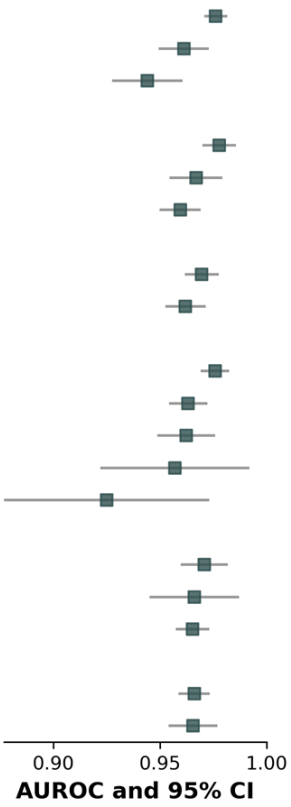
Figure 4. ROC curves depicting transdermal sensor accuracy



Note. (**A**) Receiver operating characteristic (*ROC*) curve displaying results of machine learning model for detecting drinking (>0.00% *BAC*) vs sobriety (0.00% *BAC*) in both laboratory and field contexts. (**B**) *ROC* curve for model detecting drinking vs sobriety in field contexts only. (**C**) *ROC* curve for model detecting high-risk drinking ($BAC \geq 0.08\%$) vs a combined comparison group of low-risk drinking and sobriety ($BAC < .08\%$). (**D**) – (**I**) Moderators for drinking vs sobriety detection accuracy.

Figure 5. AUROC and confidence intervals depicting transdermal sensor accuracy across moderator classes

Variable	N	AUROC	95% CI
Transdermal Device Wear Time (days)			
≤5	25,671	0.9762	0.9708, 0.9817
>5 & <10	20,539	0.9612	0.9494, 0.9729
≥10	8,460	0.9441	0.9276, 0.9606
Transdermal Device Age (months)			
≤3	14,164	0.9778	0.9699, 0.9857
>3 & <6	14,386	0.9669	0.9546, 0.9792
≥6	26,120	0.9595	0.9498, 0.9691
Sex			
Male	25,887	0.9696	0.9617, 0.9776
Female	28,783	0.9620	0.9527, 0.9714
Race			
Asian	22,444	0.9758	0.9691, 0.9825
White	23,136	0.9632	0.9542, 0.9723
Mixed	1,968	0.9624	0.9488, 0.9760
Native American	2,778	0.9570	0.9221, 0.9919
Black	4,344	0.9250	0.8769, 0.9731
Age			
>24 & <29	13,718	0.9709	0.9599, 0.9819
≥29	5,052	0.9662	0.9452, 0.9872
≤24	35,900	0.9653	0.9574, 0.9731
Drinking Patterns			
Non-Heavy Drinker	38,519	0.9661	0.9588, 0.9734
Heavy Drinker	16,151	0.9655	0.9541, 0.9770



Note. Area under the receiver operating characteristic curve (AUROC) and corresponding 95% confidence intervals across moderator classes. Heavy drinkers are defined as those who reported ≥5 binge drinking days within the past 30 days. Binge drinking refers to 4 or more standard drinks consumed in a single sitting for females and 5 or more for males. Transdermal Device Wear Time is defined as days of continuous wear time by a single participant within the study. Transdermal Device Age is defined as months elapsed since original device ship date.