



Published in final edited form as:

*Drug Alcohol Depend.* 2011 November 1; 118(2-3): 391–399. doi:10.1016/j.drugalcdep.2011.04.023.

## Contingency Management for Alcohol Use Reduction: A Pilot Study using a Transdermal Alcohol Sensor\*

Nancy P. Barnett<sup>1</sup>, Jennifer Tidey<sup>1</sup>, James G. Murphy<sup>2</sup>, Robert Swift<sup>1,3</sup>, and Suzanne M. Colby<sup>1</sup>

<sup>1</sup>Center for Alcohol and Addiction Studies, Brown University, Box G-S121-5, Providence, RI, 02912

<sup>2</sup>Department of Psychology, University of Memphis, Memphis TN, 38152

<sup>3</sup>Providence VA Medical Center

### Abstract

**Background**—Contingency management (CM) has not been thoroughly evaluated as a treatment for alcohol abuse or dependence, in part because verification of alcohol use reduction requires frequent in-person breath tests. Transdermal alcohol sensors detect alcohol regularly throughout the day, providing remote monitoring and allowing for rapid reinforcement of reductions in use.

**Methods**—The purpose of this study was to evaluate the efficacy of CM for reduction in alcohol use, using a transdermal alcohol sensor to provide a continuous measure of alcohol use. Participants were 13 heavy drinking adults who wore the Secure Continuous Remote Alcohol Monitoring (SCRAM) bracelet for three weeks and provided reports of alcohol and drug use using daily web-based surveys. In Week 1, participants were asked to drink as usual; in Weeks 2 and 3, they were reinforced on an escalating schedule with values ranging from \$5-\$17 per day on days when alcohol use was not reported or detected by the SCRAM.

**Results**—Self-reports of percent days abstinent and drinks per week, and transdermal measures of average and peak transdermal alcohol concentration and area under the curve declined significantly in Weeks 2-3. A nonsignificant but large effect size for reduction in days of tobacco use also was found. An adjustment to the SCRAM criteria for detecting alcohol use provided an accurate but less conservative method for use with non-mandated clients.

**Conclusion**—Results support the efficacy of CM for alcohol use reductions and the feasibility of using transdermal monitoring of alcohol use for clinical purposes.

### Keywords

Contingency Management; Alcohol; Transdermal Monitoring

## 1. Introduction

Contingency Management (CM) is well established as an evidence-based approach for the treatment of substance use disorders (Carroll et al., 2006; Lussier et al., 2006; Petry et al., 2006; Prendergast et al., 2006). Individuals in a CM treatment program typically receive reinforcers such as money or vouchers that are redeemable for retail goods or services

\*supplementary material for this article can be found by entering DOI

Corresponding Author: Nancy P. Barnett, Center for Alcohol and Addiction Studies, Brown University, Box G-S121-5, Providence, RI 02912, Nancy\_Barnett@brown.edu, Telephone: 401-863-6649, Fax: 401-863-6697.

following a period of abstinence that can be biochemically or otherwise objectively verified (Higgins et al., 2002; Silverman et al., 1999). Metabolites of most illicit drugs can be detected in blood or urine for at least several days after the last occasion of drug use, but alcohol can only be detected for several hours after consumption and there are no reliable biomarkers of alcohol use (Kadden, 2001), limiting the application of CM for alcohol treatment.

The few studies that have evaluated the efficacy of CM for alcohol abuse or dependence have used randomly administered or intermittent breath alcohol tests to verify abstinence (e.g., Alessi et al., 2007). Since these tests may not detect some drinking episodes, this approach may reinforce intermittent abstinence. Daily breath tests (see Petry et al., 2000) may detect heavy drinking episodes (or very heavy drinking episodes from the previous day), but participants could avoid detection by timing their drinking, and the requirement of daily breath tests is demanding for participants and time consuming for staff. Thus, the circumstances for evaluating CM for reduction in alcohol use have not been optimal.

### 1.1. Transdermal Alcohol Detection

Transdermal alcohol sensors measure the approximately 1% of ingested alcohol that is excreted through the skin (Swift, 2003), and provide a continuous estimate of blood alcohol concentration (BAC) over extended periods of time (Swift, 1993; Swift et al., 1992). One device, the Secure Continuous Remote Alcohol Monitoring bracelet (SCRAM, Alcohol Monitoring Systems; AMS) has been evaluated in controlled laboratory environments and field studies, and is currently used to monitor court-referred offenders. Sakai and colleagues (Sakai et al., 2006) conducted a laboratory trial with 24 participants who consumed alcohol and were monitored by breath alcohol concentration (BrAC) and transdermal alcohol concentration (TAC) measurements. Correlations between BrAC and TAC were 0.85 for peak alcohol concentration and 0.84 for area under the curve (AUC; which serves as an approximation of volume of alcohol consumed). However, results showed that the peak TAC occurred 2-3 hours after peak BrAC, and tended to underestimate BrAC. Participants in this study also wore the SCRAM for one week, recorded alcohol consumption in a daily log, and took one breath reading daily. Number of standard drinks reported was correlated with mean TAC ( $r = .83$ ) and with transdermal AUC ( $r = .74$ ). As in the laboratory component of this evaluation, transdermal results were not equivalent to BrAC.

Marques and McKnight (2007) also evaluated the SCRAM in laboratory and field conditions. When data for the lab and field trials were combined, results revealed the sensitivity of the SCRAM improved as TAC increased, from 65.3% at .02 g/dL to 86.5% at .08 g/dL. Specificity was 87.7%. Again, TAC curves lagged behind breath alcohol levels due to the longer time required for alcohol to be expelled through perspiration. Similar to Sakai et al. (2006), TAC peak values also tend to be lower than peak BrACs by an average value of .014 g/dL. In debriefing interviews, participants reported initial discomfort with the device but no participants dropped out of the 4-week field trial.

In summary, laboratory and field research has established the validity of transdermal alcohol monitors. Although regularly worn by court-referred offenders, no published reports of their efficacy are available and the SCRAM has not been evaluated as an intervention tool in a voluntary population. Wearers of the available transdermal devices have found them to be reasonably comfortable, though some adjustment to daily activities may be necessary. A noted limitation of the transdermal technology is that individual discrepancies from transdermal TAC to simultaneous breath alcohol measurements vary widely, in large part due to differences in skin properties (Anderson and Hlastala, 2006), which make predicting the actual BAC of an individual from transdermal readings difficult and potentially inaccurate (Sakai et al., 2006). Therefore, TAC readings may be used as a dichotomous

indicator of alcohol consumption, and TAC values themselves are meaningful because they are correlated with BAC/BrAC, but TAC levels are not an estimate of actual BAC.

