Experimental and Clinical Psychopharmacology
2014, Vol. 22, No. 1, 86–96

1064-1297/14/$12.00 DOI: 10.1037/a0034821

Predictors of Detection of Alcohol Use Episodes Using a Transdermal Alcohol Sensor

Nancy P. Barnett
Brown University

E. B. Meade
University of Delaware

Tiffany R. Glynn
Brown University

The objective of this investigation was to establish the ability of the Secure Continuous Remote Alcohol Monitoring (SCRAM) alcohol sensor to detect different levels of self-reported alcohol consumption, and to determine whether gender and body mass index, alcohol dependence, bracelet version, and age of bracelet influenced detection of alcohol use. Heavy drinking adults (N = 66, 46% female) wore the SCRAM for 1–28 days and reported their alcohol use in daily Web-based surveys. Participant reports of alcohol use were matched with drinking episodes identified from bracelet readings. On days when bracelets were functional, 690 drinking episodes were reported and 502 of those episodes (72.8%) were detected using sensor data. Using generalized estimating equations, we found no gender differences in detection of reported drinking episodes (77% for women, 69% for men). In univariate analyses, at the level of fewer than 5 drinks, women's episodes were more likely to be detected, likely because of the significantly higher transdermal alcohol concentration levels of these episodes, whereas at the level of 5 or more drinks, there was no gender difference in detection (92.6% for women, 93.4% for men). In multivariable analyses, no variables other than number of drinks significantly predicted alcohol detection. In summary, the SCRAM sensor is very good at detecting 5 or more drinks; performance of the monitor below this level was better among women because of their higher transdermal alcohol concentration levels. Individual person characteristics and bracelet features were not related to detection after number of drinks was included. Minimal bracelet malfunctions were noted.

Keywords: transdermal alcohol, alcohol sensor, alcohol detection

Objective measurement of alcohol consumption is most often collected using breath alcohol or blood tests. However, these tests are not commonly used to provide continuous measurement of alcohol use outside of a laboratory setting as they necessitate frequent observations. Furthermore, biochemical tests provide measures of very recent alcohol use and current levels of intoxication, but they do not provide information about frequency and timing of alcohol consumption. Biosensors are available that detect alcohol vapor at the surface of the skin and are worn continuously; thus, they are able to address some of the limitations of blood and breath tests. This investigation evaluated the ability of one alcohol sensor to detect episodes of alcohol consumption in naturalistic settings.

Transdermal Alcohol Sensors

Approximately 1% of consumed alcohol is excreted through the skin and can be detected with biochemical sensors (Swift, 2003). The available sensors collect transdermal alcohol concentration (TAC) readings at regular intervals and store these readings for later download, thus providing an indication of alcohol consumption across days and weeks of wear. Evaluations of sensors have concluded that TAC correlates well with simultaneously measured blood alcohol concentration (BAC) or breath alcohol concentration (BrAC; Dougherty et al., 2012; Sakai, Mikulich-Gilbertson, Long, & Crowley, 2006; Swift, Martin, Swette, LaConti, & Kackley, 1992). Sakai and colleagues (2006) administered doses of alcohol (.00, .28, or .56 g/kg) in a laboratory setting to 24 participants wearing the Secure Continuous Remote Alcohol Monitoring (SCRAM) ankle monitor from Alcohol Monitoring Systems (AMS; Littleton, CO). The researchers found a robust correlation between transdermal and breath results across subjects (peak al-

Nancy P. Barnett, Center for Alcohol and Addiction Studies and Department of Behavioral and Social Sciences, Brown University; E. B. Meade, Department of Psychology, University of Delaware; Tiffany R. Glynn, Center for Alcohol and Addiction Studies, Brown University.

This research was supported in part by Grants R21 AA015980 and R21 AA020943 from the National Institute on Alcohol Abuse and Alcoholism to Nancy P. Barnett. The funder had no further role in the study design, collection, analysis or interpretation of the data, writing the manuscript, or the decision to submit the paper for publication.

All authors declare that they have no conflicts of interest.

We thank Polly Gobin, John Hustad, Michelle Loxley, Nadine Mastroleo, James Murphy, Lindsay Orchowski, Robert Swift, and Jennifer Tidey who assisted in this research, and Tim Souza and Cheryl Eaton who prepared data.

Each author contributed significantly to the work, reviewed the manuscript, and agreed to its submission for publication.

Correspondence concerning this article should be addressed to Nancy P. Barnett, Center for Alcohol and Addiction Studies, Brown University, Box G-S121-5, Providence, RI 02912. E-mail: Nancy_Barnett@brown.edu

cohol concentration $r = .84$, $p < .01$). Dougherty et al. (2012) administered increasing doses (1–5 drinks) of alcohol to 21 participants wearing the SCRAM on multiple days and established that BrAC and TAC were highly correlated within person ($r = .91$ for women, $r = .86$ for men, $ps < .001$).

Detecting Alcohol Use Using Sensors

In validation research, TAC is validated against BrAC when alcohol use is known to have occurred (Dougherty et al., 2012; Sakai et al., 2006; Swift et al., 1992). It is also essential that these sensors be evaluated in their ability to detect alcohol use in real-world settings. For example, Marques and McKnight (2009) collected transdermal data and self-administered BrAC readings from 22 participants for 2, 4, or 6 weeks. Using TAC data from the SCRAM, the researchers correctly detected 57% of drinking episodes (defined as having a peak BrAC of ≥.02 g/dl) when the criterion for detection was TAC greater than .02 g/dl, and 79% when drinking episodes with TACs less than .02 g/dl were included. In a study in which the SCRAM was used to verify compliance with an alcohol reduction intervention, Barnett, Tidey, Murphy, Swift, and Colby (2011) used a set of criteria adapted from the SCRAM manufacturer to detect alcohol use in participants, and 91% of self-reported drinking days were detected.

Possible Influences on Detection of Transdermal Alcohol

Prior research has established that TAC correlates well with BrAC and that alcohol use episodes can be detected using TAC with specific criteria. There is also some evidence that individual differences may be associated with the detection of alcohol using transdermal sensors (Hawthorne & Wojcik, 2006), including the dose of alcohol, gender, body mass, severity of prior alcohol use, and bracelet characteristics.

Volume of Alcohol Consumed

Thus far, investigations have established that lower levels of drinking (i.e., smaller administered doses or fewer self-reported drinks) are less likely to be detected by the transdermal sensors. For example, Sakai and colleagues (2006) reported that the SCRAM was able to discriminate between lower and higher dosed participants and the authors concluded that the sensor consistently detected the consumption of approximately two standard drinks. In Barnett et al. (2011), we reported that self-reported drinking days that were not detected by TAC criteria tended to be days on which fewer drinks were consumed. Findings from these studies suggest that lower doses are less reliably detected, but no investigations have established the probability of detection of alcohol use at different numbers of drinks in the field. Furthermore, it is important to determine whether heavy drinking episodes (five or more drinks) are adequately detected, as these episodes are more likely to result in alcohol-related morbidity and mortality (Cherpitel, Pares, Rodes, & Rosovsky, 1993; Rehm, Greenfield, & Rogers, 2001; Rehm, Room, Graham, Monteiro, Gmel, & Sempos, 2003).

