

A preliminary randomized controlled trial of contingency management for alcohol use reduction using a transdermal alcohol sensor

Nancy P. Barnett¹, Mark A. Celio¹, Jennifer W. Tidey¹, James G. Murphy², Suzanne M. Colby¹ & Robert M. Swift^{1,3}

Center for Alcohol and Addiction Studies, Brown University, Providence, RI, USA,¹ Department of Psychology, University of Memphis, Memphis, TN, USA² and VA Medical Center, Providence, RI, USA³

ABSTRACT

Aims We tested the efficacy of daily contingent reinforcement for reducing alcohol use compared with (yoked) non-contingent reinforcement (NR) using a transdermal alcohol sensor to detect alcohol use. **Design** Pilot randomized controlled design with 1 baseline week, 3 intervention weeks and 1-month follow-up. **Setting** New England, USA. **Participants** Heavy drinking adults (46.7% female) not seeking treatment were randomized to (1) an escalating schedule of cash reinforcement (CR; $n = 15$) for days on which alcohol was neither reported nor detected or (2) yoked NR ($n = 15$). **Intervention and comparator** Reinforcement for CR participants started at \$5 and increased \$2 every subsequent day on which alcohol was not detected or reported, to a maximum of \$17. Participants received no reinforcement for days on which alcohol use was detected or reported, and the reinforcer value was re-set to \$5 the day after a drinking day. NR participants were yoked to the daily reinforcer value of an individual in the CR condition, in order of enrollment. Paired participants in CR and NR therefore received the same amount of money, but the amount for the NR participant was not behavior-related. **Measurements** The primary outcome was percentage of days without sensor-detected drinking. Secondary outcomes were number of consecutive days with no detected drinking, peak transdermal alcohol concentration (TAC), self-reported drinks per week and drinking below NIH low-risk guidelines. **Findings** Controlling for baseline, CR had a higher percentage of days with no drinking detected (54.3%) than NR (31.2%) during intervention weeks [$P = 0.05$, Cohen's $d = 0.74$; 95% confidence interval (CI) = 0.007–1.47]. The longest period of consecutive days with no drinking detected was 8.0 for CR versus 2.9 for NR ($P = 0.03$, $d = 0.85$; 95% CI = 0.08–1.61). Peak TAC during intervention showed a non-significant group difference ($P = 0.20$; $d = 0.48$; 95% CI = 0.00–1.18); a similar result was found for drinks per week ($P = 0.12$; $d = 0.59$; 95% CI = 0.00–1.30). Four times more participants in CR drank below NIH low-risk drinking guidelines during intervention than did participants in NR: 31.1 versus 7.1% ($P = 0.07$; $d = 0.71$; 95% CI = –0.04 to 1.46). At 1-month follow-up, the highest number of consecutive days without drinking (self-report) did not differ significantly between conditions ($P = 0.26$), but showed a medium effect size ($d = 0.44$; 95% CI = –0.32 to 1.18). **Conclusions** Cash incentives linked to a transdermal alcohol sensor can reduce heavy alcohol consumption while the incentives are in operation.

Keywords Alcohol, biosensor, contingency management, monitoring, reinforcement, transdermal.

Correspondence to: Nancy P. Barnett, Center for Alcohol and Addiction Studies, Brown University, Box G-S121-5, Providence, RI 02912, USA. E-mail: nancy_barnett@brown.edu

Submitted 8 March 2016; initial review completed 22 June 2016; final version accepted 13 January 2017

INTRODUCTION

Contingency management (CM) is among the more effective strategies for promoting abstinence in the treatment of substance use disorders [1–3]. CM protocols are designed to reinforce target behaviors (e.g. a negative urine drug

screen) that are aligned with treatment goals (e.g. maintained abstinence) by delivering a tangible reward when the target behavior occurs and withholding the reward when the target behavior does not occur [4,5]. Target behaviors must be defined clearly and measured objectively to ensure that contingent reinforcement is delivered only

when that specific behavior has occurred. Furthermore, it is critical that the target behavior is measured at a frequency that maximizes the likelihood of detecting violations of treatment goals should they occur [6,7].

Historically, the application of CM strategies to the treatment of alcohol use disorders has been limited by reliance on breath alcohol tests—typically administered in out-patient treatment settings—as the primary objective measure of alcohol use [8–10]. Given the short half-life of breath alcohol, negative breath tests ensure that individuals have not consumed alcohol within a relatively brief period (i.e. up to 12 hours [11]), but will not detect other drinking between observations. Randomly prompted daily breath testing has been employed to address this limitation [12]; however, comparison to self-report data indicated some occasions of drinking were still not captured. Ethyl glucuronide (EtG) urine testing, which allows for detection of alcohol in urine for up to 3 days [13], has also been employed within a recent CM trial [14] but the observed sensitivity of EtG (at a 500 ng/ml cut-off) detected self-reported alcohol use 50% of the time participants reported drinking 2 days prior, suggesting the need for different detection thresholds and/or monitoring frequency. One approach that can address the limitations of other objective measures of recent alcohol use is transdermal alcohol monitoring, which provides a continuous measure of alcohol excreted through the skin [15–17].