The objectives of this study were to evaluate the initial efficacy of CM for reducing drinking among heavy drinking adults, and to evaluate the feasibility and acceptability of using a transdermal alcohol sensor in the context of the intervention. This pilot study was necessary to establish the CM reinforcement protocol and data collection components, including remote self-report data collection, prior to conducting a controlled trial. In addition, the approach used by AMS to detect alcohol use in court-referred clients is overly conservative for a voluntary sample, so a final goal was to evaluate an adapted system for detecting alcohol use using SCRAM data. Participants were heavy drinking individuals recruited from the community who wore a SCRAM bracelet for 3 weeks: 1 baseline week followed by 2 weeks during which participants received monetary reinforcement on days when their transdermal alcohol levels were not detected and they did not report drinking. Within-subject analyses were conducted on alcohol use self-reports, transdermal alcohol levels, and other substance use. Acceptability of transdermal monitoring by participants was assessed, and we evaluated the correspondence between participant self report of drinking and the two objective methods (AMS and our adjusted approach) of alcohol detection.

## 2. Methods

### 2.1. Participants

Advertisements were placed in local newspapers, online using Craigslist and Facebook, and on flyers posted in public places inviting adult drinkers who were interested in reducing or stopping drinking to contact the study. A telephone screening procedure established that callers met initial inclusion criteria: 18 years or older, drinking above the national recommendations for alcohol use (U.S. Department of Health and Human Services and U.S. Department of Agriculture, 2005); drinks per week  $\geq 8$  for women,  $\geq 15$  for men), having two or more heavy drinking episodes per week, and being interested in cutting down or stopping drinking. Participants also had to have daily access to the internet and either have an acceptable landline phone or be willing to come to the research office three times per week for bracelet data downloads. Participants were excluded if they reported using drugs other than marijuana in the past month or more often than once a month in the past year, or had a score  $\geq 23$  on the Alcohol Withdrawal Symptom Checklist (Pittman et al., 2007). Since this was a feasibility and early efficacy study, individuals who were seeking treatment were provided referrals.

### 2.2. Procedures

All procedures were approved by the Brown University Institutional Review Board and participants gave written informed consent.

**2.2.1. Baseline screen**—Prior to enrolling in the study, participants completed an in-person screening, including a urine toxicological screen for drugs other than marijuana. The Clinical Institute Withdrawal Assessment for Alcohol–Revised (CIWA-AR; Sullivan et al., 1989) was administered to verify the participant was not in alcohol withdrawal. Breath alcohol was tested using the Digital Alcohol Detector (BTNX) to verify a .000 g/dL alcohol level, which is required for SCRAM bracelet installation.

**2.2.2. Baseline session and assessments**—The initial session consisted of administering baseline measures, installing the bracelet, and instructing the participant about the project procedures. The Structured Clinical Interview for DSM-IV, Substance Abuse Module (First et al., 1996) and a 30-day Timeline Followback (Sobell and Sobell, 1992,

1995) were interviewer administered. Measures of lifetime and past-month substance and tobacco use, and prior treatment experiences were self-administered. For all assessments, standard drink equivalents were used (12 oz. beer or wine cooler, 5 oz. glass of wine, 1.5 oz. of liquor).

**2.2.3. Bracelet description**—The SCRAM bracelet was installed by the research staff person who conducted the baseline session. The bracelet weighs approximately eight ounces and is locked to the ankle and worn continuously. The device contains three sensors, an electrochemical alcohol sensor that samples the ethanol vapor close to the skin and two sensors that detect changes in skin temperature and skin reflectivity (used to detect attempts to tamper with the sensor and to ensure the bracelet is not removed). The bracelet sensors take readings every 30 minutes, and once a day the bracelet transfers the readings via a wireless radio frequency signal to a modem installed in the wearer's home. Data are then transferred from the modem through the landline telephone line to AMS; the data are immediately available on a password-protected web site that can be accessed by research staff using a standard web browser. Staff can also download bracelet data in person using the AMS DirectConnect device, which attaches to the bracelet and downloads readings via a computer USB port. AMS staff make determinations and provide alerts about alcohol consumption on a daily basis on the secure web site and in email notifications. Tamper alerts are produced if the readings reflect possible interference or if a break in continuity occurs. The SCRAM bracelet was chosen over another available alcohol sensor (WrisTAS, Giner, Inc.; worn on the wrist) because the SCRAM has a lower failure rate (Marques and McKnight, 2009), cannot be removed by the wearer without cutting the strap or breaking the closure clip (allowing us to verify that the bracelet stayed on the participant), and has wireless and modem transmission technology. The SCRAM is also waterproof so can be worn continuously; the WrisTAS must be removed for showering.

**2.2.4. Daily web log**—Each morning during the three weeks of the trial, participants received an email from the project that contained a link to a secure survey. The survey contained questions about whether alcohol was consumed, and if so the number of drinks, and whether tobacco, marijuana, or other drugs were used on the previous day. The survey took less than five minutes to complete; participants received \$5 for each completed survey and a bonus of \$25 for completing 19 of the 21 daily surveys (90%) on the day they were due. Self-report indices derived from these reports were: 1) percent days self-report abstinent, 2) drinks per week, 3) drinks per drinking day, 4) percent days of cigarette use, 5) percent days of marijuana use, and 6) drug use other than marijuana.

**2.2.5. Criteria for detecting drinking with the transdermal bracelet**—The primary AMS criteria for detecting a drinking episode that we adapted for our use are: 1) at least three TAC readings over .02 g/dL, and 2) absorption rate for the episode < .05 g/dL per hour, and 3) elimination rate for the episode < .025 g/dL per hour (when peak < 0.15 g/dL) and less than 0.035 g/dL per hour (when peak > 0.15 g/dL).<sup>1</sup> AMS primarily serves clients in the criminal justice system, so these criteria for establishing a drinking event are intentionally conservative. We were interested in developing and evaluating less stringent criteria, which would possibly allow us to detect lower levels of consumption. We also wanted to determine whether a simpler set of criteria using only information available to us on the AMS user webpage (i.e., date, time, and TAC) was sufficient for our needs. We therefore modified the AMS criteria so we considered drinking to be detected by the bracelet

---

<sup>1</sup>AMS also uses data from the contact and temperature sensors, historical models, human analysts, and information from their providers (e.g., parole officers) to confirm a drinking event.

if: 1) at least one TAC reading was over .02 g/dL (not three as in the AMS system), and 2) one of the AMS absorption or elimination rate criteria above was met (AMS requires both).

**2.2.6. Bracelet data review**—Bracelet data that is transferred via modem or in person is available immediately on the AMS website for review. Every day during data collection available transdermal alcohol data were independently reviewed using the above criteria by two project staff who were blind to participant self report; ratings were immediately entered into a MS Access database. Rater agreement was excellent (Kappa = .99). Consensus ratings were made by consulting other staff when the two primary raters disagreed. For participants with modems, data were available to rate every day (for the previous day since a full day of data needed to be collected before determinations could be made), and approximately every other day for DirectConnect participants.