Severity of Alcohol Use/Dependence Status

Heavy drinkers and alcohol-dependent participants have commonly been included in research investigating transdermal sensors (Barnett et al., 2011; Sakai et al., 2006; Swift et al., 1992), but their alcohol use history has rarely been investigated as a predictor of detection of alcohol use using a sensor. In one of the only studies to consider alcohol dependence status, Sakai and colleagues (2006) collected BrAC, self-report, and transdermal sensor data on 10 dependent and 10 nondependent drinkers and found that using the sensor data, they were able to discriminate between the two types of drinkers. There is evidence that one's history of alcohol use influences his or her metabolism of alcohol, whereby elimination of alcohol is faster among those with higher alcohol use (Jones, 2008). Furthermore, there is an association between the metabolism of alcohol and risk of alcohol disorders (Edenberg, 2007). Given the associations between heavy alcohol use, alcohol disorders, and metabolism of alcohol, it is relevant to investigate whether severity of alcohol use is a predictor of the detection of alcohol use using transdermal alcohol detection methods.

Gender

Gender differences in drinking and physiology suggest that there may be a gender difference in the detection of alcohol use using a sensor. Men tend to consume more drinks per episode, which could result in a higher BAC and, for our purposes, better TAC detection among men. On the other hand, women tend to be smaller than men, so the same dose of alcohol (i.e., the same number of drinks) tends to lead to higher BACs in women, suggesting that women's drinking might show higher rates of detection. In addition, women have lower body water by weight (Watson, Watson, & Batt, 1980), so alcohol is more concentrated in women's blood and thus might result in higher detection rates. To date, only one investigation using transdermal sensors has considered gender differences in the detection of alcohol use. Marques and McKnight (2009) established that after controlling for maximum BAC and days of bracelet wear, female drinking was marginally significantly less readily detected with the SCRAM than male drinking, but this finding has not been replicated with newer sensor versions.

Body Mass

Body mass influences the processing of consumed alcohol, with higher body weights generally showing lower BAC at the same number of drinks consumed. Whether the detection of alcohol use by the SCRAM differs according to body mass has not been investigated, but might be expected.

Bracelet Characteristics

In the small literature on the validity of transdermal sensors, there is some evidence that with more days of wear, the performance of the SCRAM bracelet declined, resulting in lower detection on later days of wear (Marques & McKnight, 2009). Newer versions of the bracelet have become available since this report, but these versions have not been evaluated as to whether higher days of wear are associated with poorer detection.

Summary and Overview of the Present Study

Investigations have established that TAC correlates well with BrAC, can be used to detect known drinking episodes and days of

drinking, and can distinguish between drinking episodes with different amounts consumed. There are indications that personal characteristics may influence detection, but these effects have not been thoroughly investigated in one study. This information is necessary as it provides greater clarity about the validity of the SCRAM for use in different populations, particularly as it becomes more commonly used with nonoffenders (Leffingwell et al., 2013).

The primary purpose of this investigation was to use all available self-report and SCRAM alcohol sensor data from two investigations to establish rates of detection of self-reported episodes of alcohol use across different levels of drinking in the field and to investigate characteristics that influence detection. Although self-reported drinking is not an objective indicator of alcohol use, we used procedures that support the validity of participant report and evaluate possible sources of bias, including study design characteristics and delays in completing self-report surveys. Using detection criteria adapted from the SCRAM manufacturer and evaluated previously (Barnett et al., 2011), we initially established the agreement between episodes identified using TAC and self-report and present the proportion of self-reported drinking episodes that were detected with TAC for different numbers of self-reported drinks per episode. We next investigated demographics (gender, body mass index [BMI]), bracelet characteristics (bracelet version and length of time the bracelet had been worn), alcohol dependence, estimated BAC (eBAC), and number of drinks in the episode as predictors of drinking episode detection using the SCRAM. We expected that women, individuals with lower BMI, and episodes with higher numbers of drinks and eBAC would show higher detection rates.

Method

Data from two investigations of contingency management that used the SCRAM bracelet (Barnett, Tidey, Murphy, Colby, & Swift, 2013; Barnett et al., 2011) were used in the current study. Inclusion criteria and data collection procedures were identical for the two studies. The university institutional review board approved all procedures.

Participants

Participants were recruited through online and newspaper advertisements and with fliers posted in the community. Advertisements invited adult drinkers who were interested in reducing or stopping drinking to contact the study. During telephone screening, callers who met the following criteria were invited to enroll: (a) age 18 or older; (b) reported past-month drinking above the national recommendations for alcohol use: eight or more drinks per week for women and 15 or more drinks per week for men (U.S. Department of Health and Human Services and U.S. Department of Agriculture, 2005); (c) reported two or more heavy drinking episodes per week for the past month; (d) had an e-mail address and daily Internet access; and (e) either had a landline phone for transmission of the bracelet data or were willing to come to the research office three times each week to download data from the bracelet. Callers were not invited to participate if they reported significant alcohol withdrawal symptoms (Alcohol Withdrawal Symptom Checklist score $\geq$ 23; Pittman et al., 2007) or if they reported using drugs other than marijuana in the past month or more often than once a month in the past year. Individuals who were seeking treatment were not enrolled and were provided referrals.

Procedure

All in-person procedures were conducted at Brown University research offices. To confirm the inclusion criteria, participants received in-person screening for severe alcohol withdrawal (Clinical Institute Withdrawal Assessment for Alcohol score $\geq$ 10; Sullivan, Sykora, Schneiderman, Naranjo, & Sellers, 1989) and submitted a urine sample that was tested for drugs other than marijuana. If eligible after screening, participants gave informed consent and completed self-report and interview measures. The research interviewer collected weight and height. The SCRAM bracelet was attached to the participant's ankle and monitoring was activated through the SCRAM Website. Participants received instructions about completing daily Web surveys and setting up the SCRAM modem at home to download bracelet readings. Participants were prepared to wear the bracelet for 21 or 28 days.

The first week of both studies was a baseline week in which participants were informed that they were not expected to make any changes and no contingencies were provided. After the baseline week, participants were retained in the intervention trial if their baseline week drinking showed either (a) two TAC peaks above .08 g/dl or (b) two of the three following self-report criteria: reporting above the national weekly recommendations for their gender (8+ for women, 15+ for men), reporting two or more heavy drinking episodes (four or more drinks per episode for women, five or more for men), or showing an average drinking quantity at the heavy drinking episode level. These criteria were essentially a reliability check on our initial inclusion criteria. Thirteen participants completed a pilot trial in which all were assigned to a contingency management condition (Barnett et al., 2011), 31 were enrolled in a randomized controlled trial in which half were assigned to contingency management and half to noncontingent reinforcement (Barnett et al., 2013), and 28 were excluded after the first week for not meeting inclusion criteria. For the current investigation, since the purpose was to evaluate the sensor's ability to detect episodes of any size, all participants were included regardless of whether they were retained in one of the intervention trials after baseline.

Measures

Baseline. Measures collected at baseline included gender, age, race, ethnicity, education, marital status, height, and weight. BMI was calculated using the following formula: (mass $\times$ 703)/height^2 (National Center for Chronic Disease Prevention and Health Promotion, 2012). The research interviewer administered a 30-day timeline follow-back (Sobell & Sobell, 1992, 1995) and recorded the number of drinks and time spent drinking on each of the past 30 days. A standard drink was defined as 12 oz. of beer or wine cooler, 5 oz. of wine, or 1.5 oz. of liquor. The Alcohol and Substance sections of the Structured Clinical Interview for Diagnosis (First, Spitzer, Gibbon, & Williams, 2002) were administered to establish diagnoses.