Three studies have examined the efficacy of CM interventions using transdermal alcohol concentration (TAC) as the primary objective measure of alcohol use. In a sample of heavy drinking adults ($n = 13$), Barnett and colleagues [18] compared drinking in 1 baseline week to a 2-week period in which monetary reinforcement was provided for TAC-verified non-drinking days. Even in this small sample, statistically significant within-subject intervention effects were found on self-reported and transdermal measure outcomes. Dougherty and colleagues [19] counterbalanced CM and non-contingent monitoring to examine the efficacy of a CM protocol in which monetary reinforcement was provided when heavy drinking was not detected ($n = 26$). This study demonstrated within-subjects reductions in self-report and transdermal measures of alcohol consumption during the CM phase, and showed that individuals who experienced the CM phase first showed less drinking during the non-contingent monitoring phase than individuals who received the CM phase second, suggesting some maintenance of CM-induced behavior change. Two reports from a more recent study conducted by Dougherty and colleagues showed similar within-subject reductions in number of heavy drinking days in a sample of at-risk drinkers ($n = 82$) verified by self-report [20] and by TAC data [21] during a 12-week CM protocol. Despite the promising evaluation of CM protocols in which TAC is the primary objective measure

of the target behavior for contingent reinforcement, to date no investigations using a between-groups design have been conducted.

The purpose of this study was to evaluate the preliminary efficacy of a CM intervention for reducing alcohol use, using alcohol sensor data from the Secure Continuous Remote Alcohol Monitor (SCRAM) bracelet as the basis of the contingent reinforcement and for evaluation of the intervention outcome. We hypothesized that participants receiving contingent reinforcement would show (1) a higher percentage of days with no drinking detected, (2) longer number of consecutive days with no drinking detected, (3) lower peak TAC during intervention weeks and (4) lower average number of drinks per week compared to non-contingent reinforcement during the intervention weeks and that group differences in outcomes would be maintained at 1-month follow-up.

METHODS

Study overview and design

The investigation was conducted in southern New England, USA in a university research center. In a two-group design, following 1 baseline week, participants were assigned randomly to 3 weeks of contingent reinforcement ($n = 15$) or yoked non-contingent reinforcement ($n = 15$). Participants were again assessed 1 month after the intervention using self-reported outcomes. Given the preliminary nature of this trial, formal power analysis was not conducted a priori. The final sample of 30 provided 53% power to detect a medium size effect using regression analyses. All participants gave informed consent and the protocol was approved by the Brown University Institutional Review Board.

Study sample

Participants were recruited through flyers and online advertisements; these stated that the research was seeking adult drinkers who were interested in reducing or stopping drinking. Eligibility criteria were: (1) age ≥ 18 years; (2) past-month alcohol use above the national guidelines (> 7 standard drinks for women, > 14 for men; [22]); (3) two or more heavy drinking episodes per week in the past month; (4) interest in cutting down or stopping drinking; (5) daily internet access; and (6) an acceptable landline telephone for SCRAM modem data transmission or willingness to come to the research office for data downloads. Exclusion criteria were: (1) drug use other than marijuana more than once a month or in the past month; and (2) significant history of withdrawal as indicated by an Alcohol Withdrawal Symptom Checklist score ≥ 23 [23] or a Clinical Institute Withdrawal Assessment Revised (CIWA-Ar) score ≥ 10 [24].

Study procedures and measures

Baseline session

Prospective participants responding to online advertisements were screened for initial eligibility and scheduled for a baseline session at which a breathalyzer test verified breath alcohol concentration (BAC) = 0.000; a urine toxicology screen verified no recent drug use other than marijuana and administration of the CIWA-Ar verified self-report of minimum withdrawal symptoms. Other measures administered at baseline included demographics, the Structured Clinical Interview for DSM-IV (SCID), Substance Abuse Module [25] and a shortened version of the Treatment Services Review [26], which collected self-reported information about prior alcohol treatment. Following the baseline survey, the SCRAM bracelet was installed. Participation started on a Monday¹ and the first week was a baseline week; participants were told that they were not expected to change their drinking during this week. Participants received \$25 for the baseline session.