**2.2.7. Intervention**—In the first week, there were no contingencies for alcohol use. On the first day of Week 2 (always a Monday), participants started to receive internet-based monetary reinforcement for each day on which no self-reported drinking was reported on the web log or detected with the bracelet (hereafter called “not detected” [ND]).<sup>2</sup> The reinforcement amount for a day on which alcohol was not detected started at \$5 and increased for each consecutive ND day by \$2, for a maximum daily reinforcement amount of \$17 (on Sunday). On days when alcohol was detected, the participant received \$0, and the reinforcement amount was re-set to \$5 on the next ND day. This type of escalating schedule with re-set has been established as the most effective schedule for reinforcing drug abstinence (Higgins et al., 1994; Roll et al., 1996). The schedule also was designed to escalate so the highest reinforcer values occurred on weekends (assuming a participant did not drink prior to the weekend), when we expected that avoiding alcohol consumption would be more difficult. On the first day of Week 3 (again a Monday), the reinforcer amount was re-set to \$5. The most a participant could receive in reinforcers was \$154 total for Weeks 2-3.

Following completion of each daily web log (i.e., when the participant clicked “submit” on the self-report survey), ASP.net code queried our rater database and the self-report survey database, determined whether alcohol use had been detected by the raters or reported by the participant, and immediately provided the determination (alcohol detected or not detected) and the reinforcer amount on a web page to the participant. This web page also showed whether alcohol had been detected or not for all previous days, reinforcement amounts for each day, and the participant's total monetary balance. Information about whether alcohol was detected was not presented until after the self-report survey (reporting use for the previous day) was complete and bracelet data from that day had been rated. Actual transdermal alcohol levels were not provided to participants, only whether alcohol use had been detected. Participants also could log in to the project web site at any time using a unique username and password to see this report. During intervention weeks, staff checked with participants to confirm they were viewing the reinforcement reports. Participants could choose to receive payments in cash or a mailed money order at any time during the trial, or could wait until the end of the intervention phase to receive cash payments.

**2.2.8. Participant instructions**—At the baseline session, participants were instructed on standard drink calculations and were assisted in converting their favorite drinks to standard drink units. They were also provided handouts about the daily web logs, the reinforcement schedule, and the bracelet. Participants were informed that they could shower and wash their

---

<sup>2</sup>We considered not including participant report in the final determination, but this would have resulted in sometimes providing reinforcement when known drinking occurred.



leg around the bracelet, but the bracelet should not be immersed in water. Participants were told to avoid using household products that contained alcohol as this could result in positive alcohol readings. Participants who were given a modem were provided instructions on how to install it, and were scheduled for a daily automated download connection from the bracelet to the modem. Participants were told that any drinking might be detected, and that drinking on one day might be detected on the following day, and might result in more than one day of drinking being detected. All project materials were available to participants on the project website.

**2.2.9. Post-intervention interview**—Following the 3-week monitoring period, participants completed an in-person interview that collected information about their experiences wearing the bracelet and receiving reinforcement. Participants were asked, “Did you have any marks on your skin from the bracelet?” Additional questions measured physical and social comfort (“How physically comfortable was the bracelet?”; “How embarrassing [socially uncomfortable] was it to wear the bracelet?”), and the extent to which the bracelet interfered with work, exercise, enjoyment of life, and mood. These items were measured on a 10-point scale from 1 (low) to 10 (high). Participants also were asked: “On average, how often did you notice the bracelet while you were wearing it over the course of a day?” and “Each night, how often did you notice the bracelet while you were sleeping?” The scale for these items was 0 (never or almost never) to 5 (several times per hour).

### 2.3. Data Analyses

Self-report daily log data were summarized by week to produce percent days no alcohol use reported, number of drinks per week, drinks per drinking day, and percent days used cigarettes, marijuana, and other drugs. Staff consensus ratings and determinations from AMS for each day of monitoring (bracelet alcohol detected/not detected) were produced for each participant. We processed the transdermal alcohol readings for analysis with the use of a macro developed in our laboratory. The input to the macro was a raw Excel data file downloaded from AMS for each participant that contained date- and time-labeled TAC readings from every sample taken by the bracelet. The macro produced daily variables, including average and peak TAC, and area under the curve calculated using a formula that sums the area of trapezoids under the curve (Williamson, 2001). From these data, we calculated weekly average TAC, weekly peak TAC, and weekly sum of AUC. TAC values that were not part of a detected episode were recoded as 0 for calculating daily and weekly averages and AUC.

Paired t-tests were used to compare Week 1 to Weeks 2-3 on self-report and transdermal data. The Weeks 2 and 3 variables were averaged to create a “per week” value that was compared to Week 1.<sup>3</sup> The exception was peak TAC, for which the highest TAC value detected across Weeks 2-3 for each participant was used. Within-group effect sizes ( $d$ ) were calculated, where  $d = t/\sqrt{2(1-r)/n}$ .

## 3. Results

We enrolled 20 participants (55% female). Four showed very little drinking in the first week; two had no heavy drinking days and two had one heavy drinking day, so were removed. Two participants withdrew in the first week due to negative attention about the bracelet, and one participant cut off the bracelet and could never be reached again. A total of

---

<sup>3</sup>Weeks 2 and 3 were combined after determining that there were minimal differences between Weeks 2 and 3. There was some evidence of lower consumption in Week 2 than in Week 3, but for all outcome variables, results were identical for the individual weeks and the aggregated weeks.

13 participants completed the 3 weeks of the trial. The three participants who withdrew or could not be reached showed approximately the same number of past-month drinking days at baseline (22.3) as those who completed the trial (21.1),  $t(14) = 0.27$ , *ns*.

The final sample was 46% female, with an average age of 32 years ( $SD = 9.9$ ). Of the 13, 9 (69.2%) were white, 2 (15.4%) were black, 1 (7.7%) was Latino, and 1 (7.7%) was multiracial. All participants had a high school education or equivalent. Most ( $n = 9$ ; 69.2%) had a lifetime diagnosis of alcohol dependence, with an additional 15% ( $n = 2$ ) meeting lifetime diagnostic criteria for alcohol abuse. At baseline, participants reported drinking alcohol on 70.3% of days in the past month ( $SD = 25.8$ ), with an average of 5.9 ( $SD = 1.3$ ) drinks per drinking day, almost half (46.2%) reported using marijuana in the past month, and 53.8% reported smoking cigarettes in the past month. None of the participants reported having received substance abuse treatment, including pharmacotherapy, in the past year.