Daily self-report. At the baseline appointment, participants were provided instructions about completing daily Web surveys

and about calculating the number of drinks using standard drink units. Starting the day after enrollment, participants received an e-mail every morning containing a unique link to a brief Web survey. The survey asked participants how many drinking episodes they had on the previous day, how many drinks they consumed during each episode, and the start time of the first drink and the end time of the last drink for each episode. Participants were asked to define for themselves what a drinking episode entailed. Self-report episodes were later combined when the time interval between them was 30 min or less, as transdermal readings are taken by the SCRAM every 30 min, so drinking episodes less than 30 min apart would not be detected as distinct episodes by TAC. BAC per episode was estimated using gender, body weight, self-reported number of drinks, and time spent drinking (Matthews & Miller, 1979). Participants were paid $5 for each completed daily Web survey and a $25 bonus if 90% or more were completed on the day received.

Transdermal data. Transdermal data were collected by the SCRAMII (earlier version) and SCRAMx versions of the bracelet. There are no differences in the alcohol sensor between the two versions. The bracelet is fastened to the wearer's ankle with a locking clip that prevents the wearer from removing it without breaking the lock or cutting the strap. Every 30 min, the SCRAM draws in a sample of the vapor above the participant's skin. These readings are stored on the bracelet and can be uploaded to the SCRAM data server two ways: The bracelet can wirelessly transmit readings via radio frequency to a modem that then transmits the readings to the data server through the participant's home phone line, or data can be downloaded in person using the company's DirectConnect device that clips onto the bracelet and transfers the readings through a USB connection to a personal computer that then sends the readings to the data server via the Internet. Once received by AMS, TAC values can immediately be viewed and downloaded from a password-protected secure Website. To obtain the sensor data used in this study, we downloaded separate Excel data files for each participant. These files contained TAC values and their associated date and time stamps.

TAC episodes were identified using the following criteria: (a) one TAC reading $\geq$ .02 g/dl and (b) absorption rate $<$ .05 g/dl per hour or elimination rate $<$ .025 g/dl per hour if peak $\leq$ .15 g/dl; $<$ .035 g/dl per hour if peak $>$ .15 g/dl. Absorption rate was calculated as peak TAC divided by the time it took to rise from .000 to the peak; elimination rate was calculated as peak TAC divided by the time it took to decline from peak to .000 again. These criteria were derived from AMS criteria that are more conservative, including requiring three TAC readings $\geq$ .02 g/dl and requiring the episode to meet both absorption and elimination criteria. In our previous work evaluating our adapted criteria with human coders, we had excellent interrater reliability (intraclass correlation coefficient = .99). Sensitivity (.91) and specificity (.97) relative to self-report were excellent as well (Barnett et al., 2011). Given their performance in earlier work, we used these criteria for the current study, but provide information as well on episodes that did not meet criteria.

Bracelet characteristics. For every drinking episode, the bracelet version (SCRAMII or SCRAMx) was recorded. Bracelets are sanitized and reused; for each drinking episode, the number of days the bracelet had been in use overall before the specified episode (available in AMS records for each bracelet) and the number of days the bracelet had been worn by the specific participant were recorded.

Matching Self-Report With Transdermal Episodes

TAC episodes detected using the above criteria were matched with self-reported number of drinks in the episode by using day and time records. A number of specific data-based rules were followed. First, given the documented lag in TAC onset relative to drinking (Hawthorne & Wojcik, 2006; Sakai et al., 2006), the first pass entailed checking for TAC episodes that started within 5 hr of the completed drinking episode. Second, if two self-reported episodes occurred within the time of a single detected transdermal episode (i.e., TAC did not return to .000), the self-report episodes were combined into a single episode. In these cases, eBAC was calculated for the combined episode. Third, when two or more detected transdermal episodes occurred within the period of a single self-report episode, the transdermal episode data were combined. For these episodes, the peak TAC was the highest TAC value in the combined episode. Fourth, there were some cases in which TAC remained elevated over multiple days without returning to .000, but two or more clear TAC peaks were evident. When the participant reported distinct drinking episodes on these days, the multiday TAC episodes were manually divided at the lowest point between peaks, producing separate TAC episodes (and separate associated eBAC).

Potential Episodes Not Meeting Criteria

In the interest of fully investigating the available data, we identified all self-reported episodes that had not matched with TAC episodes meeting our criteria and determined whether any TAC elevation occurred around the time of the self-reported episode. We matched episodes with these elevations using the following guidelines: (a) if the time interval of the self-reported drinking episode overlapped the time of the TAC elevation, (b) if the start of the TAC elevation was 5 hr or less after the self-reported start time, or (c) if the TAC elevation returned to .000 g/dl 1 hr or less before the self-report start time, to allow for participant error in recording the actual time of drinking. We subsequently characterized these episodes as to whether they met either of our two primary detection criteria (i.e., peak of .02 g/dl or higher or absorption/elimination rate criteria).

Data Analysis

A person-level data set was used to describe participants; calculated variables included number of drinking days, number of drinks per drinking day, average peak TAC per episode, and drinking detection rate across episodes. For episode-level analyses, a person-period data set was used. In addition to participant characteristics of gender, BMI, and alcohol dependence, episode-level variables included number of drinks, eBAC, peak TAC, bracelet version, and number of days the bracelet had been in use prior to the day on which the episode occurred. To examine possible correlates (gender, BMI, bracelet version, number of days of bracelet use, current alcohol dependence, self-reported number of drinks, and eBAC) of TAC detection, we evaluated generalized estimating equations (GEE; Zeger & Liang, 1986) mod-

els (binomial distribution, exchangeable correlation structure, logit link, hereafter called logistic GEE) with the event-level data. Following these individual GEE analyses with the different predictors, we included significant predictors in one multivariable GEE analysis. To examine gender differences in number of drinks and peak TAC per episode, we conducted GEE models (normal distribution, exchangeable correlation structure, identity link, hereafter called linear GEE) using the event-level data. Heavy drinking (five or more drinks per episode) was investigated as a moderator of the gender association with episode detection by calculating a Heavy Drinking Episode × Gender interaction term and using it in a logistic GEE analysis to predict episode detection. Number of drinks was square-root transformed because of nonnormality.

Results

Of the 72 participants in our sample, two were removed for failing to provide at least one self-report and/or transdermal episode. Four additional participants were excluded because no valid TAC data were available during all self-reported episodes (see Missing TAC Data section for details). The remaining sample of 66 participants was 45.5% female with an average age of 30.6 years ($SD = 10.5$). Of the 66 participants, 49 (74.2%) were White, eight (12.1%) were Black, one (1.5%) was Asian, five (7.6%) were multiracial, and three (4.5%) did not report a race. Five (7.6%) were Latino. Most ($n = 62$, 93.9%) had a high school education or equivalent, with an average of 13.9 ($SD = 4.2$) years of school completed. Of the participants, 42 (63.6%) were never married, 17 (25.8%) were married or living together, and seven (10.6%) were divorced, widowed, or separated. Average BMI was 28.4 ($SD = 6.5$), and did not differ between men ($M = 28.1$) and women ($M = 28.7$), $t(63) = 0.33$, *ns*. At baseline, participants reported drinking alcohol on 20.2 ($SD = 6.8$) days in the past month, with an average of 7.2 ($SD = 2.8$) drinks per drinking day. Most met criteria for current alcohol dependence ($n = 25$, 37.9%) or alcohol abuse ($n = 11$, 16.7%).

In the baseline week, which was the only week that contained some data for all participants, there were significant differences between participants who were retained in the clinical trials and those who were excluded on number of drinks per drinking day ($M = 7.3$, $SD = 3.3$ for those included vs. $M = 4.7$, $SD = 2.4$ for those excluded), $t(68) = 3.60$, $p = .001$, and daily average eBAC ($M = .095$ g/dl, $SD = .058$ for those included vs. $M = .058$ g/dl, $SD = .062$ for those excluded), $t(63) = 2.41$, $p = .02$. There was a similar difference in daily average TAC in the first week ($M = .025$ g/dl, $SD = .028$ for those included vs. $M = .009$ g/dl, $SD = .017$ for those excluded), $t(64) = 2.52$, $p = .01$. Among participants who were included in one of the two trials, there were no differences on number of drinks per drinking day, eBAC, or daily average TAC between participants in the two intervention conditions.