Alcohol monitoring bracelet and coding for detected alcohol

The SCRAM bracelet (versions SCRAM II and SCRAMx; Alcohol Monitoring Systems; AMS) is strapped to the participant's ankle and worn continuously. The bracelet collects a sample of the vapor close to the skin every 30 minutes and stores the reading. The bracelet cannot be removed by the wearer without cutting the strap or breaking the closure clip. Research staff installed the bracelet and provided instructions to participants. Bracelet data were transmitted overnight via a SCRAM modem in the participant's home or were downloaded three times per week in the research office. Once transmitted, these data were available immediately on a secure AMS website. AMS provides data review and indicates bracelet problems and whether bracelet tampering is detected. TAC data were coded by research staff within a day of being available using criteria developed in prior work [18,27]. Alcohol use was coded as detected on days that: (1) an episode had one or more TAC readings more than 0.02 g/dl and (2) either the absorption rate for the episode < 0.05 g/dl per hour or the elimination rate for the episode < 0.025 g/dl per hour (when peak < 0.15 g/dl) and < 0.035 g/dl per hour (when peak > 0.15 g/dl).

Self-reported alcohol use

Every morning during the baseline and intervention weeks, participants received an e-mail containing a link to a web-based survey that assessed drinking on the previous day. Participants indicated the number of standard drinks (12 oz. beer, 5 oz. wine, 1.5 oz. liquor) they had consumed, from which the average number of drinks per week

was calculated. Participants received \$5 for each completed report and a \$25 bonus for submitting 90% (25 of 28) of reports on the day they were requested. A 30-day time-line follow-back (TLFB) [28,29] was administered at baseline and follow-up; from this we derived the percentage of days no drinking reported, the longest number of days with no drinking reported and average number of drinks per week. We also calculated an estimate of blood alcohol concentration (eBAC) [30] for each day, and the peak value across all days reflected highest level of intoxication in the past month.

Intervention conditions

On the first day of week 2 (always a Monday), participants received their condition assignment and instructions. The first five participants were assigned to contingent reinforcement (CR), so we would have cases for yoking to non-contingent reinforcement (NR) participants. Thereafter, participants were assigned randomly to condition using a random numbers table determined by the project coordinator and conveyed to research assistants on the day of assignment. During intervention weeks, participants in the CR group received monetary reinforcement for days when no drinking was reported or detected using TAC. When self-report was missing, only TAC was used. Participants were told that any level of drinking might be detected. Each Monday, reinforcement started at \$5 and increased \$2 every subsequent day on which alcohol was not detected or reported, to a maximum of \$17. We used this schedule because there is evidence that increasing the magnitude of reinforcement using an escalating payment schedule produces superior results [31,32]. Participants received no reinforcement for days on which alcohol use was detected or reported, and the reinforcer value was re-set to \$5 the day after a drinking day.

Participants in the NR condition were yoked to the daily reinforcer value of an individual in the CR condition, in order of enrollment. Paired participants in CR and NR therefore received the same amount of money; the only difference between conditions was whether the money received was contingent upon alcohol use reduction. NR participants were informed that they would receive a monetary value of \$0–17 each day, but that the value they received was not related to their drinking or to any other behaviors. Following every daily survey, all participants received a webpage report indicating whether alcohol had been detected on each previous day, the reinforcer value received for each day and a summary of all reinforcement to date. This information could also be accessed at any time on a password-protected website. At any time, participants in either group could receive cash for the value

¹Nine (30.0%) participants started on a Tuesday; all data for these participants were adjusted to reflect a 6-day baseline week.

of the reinforcers due to them. The maximum reinforcement value for both groups was \$231.

Post-intervention session and follow-up

At the completion of the intervention period, participants attended a post-intervention session; at this session the bracelet was removed and participants completed a survey about their experience wearing the bracelet. Questions measured physical and social comfort ('How physically uncomfortable was the bracelet?', 'How embarrassing [socially uncomfortable] was it to wear the bracelet?') on a 10-point scale from 1 'very comfortable' to 10 'extremely uncomfortable'. Interference with work, exercise, sleep, ability to concentrate and choice of clothing were measured on a 10-point scale from 1 'not at all' to 10 'completely'. Side effects of itching, sweating and skin irritation were measured on a scale from 1 'not noticeable' to 10 'unbearable'. We also asked: 'Did you have any marks on your skin from the bracelet?' and 'Would you wear the bracelet for another week?'. Participants received \$25 for this interview and received all payments due to them (most waited to receive their payments at this last session). One month after the intervention, participants completed an in-person follow-up interview for \$30.

Outcomes and data analysis

The primary outcome during intervention weeks was the percentage of days with no drinking detected using TAC.² At follow-up, we used the 30-day TLFB to produce an analogous report of self-reported percentage of days with no drinking reported. Secondary outcomes were longest number of consecutive days with no drinking detected using TAC, peak TAC across all intervention weeks, and self-reported outcomes of average number of drinks per week and the proportion of participants that stayed within two US guidelines for drinking during the intervention weeks: the dietary guideline of no more than one drink per day for women and two drinks per day for men [22] and the National Institute on Alcohol Abuse and Alcoholism (NIAAA) criteria for low-risk drinking of no more than three drinks on any single day and no more than seven drinks per week for women and four per day/14 per week for men [33].