### 3.1. Detection of Self-reported Drinking Days

Daily web logs were collected on 270 of the 271 days (99.6%) participants wore the bracelet.<sup>4</sup> Participants reported drinking on 106 of those days (39.3%) and our rating system detected 93 of the 106 days (87.7%) participants reported drinking. Of the 13 days of self-reported drinking that were not detected by our rating system, 11 had transdermal readings that did not reach .02 g/dL.<sup>5</sup> When the consumption reported by participants on these 11 days was reviewed, the average number of drinks consumed was 3.3 ( $SD = 2.0$ ), approximately half as many as on other drinking days in Week 1. The other two days of self-reported consumption that were not detected by our rating system were days on which the participant reported drinking just before midnight on one day, but due to the delay between consumption and detection of transdermal alcohol, the transdermal levels were not observed until after midnight.

There were 10 other days (3.7% of self-report days) on which we detected drinking when the participant did not report drinking. There are two possible reasons for this: participants were not telling the truth, or we detected drinking when none occurred. At the post-intervention interview, participants were asked about these discrepant days. Of the 10 discrepant days, six were from one participant who at the post-intervention interview admitted that he had lied about drinking on those days.

**3.1.1. Concordance between ratings and self-report of alcohol use**—Our rating system and the AMS alert system are compared with participant self report in Table 1. The tradeoff between using less or more stringent criteria is reflected in the cells where self-report and ratings disagree. Using our rating system, we were more likely than AMS to detect drinking when the participant reported none occurred (i.e., false-positives). Error for AMS is greatest in the cell where participant report was “yes” and AMS rated the day as “no” (i.e., false-negatives).

### 3.2. Survey Payments and Reinforcement Received

Participants received an average of \$86.77 in payments for negative readings in Weeks 2-3 (56.3% of possible amount; range 0-100%) and all reinforcement earned was redeemed. Participants received an average of \$103.84 for completing the daily web surveys, and all but one participant ( $n = 12$ ; 92.3%) received the good responding bonus.

<sup>4</sup>Two participants wore the bracelet for only six days in the baseline week.

<sup>5</sup>We cannot verify whether drinking did or did not occur on these days, however there was no incentive for reporting drinking when it did not actually occur, so we assume participants who reported drinking were telling the truth.

### 3.3. Alcohol Use

Percent days of self-reported abstinence increased from 23.3% ( $SD = 22.2$ ) in Week 1 to 82.4% ( $SD = 23.4$ ) in Week 2, and 75.8% ( $SD = 32.7$ ) in Week 3, a significant increase from the baseline to CM weeks,  $t(12) = 7.79, p < .001 (d = 2.30)$ . When our complete detection criteria were used (i.e., combining self-report and transdermal ratings), percent days ND in Week 1 was 8.8% ( $SD = 13.7$ ), 69.2% ( $SD = 31.9$ ) in Week 2, and 65.9% ( $SD = 29.5$ ) in Week 3.<sup>6</sup> The increase in percent days ND from the baseline to CM weeks was significant,  $t(12) = 8.28, p < .001 (d = 2.46)$ . The percent of the sample that had drinking detected using our complete criteria on each day is shown in Figure 1. The average number of consecutive ND days was 7.4 ( $SD = 4.3$ ; Median = 7; range 0-14). In all three weeks, weekend days had the highest drinking (i.e., lowest rates of ND).

Total number of self-reported drinks consumed per week declined from an average of 40.0 ( $SD = 21.7$ ) in Week 1 to 11.1 ( $SD = 15.0$ ) in Weeks 2-3,  $t(12) = 6.81, p < .001 (d = 1.44)$ . This reflected an average of 78.8% reduction in volume of alcohol consumed from Week 1 to Weeks 2-3, and 9 of the 13 participants (69.2%) reduced their drinking below the nationally recommended weekly limit for their gender. Drinks per drinking day reduced from 8.1 drinks ( $SD = 3.6$ ) in Week 1 to 6.9 ( $SD = 3.9$ ) in Weeks 2-3,  $t(8) = 2.10, ns (d = 0.34)$ .

### 3.4. Transdermal Readings

A total of 271 days of transdermal readings were collected (per person  $M = 20.8, SD = 0.4$ ), for 12,744 readings overall ( $M = 980$ ; range 941 – 1005). The average TAC on all days was .044 g/dL ( $SD = .047$ ) in Week 1 and .011 g/dL ( $SD = .015$ ) in Weeks 2-3. The decrease from baseline to intervention weeks was significant,  $t(12) = 3.25, p < .01 (d = 0.55)$ . The average reduction in TAC was 72% ( $SD = 28$ ). Ten participants (76.9%) showed a 50% or greater reduction in their average TAC. We also reviewed the data for days on which drinking was detected; the average TAC on these days in Week 1 was .056 g/dL ( $SD = .047$ ) and Weeks 2-3 was .030 g/dL ( $SD = .023$ ), a significant decrease,  $t(10)^7 = 2.46, p < .05 (d = 0.58)$ . The highest TAC reached on average in Week 1 was .252 g/dL ( $SD = .119$ ), and on Weeks 2-3 was .158 g/dL ( $SD = .123$ ), a significant reduction,  $t(12) = 3.32, p < .01 (d = 0.78)$ . Figure 2 reflects peak and average TAC over the 21 days of the trial. Area under the curve was highly correlated with the total volume of self-reported alcohol consumed per week ( $r = .79-.94; ps < .01$ ), and reduced significantly from Week 1 to Weeks 2-3,  $t(12) = 3.58, p < .01 (d = 0.62)$ .

**3.4.1. Individual transdermal results**—Different responses to the reinforcement weeks are reflected in the examples in Figure 3. The participant in Panel A was highly responsive, showing virtually no consumption in Weeks 2-3. The participant in Panel B showed an immediate response but some consumption returned in Week 3, and the participant in Panel C showed a reduction in use that is evident in the TAC readings, but alcohol use was detected on every day so this individual received no reinforcement.<sup>8</sup>

### 3.5. Cigarette, Marijuana, and Other Drug Use

Seven participants reported smoking cigarettes on the daily logs in the baseline week (Week 1); one more than had reported smoking at baseline. Percent days of self-reported abstinence from cigarettes among those who smoked in Week 1 was 8.2% in Week 1 ( $SD = 7.6$ ; range

<sup>6</sup>One of the differences between self-reported abstinence and our combined system is that one episode of self-reported drinking may have resulted in more than one day of transdermal levels.

<sup>7</sup>Two cases were not included in this analysis because they had no detected drinking episodes in Weeks 2-3.

<sup>8</sup>Please see supplementary materials for similar data from all participants.