Missing TAC Data

In the initial sample of 70 participants, 11 participants (15.7%) had some missing TAC data while they were wearing the bracelet. Missing data were due to two circumstances: (1) No data were collected or transferred (three participants) or (2) data were collected but there was evidence of equipment malfunction (eight participants) for a total of 66 days (of a total of 1,295 days of bracelet wear; 5.1%). GEE analyses of bracelet characteristics established that the number of days the bracelet had been worn overall (OR = 0.99, 95% CI [0.97, 1.02,] $p = .62$), the number of days worn by the participant (OR = 1.00, 95% CI [0.92, 1.10], $p = .95$), and the bracelet version (OR = 0.45, 95% CI [0.02, 9.41], $p = .61$) did not predict missing TAC data.

On days when bracelet data were determined to be invalid, there were 50 self-reported drinking episodes (6.8% of all self-reported episodes). Any transdermal data recorded when the equipment was malfunctioning precluded episode detection using TAC, so these data were excluded from further analyses. For all subsequent analyses, we evaluated the ability of the sensor to detect self-reported alcohol use when the bracelet was determined to be functioning correctly.[1] Four participants were excluded from all analyses because no valid TAC data were collected during their self-reported drinking episodes.

Daily Web Surveys

For the 66 participants with one or more days of valid bracelet data, 89.8% of the daily Web surveys were completed on the same day they were requested. Of those that were not completed on time, 97% were completed within 1 week ($M = 2.2$ days, $Mdn = 1$). Submitting a daily Web survey late was not significantly related to peak TAC on that day (TAC on on-time days = .068 g/dl; late days = .070 g/dl), $B = .00$ $p = .68$.

Bracelet Version

Of the 66 participants with valid bracelet data, 42 (63.6%) wore the SCRAMII, 20 (30.3%) wore the SCRAMx, and four (6.1%) wore both (this occurred when the bracelet was changed due to a malfunction).

Self-Reported Drinking Episodes, eBAC, and TAC

The median number of self-reported drinking episodes reported per participant was 8.0 (range = 1–30), with an average of 6.3 drinks per episode (range = 0.1–30, $SD = 4.5$). The average eBAC for self-reported episodes was .083 g/dl ($SD = .075$). The average TAC per episode was .100 g/dl ($SD = .097$). The weighted correlation between TAC and eBAC was .54, $p < .001$.[2]

Detection of Episodes

When bracelets were functional, 690 self-reported drinking episodes were reported by the 66 participants. Using our TAC detection criteria, we detected 502 of the 690 self-reported epi-

[1] According to correspondence with AMS, the majority of bracelet malfunctions occurred because of a specific bracelet component not functioning well (personal communication, Jeffrey Hawthorne, Alcohol Monitoring Systems, January 19, 2011). Because this problem has since been corrected, it is not expected that the malfunction will generalize to future users. For this reason, this article describes *functional* bracelet performance.

[2] The correlation was weighted because it was calculated at the episode level and subjects varied as to how many observations they contributed.

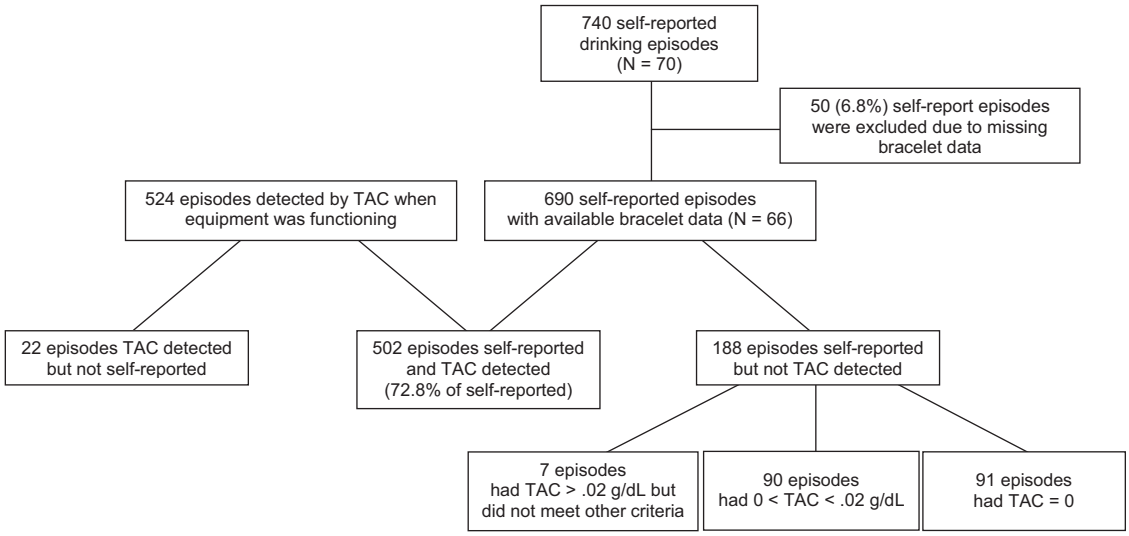

Figure 1. Flowchart of alcohol use episodes and detection.

sodes (72.8%), as shown in Figure 1. An additional 22 episodes (4.2% of all TAC episodes) met our criteria but could not be matched with a self-report episode. These episodes could be considered false positives (i.e., drinking was detected when none occurred) or false self-reports (i.e., drinking occurred but was not reported). Using GEE, we established that the 22 TAC episodes that did not match with a self-reported episode had significantly lower peak TAC (M = .047 g/dl, SD = .035) than the 502 detected episodes that matched with a self-report (M = .132 g/dl, SD = .091), $B = -.06$, $p = .001$. We also reviewed the AMS ratings for these episodes, and found that only 4 of the 22 were identified as confirmable drinking events by AMS.

There were 188 reported drinking episodes that were not detected using our TAC criteria (see Figure 1). Of these 188 episodes, seven had TAC elevation greater than .02 g/dl (M = .180 g/dl, SD = .114, range = 0.031–0.324) but were not detected because of failure to meet absorption or elimination criteria. Ninety episodes were not detected because of TAC elevation less than 0.02 g/dl; these episodes had an average peak TAC of .010 g/dl (SD = .004, range = .005–.019 g/dl). Ninety-one reported drinking episodes had no associated TAC elevation. For the 181 episodes that were not detected due to not reaching .02 g/dl (i.e., the 90 that were between .00 and .02 g/dl and the 91 that were .00), the average number of drinks per episode was 2.5 (SD = 1.6, Mdn = 2.0). More information about the self-reported drinking episodes is in Table 1.