For between-group comparisons, independent-samples *t*-tests and χ^2 tests were used. For outcome analyses conducted on the 3 intervention weeks, multiple regression was used with the baseline observation week covaried. For follow-up analyses, we used regression with the baseline session values covaried. Cohen's *d* was calculated from regression output ($2 t/\sqrt{df}$). We also conducted a survival analysis using Cox regression in which the number of days to first drink during the intervention weeks was the measured effect.

Missing data

Some bracelet data were missing for seven participants (23.3%), representing 5.5% of all monitored days. Missing TAC was due to: (1) no data collected by the bracelet (i.e. bracelet malfunction; $n = 3$); (2) temporary bracelet removal ($n = 3$; two participants were detained by security/police and one who said the bracelet 'flew off' while she was dancing)—these were detected by AMS as tampers; and (3) undetermined ($n = 1$; possible malfunction, possible environmental alcohol). For the latter two conditions, data were removed by research staff from the data set. Summary variables were calculated using available data.

RESULTS

Of the 69 participants enrolled at baseline based on their eligibility on the telephone screen, 16 were determined ineligible at the in-person baseline appointment because they did not meet both drinking inclusion criteria on the TLFB. In the baseline week, 19 were withdrawn because they did not meet inclusion criteria by self-report and TAC (2 or more drinking days). An additional three participants withdrew during the baseline week due to skin irritation, work clothing restrictions and travel. Figure 1 shows eligibility, enrollment and retention through all study phases. Information about participants is shown in Table 1. There were no significant differences between intervention groups on demographics, background information, self-reported drinking at baseline, use of a modem for data transmission and self-reported drinking and transdermal variables during the baseline observation week. There were no baseline differences

²The nature of alcohol metabolism and of the BAC/TAC curve is such that alcohol can be detected long after consumption. Thus, TAC derived from alcohol consumed on one day will commonly 'cross over' into the second day (i.e. cross over midnight). This creates a dilemma when determining whether to consider the second day as a day on which alcohol was consumed. On one hand, alcohol may have been consumed after midnight and therefore impairment on the second day was possible. On the other hand, we commonly record a drinking episode as occurring on the day it began (e.g. in the time-line follow-back measure), and disregard whether drinking crosses over midnight or whether intoxication continues past midnight. Furthermore, counting the second day as a drinking day may exaggerate intervention group differences, as any day on which TAC crossed over into a second day would be counted as 2 days of drinking when it may only reflect one drinking episode. However, it is not unusual to see a single TAC episode cross over multiple days without returning to 0, reflecting continued drinking on subsequent days, so counting the day on which the episode began as the only day of drinking would, in some cases, also be inaccurate. In summary, a day when TAC 'cross-over' is detected may or may not indicate drinking on that day. For this reason, we have analyzed our data using both ways: first, counting the second day on which a transdermally detected episode continued as a day of drinking and secondly, counting only the first day on which a transdermally detected episode occurred as a drinking day.

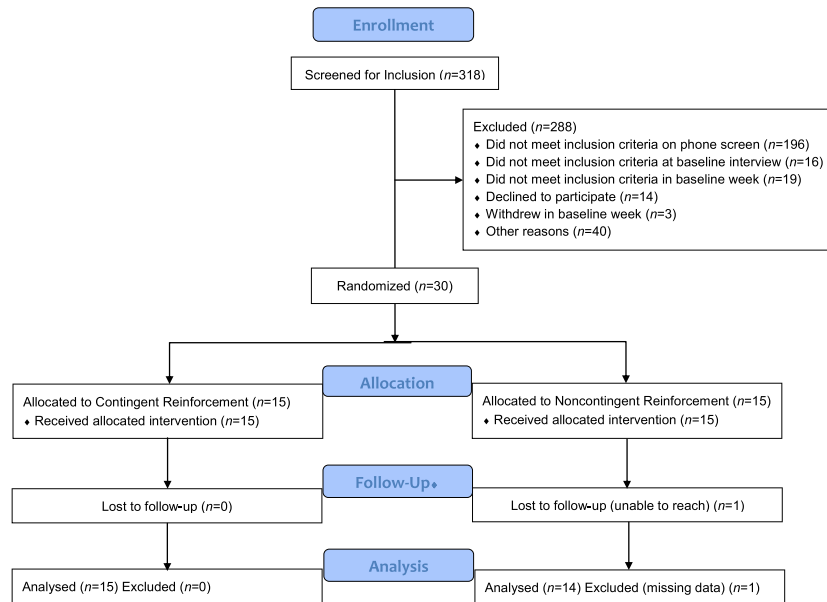


Figure 1 Flow of participants through trial [Colour figure can be viewed at wileyonlinelibrary.com]

between participants who used a modem (30%) and those who did not.