0-14.3) and 28.6% in Weeks 2-3 ( $SD = 29.5$ ; range 0-71.4),  $t(6) = 1.84$ ,  $ns$  ( $d = 0.91$ ). Among those who used any marijuana during the weeks of the trial ( $n = 7$ ), percent days used marijuana was 22.4% in Week 1 ( $SD = 24.6$ ) and 27.6% in Weeks 2-3 ( $SD = 37.4$ ), a nonsignificant change,  $t(6) = 0.57$  ( $d = 0.14$ ). No other drug use was reported in any of the weeks.

### 3.6. Participant Reports about Bracelet Wear and Online Reporting

Participants reported moderate physical discomfort ( $M = 4.8$ ;  $SD = 1.9$ ; range 3-9) and embarrassment about the bracelet ( $M = 6.3$ ;  $SD = 2.7$ ; range 1-10). Although no serious bruising or irritation was noted upon bracelet removal, 61.5% of participants reported marks on their skin from the bracelet. Participants reported noticing the bracelet over the course of the day on average every hour ( $M = 3.1$ ;  $SD = 1.4$ ; range 1-5), and one to two times each night while sleeping ( $M = 1.2$ ;  $SD = 1.2$ ; range 0-4). Interference with daily activities was lowest for normal work ( $M = 2.2$ ;  $SD = 1.9$ ) and highest for exercise ( $M = 4.3$ ;  $SD = 3.5$ ).

Participants provided positive reports about the online self-report data collection process, stating that surveys were easily accessed and completed by computers or Smartphones. Participants appreciated receiving information about their transdermal levels on a daily basis, with several suggesting that they would have liked to have more detail about their alcohol levels (rather than just the report that alcohol use was detected or not).

### 3.7. Reactivity to Wearing the Bracelet

Wearing the bracelet prior to the intervention week did not appear to affect drinking significantly in Week 1. On the 30-day TLFB at baseline participants reported drinking on 70.3% of days; during Week 1 while wearing the bracelet, participants reported drinking on 76.7% of days, a nonsignificant difference,  $t(12) = 1.49$  ( $d = 0.26$ ).

### 3.8. Data Downloads

Six of the participants (46.2%) were able to use the modem system to transmit their data; the rest came to the research laboratory to download data directly three times per week. Those who were not able to use the modem system did not have a landline telephone or had a telephone service provider that did not work with the modem. With one modem participant, problems with data transmission occurred because the modem appeared not to be working consistently.

## 4. Discussion

This pilot study evaluated the efficacy of contingency management for alcohol use reductions among heavy drinking adults using a transdermal device originally developed for use with problem drinkers who are ordered by a court to abstain from alcohol. Heavy drinking participants reduced their alcohol use substantially in the weeks when contingent reinforcement was provided. The within-subjects reductions were found for self-reported days and the combined system of self-report and transdermal ratings, on drinks per week, and on transdermal measures reflecting average alcohol level, peak alcohol level, and weekly volume of consumption. These reductions were statistically significant, even in this small sample, and effect sizes were in the medium to large range, with the majority reducing their weekly drinking below national recommendations. Despite intentionally providing the highest reinforcer values for weekend days (assuming continuous ND), weekends had the lowest ND rates. Reinforcement schedules may require adjustment to determine the value that is most effective, particularly on historically heavier drinking days.

We found a nonsignificant change in drinks per drinking day with a small effect size, which indicates that participants did not make an effort to reduce their volume of drinking on days when they decided to drink. This may indicate that competing reinforcers (such as social reinforcement for drinking) outweighed the value of the project reinforcers, which would explain the higher levels of drinking on weekends as well, or alternately, that once drinking was initiated, abstinence plans had failed (i.e., the abstinence violation effect; Marlatt, 1985). However, the average TAC on drinking days and the peak TAC declined significantly, with moderate to large effect sizes, which provides some evidence of changes in drinking behavior on drinking days. In other words, the differences in average and peak alcohol levels on drinking days appear to be greater than the difference in self-reported number of drinks, which suggests possible harm reduction efforts on drinking days.

There were no contingencies for cigarette or other drug use and significant changes in marijuana, tobacco, or other drugs were not found. However, the nonsignificant increase in days abstinent from cigarettes had a large effect size, so might be significant in a larger sample. It should be noted however, that heavy drug users were screened out of this sample, the subsamples of marijuana users and cigarette smokers were very small, and biological confirmation was not conducted.

One of our primary objectives was to determine whether adjusting the AMS criteria for detection of alcohol would result in acceptable detection statistics. We established that when an episode reached .02 g/dL TAC once, and met either the AMS absorption *or* elimination criteria, sensitivity and specificity (relative to self-report) were good. For this population, our adjustment was an improvement over using AMS alerts since it resulted in fewer false-negatives (i.e., missed episodes of reported drinking). If we did not have (or trust) self report, the adjusted rating system would have missed 9.5% of drinking days; if we had relied on the AMS system we would have missed (and erroneously provided reinforcement on) 24.1% of drinking days. However, using our less conservative rating system required daily review of data, which is more time consuming than using AMS determinations.

We considered participant self-report in our determination of whether to provide the reinforcer. This was an important decision point, and an argument could be made that reinforcement should only be provided when a substance is objectively detected. However, this would have resulted in knowingly providing reinforcement when alcohol use had been reported, which is not consistent with behavioral reinforcement principles. Contingency management protocols in treatment settings have previously ignored self-report in the determination of reinforcement (e.g., Petry et al., 2000); as described above, this would have resulted in 9.5% of known drinking days being reinforced.

We determined that one participant intentionally reported no drinking when he had indeed drank. This individual admitted that he had hoped his drinking would not be detected so he would receive the reinforcement. It is possible that other participants underreported drinking as well, which would compromise the basis of our detection statistics (i.e., because they were established relative to self-report), but would not affect the CM findings. Indeed, one would expect that individuals who are being reinforced for not using substances would at times underreport their use, underscoring the importance of using objective measures.

Problems that participants noted with the bracelet were moderate and were primarily related to embarrassment and physical discomfort. Both participants who withdrew from the trial had experienced negative attention from the bracelet. Marks on the skin and interference with physical activity due to irritation were commonly reported, but none prompted complaints or removal. Nevertheless, wearability issues may limit the utility of similar transdermal devices for voluntary populations. Prior reports with the SCRAM have noted

water accumulation problems which caused a decline in accuracy over time (Marques and McKnight, 2009). We noted no changes in accuracy over the course of the study but participants wore the bracelet for only three weeks, so our study could not adequately evaluate this issue.

One of the strengths of the AMS system is that transdermal data can be transmitted via modem and reviewed daily, allowing for regular reinforcement of the targeted behavior. The self-report data and all readings for a day must be available before a determination can be made about alcohol consumption, so the soonest a reinforcer could be provided is the following day. Using the bracelet data and remote transmission allowed us to provide reinforcement more quickly than most clinical applications of CM. Although more than half of the participants did not use the modem system, we were able to download bracelet data in person from these participants and provide a report approximately every second day.