Putative Predictors of Episode Detection

Gender. Women reported 335 drinking episodes (48.6% of 690 episodes) and men reported 355 episodes (51.4%). The detection rates using TAC were 77.0% for women's episodes and 68.7% for men's; using GEE, we found a nonsignificant difference (OR = 0.66, 95% CI [0.36, 1.21], p = .18). Given the interest in determining detection rates of heavy drinking, we calculated an interaction term with gender and a dichotomous term of five or more drinks and entered the interaction into a logistic GEE. The interaction term was significant (OR = 19.64, 95% CI [4.68, 82.35], $p < .001$), and indicated that at fewer than five self-reported drinks, women's episodes were more likely to be detected (53.4% vs. 32.6% for men), but at the level of five or more drinks, there was no gender difference in detection (92.6% of women's vs. 93.4% of men's). Of note is that at the level of fewer than five drinks, women drank an average of 2.6 drinks (SD = 1.1) per episode, compared with men's 2.4 (SD = 1.2) drinks, a nonsignificant difference, B = .09, p = .11, whereas at the higher level of five or more drinks, men showed significantly higher number of drinks per episode (M = 9.8, SD = 4.9 vs. M = 7.9, SD = 2.9 for women), $B = -.26$, p = .02. The detection of episodes at different numbers of drinks is displayed in Figure 2.

To explore further the gender differences in detected and undetected episodes, we compared the peak TAC of all 690 self-reported episodes between genders (see Figure 3). Self-reported episodes that showed no TAC elevation were assigned a value of .000 g/dl. Using linear GEE with peak TAC as the dependent variable, we established that when four or fewer drinks were consumed, women reached higher peak TAC (M = .038 g/dl, SD = .047) than men (M = .023 g/dl, SD = .036), B = .02, p = .02. The same was true when eBAC was the dependent variable; women reached higher eBAC (M = .041 g/dl, SD = .037) than men (M = .022 g/dl, SD = .020), B = .02, $p < .001$. At drinking levels higher than four drinks, there were no differences between the peak TAC of women (M = .146 g/dl, SD = .096) and men (M = .146 g/dl, SD = .094), $B = -.01$, p = .79. However, women showed higher eBAC (M = .131 g/dl, SD = .069 vs. M = .103 g/dl, SD = .075 for men), B = .03, p = .03. Therefore, at the lower level of four or fewer drinks, despite no difference in number of drinks between men and women, women's episodes reached higher TAC and eBAC and were more likely to be detected than men's. At the higher level of five or more drinks, men drank more per episode and women had higher episode eBACs, but there were no gender differences in detection or in peak TAC.

Table 1
Characteristics of Self-Reported Drinking Episodes

Criteria						
TAC (g/dL)	Met absorption/ elimination	Count	%	Number of drinks *M* (*SD*)	Peak TAC (g/dl) *M* (*SD*)	eBAC (g/dl) *M* (*SD*)
≤ .02	Yes	502	72.8	7.71 (4.42)	.132 (.091)	.102 (.075)
≥ .02	No	7	1.0	5.36 (3.33)	.180 (.114)[a]	.080 (.054)
.01 ≤ and < .02	N/A[b]	41	5.9	3.23 (1.56)	.014 (.003)	.040 (.034)
.00 < and < .01	N/A[b]	49	7.1	2.49 (1.53)	.007 (.001)	.026 (.032)
0	N/A[b]	91	13.2	2.09 (1.61)	.000	.027 (.032)
Total		690	100	6.31 (4.51)	.100 (.097)	.083 (.073)

Note. TAC = transdermal alcohol concentration; eBAC = estimated blood alcohol concentration.
[a] This average TAC is considerably higher than the episodes that were detected, suggesting that the individuals had absorption and elimination rates that were outside the norm or that environmental alcohol contributed to these episodes. [b] Absorption/elimination rates are not meaningful for such small peaks.

BMI. As participants' BMI increased, drinking detection rate (i.e., the proportion of self-reported episodes detected) tended to decrease, $r(65) = -.32$, $p = .01$.[3] Using the episode-level data, we found that the average BMI of individuals with detected episodes was 26.6 ($SD = 5.7$), the BMI of individuals with nondetected episodes was 28.7 ($SD = 7.4$), and BMI was a significant predictor of episode detection (OR = 0.95, 95% CI [0.92, 0.99], $p = .006$). Using BMI categories of normal (18.5–24.9; 38.5% of sample), overweight (25.0–29.9; 29.2% of sample) and obese (≥30; 32.3% of sample; National Center for Chronic Disease Prevention and Health Promotion, 2012), we established that participants with a BMI in the normal range had a detection rate of 70.6%; overweight and obese participants' rates were 71.0% and 56.8%, respectively.

Current alcohol dependence. Participants with a current diagnosis of alcohol dependence drank an average of 6.6 drinks ($SD = 3.3$) per drinking day, and those without dependence averaged 5.7 drinks ($SD = 2.8$) per drinking day, and episodes reported by participants with dependence were more likely to be detected (80.1%) compared with those reported by nondependent participants (67.6%), (OR = 1.93, 95% CI [1.01, 3.70], $p = .047$).

Bracelet version and number of days of bracelet use. As described previously, 72.8% of self-reported drinking episodes were detected using our TAC criteria. The proportion of episodes detected by the SCRAMII was 76.7% and by the SCRAMx was 67.1%. Using logistic GEE, we found that the SCRAMII showed higher rates of detection (OR = 0.54, 95% CI [0.32, 0.90], $p = .02$). The total number of days a bracelet had been worn (by the current participant and all other users) before the day on which a self-reported episode was reported ranged from 0 to 317; the average for self-reported episodes that were detected was 87.2 ($SD = 76.8$), and for nondetected episodes was 84.4 ($SD = 76.0$); this was not a significant predictor of whether the episode was detected (OR = 1.00, 95% CI [0.999, 1.00], $p = .33$). Number of days that the bracelet had been worn by the participant only ($M = 9.8$, $SD = 7.6$ for detected episodes and $M = 8.2$, $SD = 7.7$ for nondetected episodes) also was not a significant predictor of detection (OR = 1.01 95% CI [0.99, 1.03], $p = .37$).

eBAC. The average eBAC for TAC-detected self-reported drinking episodes was .102 g/dl ($SD = .078$) and .032 g/dl ($SD = .033$) for nondetected episodes. Using logistic GEE, we found that eBAC was a significant predictor of episode detection (OR = 20.1, 95% CI [8.90,45.37], $p < .001$).

Self-reported number of drinks. The average self-reported number of drinks was 7.7 drinks ($SD = 4.4$) for detected episodes and 2.6 drinks ($SD = 1.8$) for nondetected episodes. Using logistic GEE, we found that (transformed) number of drinks was a significant predictor of episode detection (OR = 21.0, 95% CI [11.79, 37.41], $p < .001$).

Research design elements. To ensure that episode detection was not an artifact of the research design components, we investigated stage of the study (baseline vs. intervention), and assignment to contingent reinforcement or noncontingent reinforcement (for participants who were included in one of the intervention trials; $n = 44$ intervention weeks only). Neither stage of the study (OR = 0.93, 95% CI [0.63, 1.37], $p = .71$) nor condition (OR = 1.15, 95% CI [0.55, 2.41], $p = .71$) was significantly associated with detection of self-reported drinking episodes.

Multivariable analysis. Variables that were significantly associated with episode detection—Gender × Heavy Drinking interaction, BMI, alcohol dependence, bracelet version, and number of drinks (transformed)—were included together in a logistic GEE analysis. eBAC was not included in this analysis because of its high collinearity with number of drinks ($r = .77$, $p < .001$). The model revealed that number of drinks had the only significant association with TAC detection. Odds ratios and 95% confidence intervals are presented in Table 2.

Discussion

In a sample of adults screened for heavy drinking, we detected almost three of every four self-reported drinking episodes using the SCRAM alcohol sensor. As expected, the more drinks consumed in the episode, the greater likelihood of detection; above the heavy drinking threshold of five or more drinks, detection of drinking was very high (93.0%). Also at the univariate level, lower BMI, alcohol dependence, older bracelet version, higher episode eBAC, and higher number of drinks were significantly associated with episode detection. However, when these variables (except eBAC due to collinearity) were included in a multivariable analysis, only number of drinks remained significant. Therefore, we may conclude that no variables (of those that we explored) other than alcohol consumption are associated with detection using the

[3] Height was missing for one participant so BMI could not be calculated.