Correspondence between self-report and TAC data

The daily web-based survey completion rate was 99.4%. Inter-rater reliability of coding the TAC episode criteria was excellent [correlation coefficient (ICC) = 0.98]. The correlation between self-reported number of drinks and

area under the transdermal curve was 0.79, $P < 0.001$, indicating a good correspondence between TAC and self-report measures of volume, and consistent with other work [17,18].

Outcomes during intervention weeks

Means and standard deviations (SD) for outcome variables during the baseline week and intervention weeks are in

Table 1 Participant demographic and background information.

Variable	Whole sample (n = 30)	Contingent (n = 15)	Non-contingent (n = 15)	P-value
Female	14 (46.7%)	7 (46.7%)	7 (46.7%)	1.0
Age (years, range = 21–57)	28.9 (9.3)	28.3 (7.9)	29.5 (10.7)	0.72
Race				
White	23 (76.7%)	12 (80.0%)	11 (73.3%)	0.28
Black	3 (10.0%)	0	3 (20.0%)	
Asian	1 (3.3%)	1 (6.7%)	0	
Multi-racial	1 (3.3%)	1 (6.7%)	0	
Did not answer	2 (6.7%)	1 (6.7%)	1 (6.7%)	
Hispanic/Latino/a	4 (13.3%)	1 (6.7%)	2 (20.0%)	0.28
Weight	179.2 (42.9)	175.7 (41.6)	182.8 (45.3)	0.66
High school graduate/GED	27 (90.0%)	14 (93.3%)	13 (86.7%)	0.54
Baseline alcohol use				
Percentage of drinking days past month	67.8%	66.2%	69.3%	0.69
Drinks per week past month	35.9 (14.5)	34.1 (13.6)	37.7 (15.5)	0.51
Alcohol abuse or dependence diagnosis ^a	18 (62.1%)	8 (57.1%)	10 (66.7%)	0.60
Received treatment for alcohol in past year	2 (6.7%)	0	2 (13.3%)	0.14
Used SCRAM modem for data transmission (versus in-person downloads)	9 (30.0%)	4 (26.7%)	5 (33.3%)	0.69

Mean (standard deviation) presented unless indicated otherwise. ^aSCID diagnosis was missing for one participant in the contingent group. GED = general educational development; SCID = Structured Clinical Interview for DSM-IV; SCRAM = Secure Continuous Remote Alcohol Monitor.

Table 2 Intervention group differences on alcohol outcomes.

Outcome measure	Contingent (<i>n</i> = 15) mean (SD)	Non-contingent (<i>n</i> = 15) mean (SD)	<i>P</i> -value	Cohen's <i>d</i> (95% CI)
Primary outcome				
Percentage of days no drinking detected (counting 'cross-over' days as drinking days) ^a				
Baseline week	19.1 (17.8)	23.0 (17.7)	0.053	0.74 (−0.007 to 1.47)
Intervention weeks	54.3 (36.7)	31.2 (28.9)		
(not counting 'cross-over' days as drinking days) ^a				
Baseline week	39.0 (18.4)	37.3 (19.8)	0.083	0.67 (−0.07 to 1.40)
Intervention weeks	65.7 (30.7)	47.7 (27.2)		
Follow-up (self-report only)	55.6 (25.2)	56.2 (29.1)	0.902	0.05 (0.00 to 0.31)
Secondary outcomes				
Longest number of consecutive days with no drinking detected				
Baseline week	1.1 (1.2)	1.3 (1.0)	0.031	0.85 (0.08 to 1.61)
Intervention weeks	8.0 (7.5)	2.9 (2.2)		
Follow-up (self-report only)	7.1 (5.9)	5.7 (7.6)	0.260	0.44 (−0.32 to 1.18)
Peak TAC				
Baseline week	0.191 (.111)	0.198 (.073)	0.200	0.48 (0.00 to 1.18)
Intervention weeks	0.189 (.123)	0.241 (.094)		
Follow up peak eBAC (self-report)	0.145 (.082)	0.184 (.101)	0.625	0.20 (0.00–0.94)
Average number of drinks per week (self-report)				
Baseline week	37.4 (20.3)	40.8 (21.6)	0.118	0.59 (0.00 to 1.30)
Intervention weeks	18.5 (18.9)	30.2 (18.7)		
Follow-up	19.0 (12.2)	21.2 (18.9)	0.895	0.05 (−0.69 to 0.79)

TAC = transdermal alcohol concentration; eBAC = estimated breath alcohol concentration. Peak TAC and peak eBAC reflect the highest values across all collected days. Longest number of consecutive days with no drinking detected at follow-up was square root-transformed for analyses; untransformed values are presented. Follow-up values reflect past 30 days CI = confidence interval; SD = standard deviation. Regression analyses for intervention weeks controlled for baseline week values. Analyses for follow-up (all self-report) controlled for baseline values from the time-line follow-back. ^a'Cross-over' days were days when a TAC-identified drinking episode started on one calendar day and crossed over at midnight into the next day. The two analyses in this table treated the second day of drinking differently, either counting it as a drinking day or not.