#### 4.1. Limitations

The sample was small and the length of intervention was short, with no randomization to conditions, no control group, and no follow-up assessment after intervention completion. To minimize the risk that drug use would increase upon the reduction in alcohol use, we excluded regular users of substances other than marijuana. We included participants who were interested in cutting down or stopping alcohol use, but excluded those who were seeking immediate treatment, therefore our participants most likely did not reflect the most severe end of the continuum of treatment-seeking individuals.

Transdermal devices do not detect all alcohol use. Indeed, about 1 in 12 self-reported days of alcohol use were not detected by the transdermal monitor because the TAC was below the minimum of .02 g/dL. This is a lower false-negative rate than reported by others (Marques and McKnight, 2009), which could reflect differences in the standard (ours was self-report, Marques and McKnight used BrAC), or could indicate that ours was a heavier drinking sample (resulting in fewer low drinking episodes and making detection by the bracelet more likely), and/or that the SCRAM device has improved in sensitivity. Using a .02 g/dL TAC level as the minimum for alcohol use detection is the standard used by AMS and other evaluations of the transdermal devices (Marques and McKnight, 2009), and we also determined in earlier piloting work that interrater reliability of transdermal data was poor when episodes did not reach .02 g/dL. Although we were not able to verify absolute abstinence, the adjusted rating system did detect all but the very lowest levels of drinking, and other CM interventions that use urine tests to detect substance use have similar limitations (i.e., depending on timing of the test, may not detect episodes of use).

There were days that we detected a positive TAC when participants reported not drinking; in most cases, this was due to TAC extending into the next calendar day. For these days, we retained the positive rating for the second day because it is impossible from the TAC readings to determine exactly when alcohol was consumed, and because there was some likelihood that participants had consumed alcohol past midnight and/or were under the influence of alcohol on that second day. We did explain to participants that drinking on one day might be detected on the following day. If anything, this approach provides a conservative indication of the efficacy of CM on TAC levels, as one drinking day may result in two TAC positive days (and no reinforcement on either day). It may be advisable for future studies to consider shifting the hours of a “day” to reflect the hours that an individual may still be affected by alcohol (e.g., shift the “day” hours to be 6 AM – 6 AM so drinking that occurs on one night and carries over to the next morning is counted as one day). An alternative would be to assign drinking “episodes” to the day when the episode was first detected, though this solution would not be satisfactory for episodes that continue for more than a day.

## 4.2. Future Research and Clinical Implications

Controlled experimental designs with larger samples and a longer reinforcement phase are needed to determine whether CM for alcohol use reductions using transdermal alcohol sensors is an efficacious intervention approach. In addition, the maintenance of intervention gains needs to be established using post-intervention follow-up assessments. As with other substance abuse treatment (Carroll et al., 2006; Carroll and Rounsaville, 2007), CM for alcohol use disorders with voluntary patients may be most effective when combined with other treatment approaches, but this too requires further study. Measuring possible complementary benefits of this alcohol intervention (i.e., a decrease in tobacco) and possible substitute use of other substances is important as well.

We found considerable variability in alcohol use during the reinforcement phase. Some participants showed dramatic reductions, some showed very little behavior change, and some participants showed reductions in alcohol use that were not reinforced because they did not reduce their TAC levels below our targeted criterion. Future research could examine whether using a reduction or shaping phase would be effective with individuals who are unable or unwilling to maintain continuous drinking reductions. Modifying the reinforcement schedule to a variable ratio schedule might require fewer resources, but a fixed schedule as used in the present study is likely to have the most rapid effect, which is desirable in individuals who are initiating a reduction in use.

Using the SCRAM to detect alcohol use addressed some of the limitations of using daily breath tests for CM interventions: it did not require the staff and participants to meet daily to collect the breath readings, and it measured alcohol use approximately every 30 minutes, so participants could not avoid detection by timing their drinking. The SCRAM readings also provide information about the volume of alcohol consumed (using the area under the transdermal curve) which breath tests cannot, and the average and peak transdermal alcohol levels provide considerably more information than a once-daily breath test. However, like breath tests, the SCRAM can miss low levels of alcohol use, and the demand on staff and participants is still considerable, the bracelets and monitoring are expensive, and data management is more extensive. There have not been comparisons of the sensitivity of breath tests and the SCRAM to detect alcohol use in a clinical population; such a comparison would help establish the costs and benefits of the two approaches.

Participants in our study reported new levels of insights about their drinking and curiosity about their TAC levels. Some expressed frustration about not being able to reduce their alcohol use as much as they wanted. These reactions present opportunities for counseling about behavior change. Treatment providers could use the objective data from the sensors to develop or adjust treatment plans, provide feedback to clients, and measure outcomes. Participants in our study were not concerned about being monitored, but stigma related to the bracelet's appearance may limit acceptance by clients, and costs may limit adoption in clinical settings.

In summary, this pilot trial showed the feasibility and efficacy of CM for alcohol use reduction using transdermal technology to measure alcohol use remotely and continuously. It extends the potential utility of CM for alcohol dependence and suggests the potential of transdermal monitors for clinical use.

## Supplementary Material

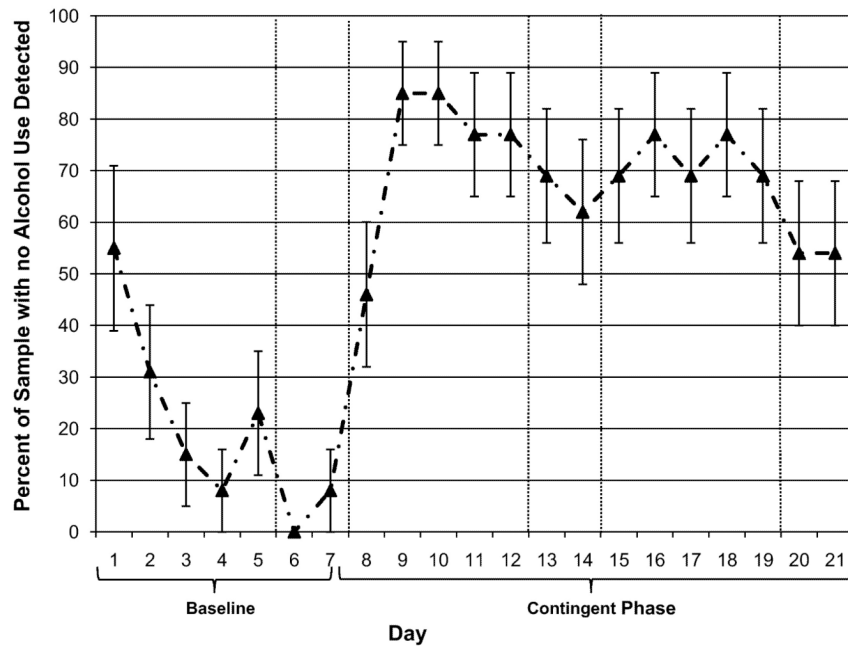
Refer to Web version on PubMed Central for supplementary material.