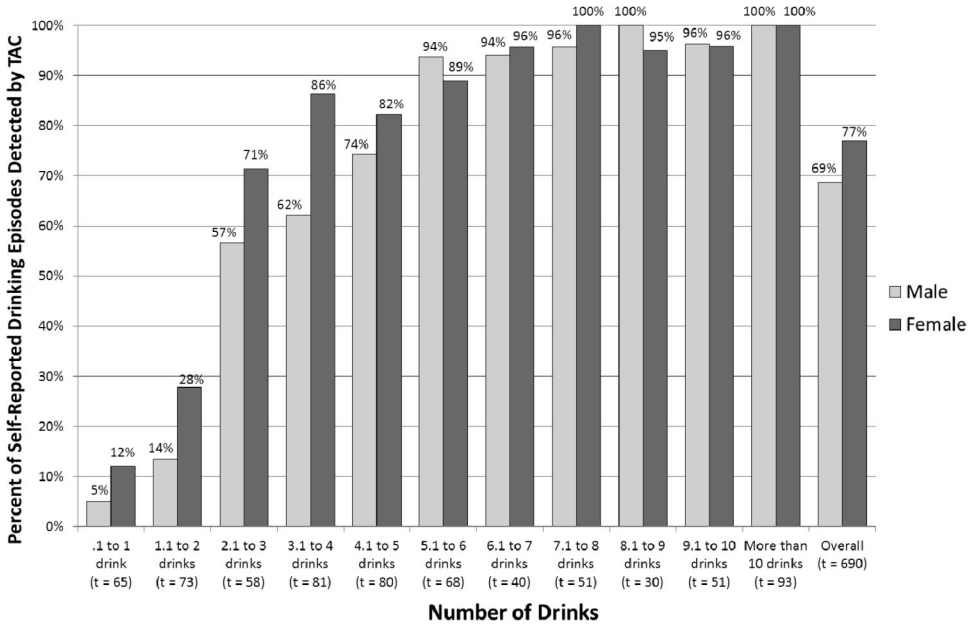

Figure 2. Percentage of 690 self-reported drinking episodes ($N = 66$) detected using transdermal alcohol concentration (TAC) at different numbers of drinks by gender. t = the number of episodes at each level of drinking.

SCRAM. This is a confirmation of the utility of the SCRAM bracelet, as it suggests that the bracelet should not show varying performance for different users.

Initially, we found that at the level of fewer than five drinks, women's episodes were significantly more likely than men's to be detected; indeed, women had higher peak TAC and eBAC than

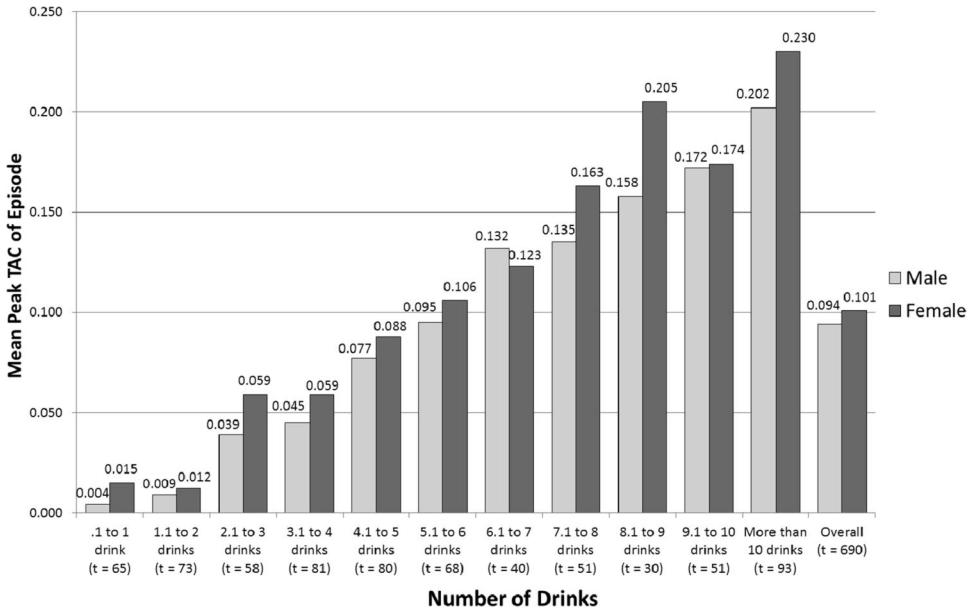

Figure 3. Peak transdermal alcohol concentration (TAC) levels of 690 self-reported drinking episodes at different numbers of self-reported drinks by gender. t = the number of episodes at each level of drinking. All self-reported episodes are included, regardless of whether they were detected using TAC criteria. Self-reported episodes that were not detected using TAC criteria were assigned TAC = .000 g/dl. This allows for the gender comparison on TAC across all self-reported episodes and reflects the lower TAC of men's episodes.

Table 2
Generalized Estimating Equations Analysis of Self-Reported Episode Detection

Variable	*B*	*SE B*	OR	95% CI
Gender × Heavy Drinking	−.20	.56	0.82	[0.28, 2.44]
Body mass index	−.03	.03	0.97	[0.91, 1.03]
Alcohol dependence	.42	.40	1.52	[0.69, 3.32]
Bracelet version	−.41	.39	0.67	[0.31, 1.43]
Estimated BAC	−.56	.62	0.57	[0.17, 1.93]
Number of drinks	3.30	.37	27.19*	[13.11, 56.39]

Note. The dependent variable was coded as self-reported episode not detected using transdermal alcohol concentration (TAC) = 0, self-reported episode detected using TAC = 1. Gender × Heavy Drinking interaction term coded as four or fewer drinks = 0, 5 or more drinks = 1; Alcohol dependence: nondependent = 0, dependent = 1; Bracelet version: SCRAMII = 0, SCRAMx = 1; Number of drinks was square-root transformed because of nonnormality.
* $p < .001$.

men at this lower level of drinking. Higher TAC and eBAC in women at similar levels of consumption would be expected; such differences in measured BAC have been explained by differences in gastric metabolism (Baraona et al., 2001), body composition, and absorption (Thomasson, 1995). The implication for this finding is that the SCRAM appears to be less sensitive to the drinking of men than women when they drink below the heavy drinking threshold, even though no differences in number of drinks were found at this lower level.

The TAC criteria we used to detect self-reported alcohol use were less conservative than those used by AMS to confirm alcohol use among SCRAM wearers, but more conservative than using a simple TAC level of .02 g/dl. In our data, there were seven self-reported episodes that showed an elevation in TAC that reached .02 g/dl or higher, but did not meet absorption and elimination criteria and therefore were not considered detected using our criteria. Generally, a level of .02 g/dl is a reasonable threshold to use; as emphasized by Marques and McKnight (2009), this is a standard used by BrAC and interlock devices for legal purposes. Had we used just the .02 g/dl threshold without the requirement for absorption or elimination, these seven episodes would have been detected, increasing our detection 1%. However, not having the absorption/elimination criterion would also result in an unknown number of TAC episodes that reached .02 g/dl but did not match with a self-reported episode, in effect raising the number of cases in the cell on the flowchart that currently contains 22 TAC detected episodes (i.e., false positives), thereby likely offsetting the more liberal detection criteria with a higher false-positive rate. In addition, 13% of the self-reported episodes did not meet our criteria because they had a TAC level less than .02 g/dl (but greater than .00). Using a lower TAC threshold to detect drinking would also result in a related increase in false positives (i.e., TAC detected alcohol use that did not occur). Dougherty and colleagues (2012) conducted a laboratory study in which they determined that a TAC level of .011 g/dl could distinguish between one drink and more than one drink (when drinking was known to have occurred). Thus, it may be possible to use lower levels of TAC to detect drinking, but whether lower levels can be used while preserving high specificity (i.e., low false positives) has not been determined. There are a number of decisions that must be made when evaluating biosensors, including what detection criteria to use (which may differ depending on the population under study) and what standard to use (i.e., biochemical verification or self-report); these decision rules will obviously influence detection statistics. The focus of this study was not to compare different thresholds and the trade-off between sensitivity and specificity that results; we used the criteria that had performed well in previous work. Future investigations might include comparisons of different detection criteria in different populations.