Table 2. Participants received an average of \$95 in reinforcement (same for both groups because of yoking; 41% of the possible amount; range 0–100%).

Primary outcome

Percentage of days no alcohol detected (TAC)

Figure 2 shows the percentage of days with no alcohol detected by week and group. With baseline week values covaried, the significance of intervention condition was $P < 0.10$, with the CR group showing fewer days of drinking than the NR group, both when 'cross-over' TAC was counted as a drinking day and when it was not.³

Secondary outcomes

Longest number of consecutive days with no drinking (TAC)

During intervention weeks, the longest period of no drinking was significantly longer in the CR group compared to the NR group after controlling for the baseline week ($\beta = 0.41$, $P = 0.031$). In addition, a significantly higher proportion of CR participants had 1 week or more with no drinking detected (40.0 versus 7.1%, $\chi^2_{(df = 1, n = 29)} = 4.3$, $P = 0.039$).

The survival plot for number of days to first detected drinking is in Fig. 3. The contingencies had a hazard ratio of 2.14 [95% confidence interval (CI) = 0.92–4.99], $P = 0.079$, indicating that the failure rate for the NR

³When self-reported and detected drinking days were both included as indicating no drinking, effects became statistically significant, both when 'cross-over' TAC was counted as a drinking day ($\beta = 0.41$, $P = 0.029$) and when it was not ($\beta = 0.37$, $P = 0.047$). It is possible that self-report captured days of drinking that TAC did not ([27]), but the two groups also had different incentives for reporting no drinking, so we cannot determine if the findings that include self-report are due to the CR group reporting less drinking or actually drinking less.

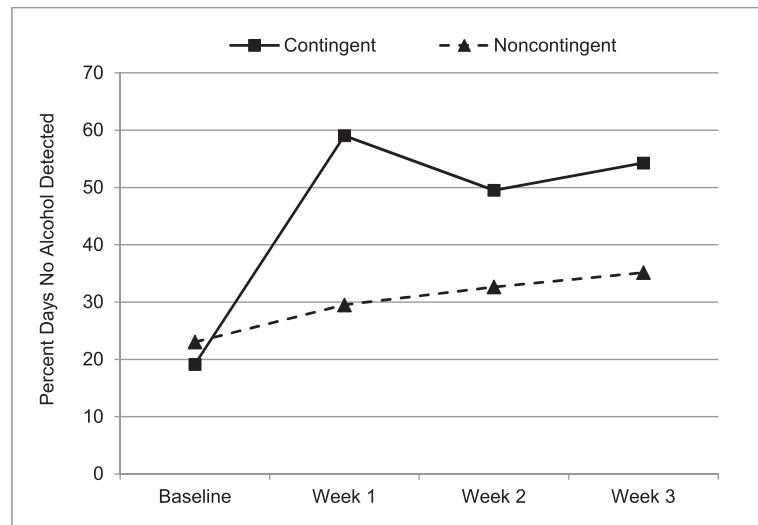


Figure 2 Percentage of days with no alcohol detected by intervention group and week. Weeks 1–3 are intervention weeks

group was more than two times higher than the CR group.

Peak TAC overall

The CR group showed essentially no change in peak TAC from baseline to intervention weeks, whereas the NR group showed a 22% increase in peak TAC from the baseline week to the intervention weeks (see Table 2). The group difference was not significant, but showed a medium effect size.

Average number of drinks per week (self-report)

The average number of drinks per week declined from the baseline week to the (average of the) 3 intervention weeks for both groups: by 26% in the NR group and by 51% in the CR group. Within both groups this decline was

significant using paired *t*-tests: $t(14) = 2.65$, $P = 0.019$ for CR, and $t(14) = 2.24$, $P = 0.042$ for NR. Controlling for the baseline week, the between-groups difference was nonsignificant for the intervention weeks, with a medium effect size.