## References

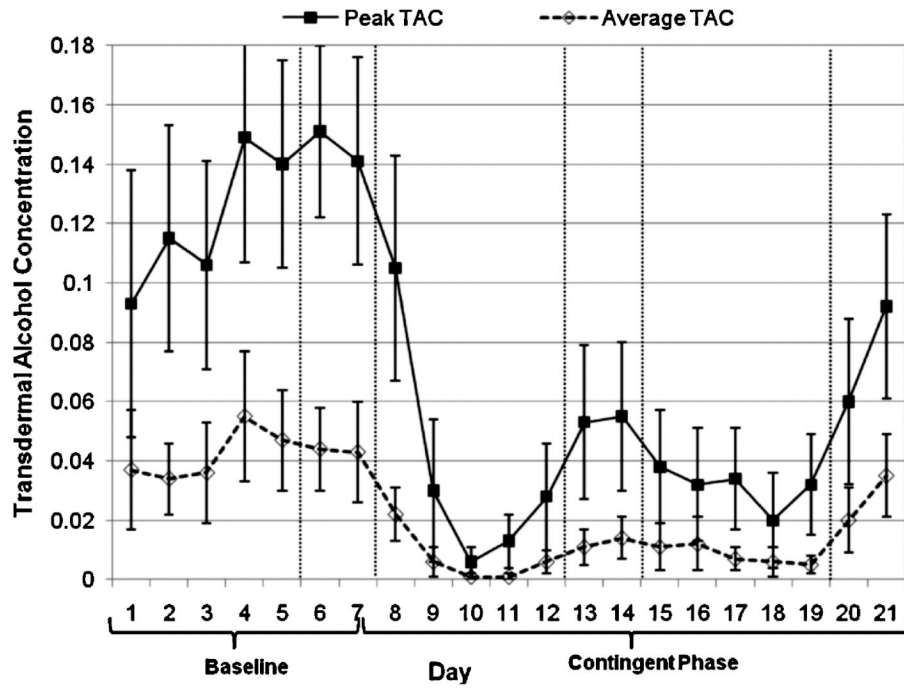
- Alessi SM, Hanson T, Wieners M, Petry NM. Low-cost contingency management in community clinics: delivering incentives partially in group therapy. *Exp Clin Psychopharmacol.* 2007; 15:293–300. [PubMed: 17563216]
- Anderson JC, Hlastala MP. The kinetics of transdermal ethanol exchange. *J Appl Physiol.* 2006; 100:649–655. [PubMed: 16239611]
- Carroll KM, Easton CJ, Nich C, Hunkele KA, Neavins TM, Sinha R, Ford HL, Vitolo SA, Doebrick CA, Rounsaville BJ. The use of contingency management and motivational/skills-building therapy to treat young adults with marijuana dependence. *J Consult Clin Psychol.* 2006; 74:955–966. [PubMed: 17032099]
- Carroll KM, Rounsaville BJ. A perfect platform: combining contingency management with medications for drug abuse. *Am J Drug Alcohol Abuse.* 2007; 33:343–365. [PubMed: 17613963]
- First, MB.; Gibbon, M.; Spitzer, RL.; Williams, JBW. *User's Guide for the Structured Clinical Interview for DSM-IV Axis I Disorders-Research Version.* Biometrics Research Department, New York State Psychiatric Institute; New York: 1996.
- Higgins ST, Alessi SM, Dantona RL. Voucher-based incentives. A substance abuse treatment innovation. *Addict Behav.* 2002; 27:887–910. [PubMed: 12369474]
- Higgins ST, Budney AJ, Bickel WK, Foerg FE, Donham R, Badger GJ. Incentives improve outcome in outpatient behavioral treatment of cocaine dependence. *Arch Gen Psychiatry.* 1994; 51:586–576.
- Kadden RM. Behavioral and cognitive-behavioral treatments for alcoholism: research opportunities. *Addict Behav.* 2001; 26:489–507. [PubMed: 11456073]
- Lussier JP, Heil SH, Mongeon JA, Badger GJ, Higgins ST. A meta-analysis of voucher-based reinforcement therapy for substance use disorders. *Addiction.* 2006; 101:192–203. [PubMed: 16445548]
- Marlatt, GA. Cognitive factors in the relapse process. In: Marlatt, GA.; Gordon, JR., editors. *Relapse Prevention: Maintenance Strategies in the Treatment of Addictive Behaviors.* Guilford; New York: 1985.
- Marques, PR.; McKnight, AS. *Evaluating transdermal alcohol measuring devices (DOT HS 810 875).* National Highway Traffic Safety Administration; Washington, DC: 2007.
- Marques PR, McKnight AS. Field and laboratory alcohol detection with 2 types of transdermal devices. *Alcohol Clin Exp Res.* 2009; 33:703–711. [PubMed: 19170663]
- Petry NM, Alessi SM, Carroll KM, Hanson T, MacKinnon S, Rounsaville B, Sierra S. Contingency management treatments: reinforcing abstinence versus adherence with goal-related activities. *J Consult Clin Psychol.* 2006; 74:592–601. [PubMed: 16822115]
- Petry NM, Martin B, Cooney JL, Kranzler HR. Give them prizes and they will come: contingency management for treatment of alcohol dependence. *J Consult Clin Psychol.* 2000; 68:250–257. [PubMed: 10780125]
- Pittman B, Gueorguieva R, Krupitsky E, Rudenko AA, Flannery BA, Krystal JH. Multidimensionality of the Alcohol Withdrawal Symptom Checklist: a factor analysis of the Alcohol Withdrawal Symptom Checklist and CIWA-Ar. *Alcohol Clin Exp Res.* 2007; 31:612–618. [PubMed: 17374040]
- Prendergast M, Podus D, Finney J, Greenwell L, Roll J. Contingency management for treatment of substance use disorders: a meta-analysis. *Addiction.* 2006; 101:1546–1560. [PubMed: 17034434]
- Roll JM, Higgins ST, Badger GJ. An experimental comparison of three different schedules of reinforcement of drug abstinence using cigarette smoking as an exemplar. *J Appl Behav Anal.* 1996; 29:495–504. [PubMed: 8995832]
- Sakai JT, Mikulich-Gilbertson SK, Long RJ, Crowley TJ. Validity of transdermal alcohol monitoring: fixed and self-regulated dosing. *Alcohol Clin Exp Res.* 2006; 30:26–33. [PubMed: 16433729]
- Silverman K, Chutuape MA, Bigelow GE, Stitzer ML. Voucher based reinforcement of cocaine abstinence in treatment-resistant methadone patients: effects of reinforcer magnitude. *Psychopharmacologia.* 1999; 146:128–138.