It should be noted as well that the 22 TAC episodes that did not match with self-report had lower peak TACs than episodes that did match self-report. These may have been actual drinking episodes not reported by participants or the TAC elevations may have been due to environmental alcohol; we are unable to determine which. Alcohol in the environment, including in body care products or in the work environment can result in TAC readings and wearers are instructed about taking care to avoid such exposure. Using a .02 g/dl threshold, Marques and McKnight (2009) reported essentially no false positives, stating that exposure to environmental alcohol produces a different TAC pattern than consumed alcohol, but did not report the number of times this occurred in their sample. These TAC patterns are characterized by a "spike" of TAC that tends to decline much more quickly than consumed alcohol, but may still be mistaken for alcohol ingestion. In addition to their more strict criteria, AMS conducts visual inspections of suspected drinking episodes to eliminate such false positives, and in fact AMS identified only 4 of these 22 episodes as confirmable drinking events. Again, this suggests that the AMS criteria will have lower false positives, but may also miss valid (though lower-level) drinking episodes.

This investigation provides updated information about the reliability of the SCRAM. In earlier reports, failure of the bracelet was moderate and there was some indication that length of wear by the participant was a predictor of sensor failure (Marques & McKnight, 2009). We evaluated two versions of the SCRAM and found malfunctions on at least 1 day for 11 of our original 70 participants (15.7%), for 5.1% of all days of wear, resulting in missing TAC data for 50 of the original 740 (6.8%) self-reported drinking episodes, with no differences between the SCRAM II and SCRAMx. Contrary to our expectations and those of previous research, we did not find that length of wear (either length of total wear by all previous users or by our study participant) was a factor in bracelet performance, suggesting that the earlier problems with loss of accuracy had been resolved in later bracelet versions. We did find in univariate results that the SCRAMII detected a higher proportion of self-reported drinking episodes than the more recent version of the bracelet (SCRAMx) but caution against making too much of this finding, because when included in the multivariable analyses, the bracelet version difference was no longer significant, suggesting that any performance difference in the two versions was negligible.

An estimate of BAC was calculated for each drinking episode, and the weighted correlation ($r = .54$) between this estimate and TAC was considerably smaller than those between TAC and BrAC reported in laboratory investigations using the SCRAM ($r = .84-.91$; Dougherty et al., 2012; Sakai et al., 2006). This is not surprising given that laboratory-based alcohol administration studies measure and control the amount and pace of alcohol consumed

and use a calibrated tool for measuring BrAC, resulting in less measurement error. Furthermore, eBAC is limited because population averages (for metabolic rate and gender-specific body water) are used, and measurement of length of drinking episodes may be inaccurate. In addition, in this study, we combined self-report episodes when they were represented by one TAC curve, which could lead to inaccurate estimates of BAC for these episodes. Nevertheless, we did establish that in univariate analysis the estimate of BAC was a very good predictor of episode detection, but because eBAC had high collinearity with number of drinks, we did not include it in the final analyses. Estimating BAC allowed for comparisons with prior research that used measured BrAC, but this investigation was not designed to provide the optimal alternative physiological measure of BAC to TAC and so estimates of BAC should be interpreted with caution.

There is one methodological issue that warrants commentary: The detection statistics found in this investigation (73% of self-reported episodes) is quite different from information reported in Barnett et al. (2011), in which we reported a sensitivity of .91 (i.e., 91% of self-reports were detected by TAC). The difference in the methods between these two evaluations was that in Barnett et al. we detected drinking *days,* not drinking *episodes*. Any difference in the ability of the sensor to detect drinking is in large part a function of whether or not a drinking episode that begins on one day and that shows a TAC curve that carries over to the next day is counted as drinking on both days. An episode-based determination would count the episode as only occurring on the first day. A day-based determination may result in higher detection rates because of TAC curves carrying over to a second day, primarily because any drinking (later) on the second day would have a chance of being counted as detected from the crossover TAC earlier in the day and/or from detected TAC later in the day. Using self-reported time of drinking would help establish whether drinking actually occurred after midnight on a day, but would not reflect whether an individual was under the influence of alcohol on that second day, which is an important consideration. This research is still in early stages; future work should sort out the best methods for establishing whether drinking occurred on a particular day.

Limitations

The primary limitation of this investigation was that we did not have an objective measure such as BAC or BrAC to use as the standard against which to compare the TAC detection methods. Therefore, we do not know with certainty whether drinking occurred. However, we had daily self-reports that contained information about the timing of drinking episodes, and participants were well prepared to report drinks using standard drink units. The investigations from which the data were derived were intervention studies, so on some days participants may have been motivated to underreport their drinking, in which case the TAC and self-report would not have agreed. However, there was no indication that either stage of the trial (baseline vs. intervention) or the intervention condition (contingency management vs. noncontingent reinforcement) was related to the detection of drinking. Thus, the procedures and findings support the accuracy of participant self-report. Measuring BrAC multiple times daily and during self-reported episodes would have confirmed that drinking did occur, and might have provided some adjustment of the detection statistics (e.g., it might have helped us understand the circumstances when TAC episodes and self-report episodes did not match). However, using such methods increases the burden and expense of such research. As noted above, estimates of BAC are imperfect, and future research is needed to improve the estimation of BAC when objective measures are not available or practical. There are other characteristics that might influence detection properties, including skin characteristics (Hawthorne & Wojcik, 2006; Swift et al., 1992) and rate of consumption (Marques & McKnight, 2009). We did not define the boundaries of a drinking episode a priori but combined episodes that were close together in time; it is possible that a different approach would result in different results. Data for this investigation were from studies conducted with heavy drinkers using a specific alcohol sensor; findings will not be applicable to other alcohol biosensors and may not generalize to other populations.

Real-World Implications

Although the research participants in this study were not mandated to wear the SCRAM, it is important to consider how these research findings are relevant for court-mandated alcohol monitoring. First, false negatives (i.e., the failure to detect a real drinking episode) are more likely to happen when the drinking episode is smaller, whereas missing a drinking episode using the SCRAM when the number of drinks is five or greater is unlikely. That is, heavy drinking episodes, which are more likely to be associated with significant impairment, have a very high likelihood of being detected. Second, it is important to emphasize that false positives (i.e., identifying a drinking episode that did not occur) are likely to be lower in the real world using AMS's system for detecting alcohol use. AMS has more conservative criteria for detecting alcohol use and takes additional steps including visual inspection of suspected drinking events to determine the likelihood that drinking occurred; this more conservative approach has a very low likelihood of false positives. We did not compare our findings with AMS alerts, but in earlier work we found the false-positive rate using AMS criteria was zero (Barnett et al., 2011). Finally, device failure resulting in missing data (which could reflect a possible false negative if drinking occurred and was not detected) was in part due to a temporary malfunctioning component and reflected a small proportion of days and episodes.[4] Of note is that the participants in this study were volunteers with very high rates of compliance; an offender population will have different circumstances, may show different drinking episode characteristics, and is likely to differ considerably in behavioral compliance.