Low-risk drinking (self-report)

On average, the CR group stayed within the dietary guideline ($\leq$ one drink per day for women; $\leq$ two drinks per day for men) on 61.0% of days and the NR group stayed within the guideline on 44.1% of days ($\beta = 0.25$, $P = 0.18$), with a medium effect size of $d = 0.51$ (95% CI = -0.23 to 1.25). The percentage of weeks in which participants stayed under the National Institutes of Health (NIH) daily ($\leq$ three for women; $\leq$ four for men) and weekly ($\leq$ seven for women; $\leq$ 14 for men) limits were 31.1% in the CR group

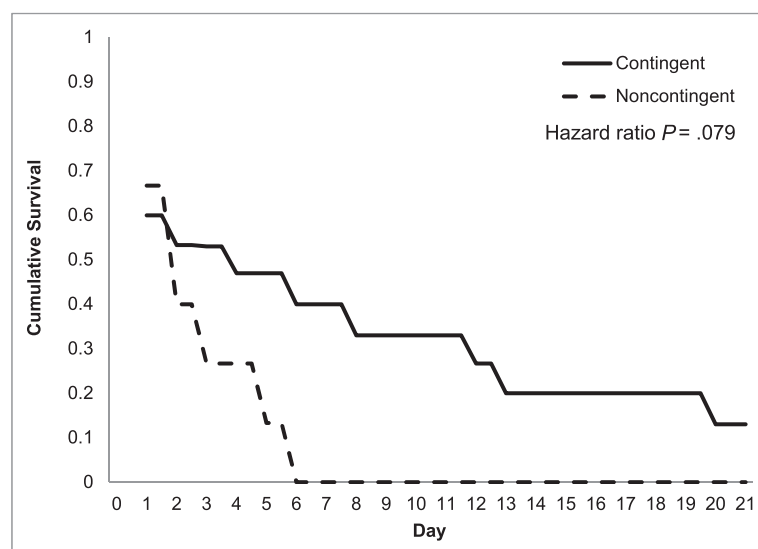


Figure 3 Survival curves for days to first detected drinking in intervention groups during intervention weeks

and 7.1% in the NR group ($\beta = 0.35$, $P = 0.07$), with a medium-to-large effect size of $d = 0.71$ (95% CI = -0.04 to 1.46) Three participants met the NIH criteria for all 3 weeks; all were in the CR group.

Post-intervention

One person in the NR group did not complete the post-intervention survey. Participant responses to wearing the SCRAM bracelet are in Table 3. Physical discomfort had the most negative rating in the sample as a whole. The most problematic side effect was itching, but the average score for the side effects items was 2.9 (SD = 1.6) on a scale of 1–10, suggesting that overall side effects were minimal. There were no differences between groups on any of these items.

Follow-up outcomes

The completion rate for the 1-month follow-up assessment was 93%. Group differences were in the hypothesized direction, but no outcomes were significantly different. On drinks per week the CR group showed very little within-group change from post-intervention to follow-up, $t(14) = 0.08$, $P = 0.94$; $d = 0.02$ (95% CI = -0.49 to 0.53), but the NR group showed improvement, $t(12) = -2.53$, $P = 0.027$; $d = 0.70$ (95% CI = 0.07 – 1.29), indicating that the reduction in the between-groups differences between intervention and follow-up was due to the NR group reducing drinking at follow-up, not to an increase in drinking in the CR group.

DISCUSSION

This study is the first between-groups controlled trial of contingent reinforcement for alcohol use reduction using transdermal alcohol detection as a continuous objective indicator of alcohol use. Participants who were reinforced on a daily basis had approximately one-and-a-half as many days with no detected alcohol use, and sustained not drinking for almost three times longer than the comparison group. Data from the transdermal monitor showed that the highest TAC level during intervention weeks among those in the CR group was 22% lower than among those in the NR group. Furthermore, when we considered the clinical impact of the intervention, we found that participants in CR were more than three times more likely than those in NR to stay within NIH drinking guidelines. Although most findings were not statistically significant, during the intervention weeks all outcomes had medium-to-large effect sizes in the hypothesized direction.

There were no significant differences in outcomes between intervention groups 1 month after the intervention, but two of the four outcomes were in the expected direction, with small or small-to-medium effect sizes, including an estimate of BAC. We note that across most outcomes the control group showed improvement between the intervention period and follow-up, with minimal change in the intervention group. This pattern of findings suggests that the effect of the contingent reinforcement did not decay in the CR group, but the mechanism through which the NR group improved from baseline throughout the intervention weeks and the follow-up observation is unknown.

Table 3 Participant response to wearing alcohol monitoring bracelet.