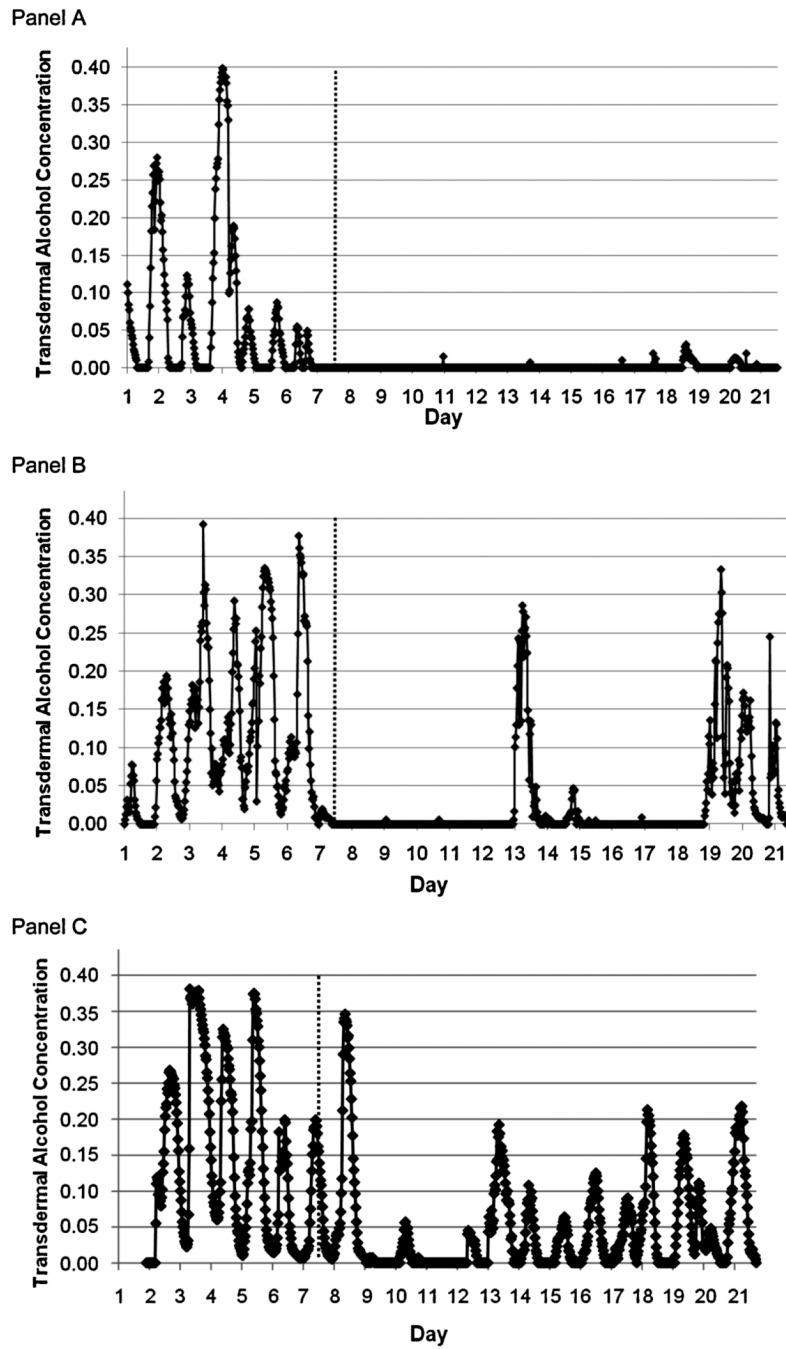
- Sobell, LC.; Sobell, MB. Timeline Followback: A technique for assessing self-reported ethanol consumption. In: Allen, J.; Litten, RZ., editors. *Measuring alcohol consumption: Psychosocial and Biological Methods*. Humana Press; Totowa, N.J.: 1992. p. 41-72.
- Sobell, LC.; Sobell, MB. *Alcohol Timeline Followback users' manual*. Addiction Research Foundation; Toronto, Canada: 1995.
- Sullivan JT, Sykora K, Schneiderman J, Naranjo CA, Sellers EM. Assessment of alcohol withdrawal: the revised clinical institute withdrawal assessment for alcohol scale (CIWA-Ar). *Br J Addict Alcohol Other Drugs*. 1989; 84:1353–1357.
- Swift RM. Transdermal measurement of alcohol consumption. *Addiction*. 1993; 88:1037–1039. [PubMed: 8401157]
- Swift RM. Direct measurement of alcohol and its metabolites. *Addiction*. 2003; 98 2:73–80. [PubMed: 14984244]
- Swift RM, Martin CS, Swette L, LaConti A, Kackley N. Studies on a wearable, electronic, transdermal alcohol sensor. *Alcohol Clin Exp Res*. 1992; 16:721–725. [PubMed: 1530135]
- U.S. Department of Health and Human Services and U.S. Department of Agriculture. *Dietary Guidelines for Americans*. 6. Washington, DC: US Government Printing Office; 2005.
- Williamson, D. [accessed on November 24, 2009] The area under a curve. 2001. <http://business.fortunecity.com/discount/29/areacurvweb.html>



**Figure 1. Percent of sample for which no alcohol was reported or detected by day**  
*Note.* Day 1 is Monday. Vertical lines denote weekends.



**Figure 2.** Peak and average transdermal alcohol concentration by day  
*Note.* Day 1 is Monday. Vertical lines denote weekends.



**Figure 3. Three individual participant readings by day**  
*Note.* Day 1 one is Monday. Days 1–7 were the baseline period.

**Table 1**  
**Participant report of drinking and ratings of transdermal readings**

	Participant report		Participant report	
	No	Yes	No	Yes
<b>Research Rater Consensus</b>	No	129   13   142	No	135   31   166
	Yes	4   124   128	Yes	0   104   104
		133   137   270		135   135   270
		Sensitivity: .91		Sensitivity: .76
		Specificity: .97		Specificity: 1.0
		False-positive: 3.0%		False-positive: 0
		False-negative: 9.5%		False-negative: 24.1%

*Note.* AMS = Alcohol Monitoring Systems. Days on which alcohol use was detected due to transdermal levels crossing over midnight ( $n = 25$ ) were recoded as self-report "yes" for the purposes of establishing agreement. The days that we had confirmed from a participant that he had lied about drinking ( $n = 6$ ) were moved to the "participant report yes" cell as well for this comparison.