Conclusions

This investigation included a gender-balanced sample of heavy drinking adults who together reported a large number of drinking episodes with good variability in number of drinks per episode. Strengths of the study include high response rates for the daily Web surveys, resulting in very low levels of missing self-report,

[4] We did not record whether AMS identified missing data or malfunctions. In some cases (e.g., when we could see immediately that the data were missing), we notified AMS; in other cases, AMS alerted us that there were anomalies in the data that required us to replace a bracelet.

and a low proportion of TAC episodes that were not matched with a self-reported drinking episode (5.1% of all TAC episodes), indicating that the false-positive rate (or underreporting) was low. Findings established that number of drinks consumed by participants is the primary determinant of detection of drinking episodes, with particularly good detection at the level of five or more drinks and with some gender differences in detection at lower drinking levels.

References

Baraona, E., Abittan, C. S., Dohmen, K., Moretti, M., Pozzato, G., Chayes, Z. W., . . . Lieber, C. S. (2001). Gender differences in pharmacokinetics of alcohol. *Alcoholism: Clinical and Experimental Research, 25,* 502–507. doi:10.1111/j.1530-0277.2001.tb02242.x

Barnett, N. P., Tidey, J., Murphy, J. G., Colby, S. M., & Swift, R. (2013). *A controlled trial of contingency management for alcohol use reduction using a transdermal alcohol sensor*. Manuscript in preparation.

Barnett, N. P., Tidey, J., Murphy, J. G., Swift, R., & Colby, S. M. (2011). Contingency management for alcohol use reduction: A pilot study using a transdermal alcohol sensor. *Drug and Alcohol Dependence, 118,* 391–399. doi:10.1016/j.drugalcdep.2011.04.023

Cherpitel, C. J., Pares, A., Rodes, J., & Rosovsky, H. (1993). Drinking in the injury event: A comparison of emergency room populations in the United States, Mexico, and Spain. *International Journal of the Addictions, 28,* 931–945.

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. doi:10.1037/a0029021

Edenberg, H. J. (2007). The genetics of alcohol metabolism: Role of alcohol dehydrogenase and aldehyde dehydrogenase variants. *Alcohol Research & Health, 30,* 5–13.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2002). *Structured Clinical Interview for DSM–IV–TR Axis I Disorders, Research Version, Patient Edition (SCID-I/P)*. New York, NY: Biometrics Research, New York State Psychiatric Institute.

Hawthorne, J. S., & Wojcik, M. H. (2006). Transdermal alcohol measurement: A review of the literature. *Canadian Society of Forensic Science, 39,* 65–71.

Jones, A. W. (2008). Ultra-rapid rate of ethanol elimination from blood in drunken drivers with extremely high blood-alcohol concentrations. *International Journal of Legal Medicine, 122,* 129–134. doi:10.1007/s00414-007-0181-7

Leffingwell, T. R., Cooney, N. J., Murphy, J. G., Luczak, S., Rosen, G., Dougherty, D. M., & Barnett, N. P. (2013). Continuous objective monitoring of alcohol use: Twenty-first century measurement using transdermal sensors. *Alcoholism: Clinical and Experimental Research, 37,* 16–22. doi:10.1111/j.1530-0277.2012.01869.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. doi:10.1111/j.1530-0277.2008.00887.x

Matthews, D. B., & Miller, W. R. (1979). Estimating blood alcohol concentration: Two computer programs and their applications in therapy and research. *Addictive Behaviors, 4,* 55–60. doi:10.1016/0306-4603(79)90021-2

National Center for Chronic Disease Prevention and Health Promotion. (2012). *Body mass index*. Retrieved from http://www.cdc.gov/healthyweight/assessing/bmi/index.html

Pittman, B., Gueorguieva, R., Krupitsky, E., Rudenko, A. A., Flannery, B. A., & Krystal, J. H. (2007). Multidimensionality of the Alcohol Withdrawal Symptom Checklist: A factor analysis of the Alcohol Withdrawal Symptom Checklist and CIWA-Ar. *Alcoholism: Clinical and Experimental Research, 31,* 612–618. doi:10.1111/j.1530-0277.2007.00345.x

Rehm, J., Greenfield, T. K., & Rogers, J. D. (2001). Average volume of alcohol consumption, patterns of drinking, and all-cause mortality: Results from the U.S. National Alcohol Survey. *American Journal of Epidemiology, 153,* 64–71. doi:10.1093/aje/153.1.64

Rehm, J., Room, R., Graham, K., Monteiro, M., Gmel, G., & Sempos, C. T. (2003). The relationship of average volume of alcohol consumption and patterns of drinking to burden of disease: An overview. *Addiction, 98,* 1209–1228. doi:10.1046/j.1360-0443.2003.00467.x

Sakai, J. T., Mikulich-Gilbertson, S. K., Long, R. J., & Crowley, T. J. (2006). Validity of transdermal alcohol monitoring: Fixed and self-regulated dosing. *Alcoholism: Clinical and Experimental Research, 30,* 26–33. doi:10.1111/j.1530-0277.2006.00004.x

Sobell, L. C., & Sobell, M. B. (1992). Timeline follow-back: A technique for assessing self-reported alcohol consumption. In R. Z. Litten & J. P. Allen (Eds.), *Measuring alcohol consumption: Psychosocial and biochemical methods* (pp. 41–72). Totowa, NJ: Humana Press. doi:10.1007/978-1-4612-0357-5_3

Sobell, L. C., & Sobell, M. B. (1995). *Alcohol Timeline Followback users' manual*. Toronto, ON: Addiction Research Foundation.

Sullivan, J. T., Sykora, K., Schneiderman, J., Naranjo, C. A., & Sellers, E. M. (1989). Assessment of alcohol withdrawal: The revised Clinical Institute Withdrawal Assessment for Alcohol Scale (CIWA-Ar). *British Journal of Addiction, 84,* 1353–1357. doi:10.1111/j.1360-0443.1989.tb00737.x

Swift, R. M. (2003). Direct measurement of alcohol and its metabolites. *Addiction, 98* (Suppl. 2), 73–80. doi:10.1046/j.1359-6357.2003.00605.x

Swift, R. M., Martin, C. S., Swette, L., LaConti, A., & Kackley, N. (1992). Studies on a wearable, electronic, transdermal alcohol sensor. *Alcoholism: Clinical and Experimental Research, 16,* 721–725. doi:10.1111/j.1530-0277.1992.tb00668.x

Thomasson, H. R. (1995). Gender differences in alcohol metabolism: Physiological responses to ethanol. In M. Galanter (Ed.), *Recent developments in alcoholism* (Vol. 12, pp. 163–179). New York, NY: Plenum Press.

U.S. Department of Health and Human Services and U.S. Department of Agriculture. (2005). *Dietary guidelines for Americans* (6th ed.). Washington, DC: U.S. Government Printing Office.

Watson, P. E., Watson, I. D., & Batt, R. D. (1980). Total body water volumes for adult males and females estimated from simple anthropometric measurements. *American Journal of Clinical Nutrition, 33,* 27–39.

Zeger, S. L., & Liang, K. Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics, 42,* 121–130. doi:10.2307/2531248

Received March 1, 2013
Revision received July 29, 2013
Accepted August 21, 2013 ■