Variable	Whole sample ($n = 29$)	Contingent ($n = 15$)	Non-contingent ($n = 14$)	P^a
Physical discomfort ^b	5.86 (1.89)	5.53 (1.96)	6.21 (1.81)	0.34
Social discomfort ^b	4.59 (2.57)	4.40 (1.92)	4.79 (3.19)	0.69
Interference with daily activities ^c				
Choice of clothing	5.28 (3.49)	5.73 (3.39)	4.79 (3.66)	0.48
Exercise	3.24 (2.57)	3.13 (2.75)	3.36 (2.47)	0.82
Sleep	2.97 (1.99)	2.93 (2.15)	3.00 (1.88)	0.93
Work	1.82 (1.49)	1.71 (1.14)	1.93 (1.82)	0.71
Ability to concentrate	1.72 (1.28)	1.80 (1.42)	1.64 (1.15)	0.75
Side effects ^d				
Itching	3.36 (2.31)	3.50 (2.59)	3.21 (2.08)	0.75
Sweating	2.89 (2.08)	3.07 (1.77)	2.71 (2.40)	0.66
Skin irritation	2.28 (1.83)	2.60 (2.17)	1.93 (1.39)	0.80
Marks on skin (%)	52%	47%	57%	0.57 ^e
Willing to wear the bracelet for another week (%)	79%	87%	69%	0.57 ^e

Mean (standard deviation) presented unless otherwise indicated. One participant did not complete this assessment. ^aThe P -values presented were derived from independent-sample t -tests, unless indicated otherwise. ^bPhysical discomfort and social discomfort were measured using a 10-point scale (10 = 'extremely uncomfortable'). ^cThe degree of interference with daily activities was measured using a 10-point scale (10 = 'completely' interfered). ^dSide effects were measured using a 10-point scale (10 = 'unbearable'). ^e χ^2 test.

Our findings are consistent with previous studies using single-group designs in which participants showed change over time in response to contingent reinforcement for reducing alcohol use on both transdermal and self-reported outcomes [18,20,21], and maintained that change in the short term [20,21]. Two other trials of CM used breath alcohol [12] and EtG [14] for objective verification; both these studies showed promise for the use of objective measures in a CM study but had no follow-up beyond the intervention weeks. It is important to note that there are important trade-offs for the different alcohol biomarkers used in CM interventions. The SCRAM collects information without required action from the user and provides multiple observations per day, whereas breath tests require many prompted self-tests or staff-observed observations which interfere with daily life, and EtG has a detection window of only 24–48 hours [34,35]. Being able to monitor alcohol use remotely without any contact with the participant has the potential to be a cost-effective approach for addressing problematic alcohol use. Furthermore, participants were generally accepting of the SCRAM technology, the physical side effects were minimal and the social discomfort was tolerable.

We used cash incentives because studies have suggested that participants value cash more than they value vouchers that can be exchanged for retail goods [36,37]; there is also evidence that cash reinforcers and vouchers do not differ in efficacy, and that cash payment does not increase craving or high-risk behaviors [38]. However, there is considerable research showing that CM using prize-based vouchers for patients in treatment is efficacious [39–41] and more cost-effective than cash [42].

Strengths and limitations

This study contained several methodological strengths, including a baseline week that provided an objective indicator of comparability between groups prior to randomization, an escalating reinforcement schedule with yoking to control for the monetary value of the reinforcers received by the CR group, a daily web-based report and the ability for remote sensor data uploads, thus minimizing staff effort and participant burden.

A primary limitation is that we did not have an objective indicator of drinking during the follow-up phase, so comparing the intervention phase in which the objective verification was present to the baseline and follow-up phases in which there was no objective indication of drinking might not be appropriate. We used midnight as the end of the calendar day, which increased the likelihood of finding drinking episodes that crossed over from one day to the next, thereby complicating the identification of drinking days. We recommend that future researchers convert the measured 'day' to reflect awake hours on drinking days

(including early morning hours). Reinforcers were withheld based on self-report or detected levels of alcohol, and there were some drinking days that we would not have detected (using just TAC) if participants had not reported alcohol use (see [18] for overview); however, our outcome analyses included only days on which alcohol use was detected. In some cases reinforcement was delayed until in-person bracelet downloads were completed, and most participants did not 'cash out' until the study end; these delays in the immediacy of reinforcement could reduce efficacy. The enrolled sample may have had less severe substance use history than a typical treatment sample; approximately half were alcohol-dependent; most had not had previous treatment.

There are some limitations of the transdermal technology that are relevant. Transdermal alcohol is not equivalent to breath or blood alcohol levels, so it is impossible to determine level of impairment using the familiar metric of BrAC/BAC, although efforts are under way to address this limitation [43–47]. The TAC curve is also delayed relative to BAC, so it is difficult to use sensor data for real-time assessment of contextual factors' effect on drinking. The TAC level and TAC curve must meet specific criteria, so some drinking may not have been detected, but these would tend to be low-risk episodes [27]. One complication with using sensor data to establish adherence to treatment expectations is that the data have to be viewed and coded either by a coder or by software [48], a more time-consuming process than viewing toxicology test results. Finally, bracelet malfunctions do occur, but in this study malfunctions were lower than previous reports, some of which used earlier SCRAM models (see [27] for more detail).

Clinical implications and future directions

Research has demonstrated that contingent reinforcement for abstinence is effective in reducing substance use among those in behavioral or pharmacotherapy treatment [49], and that reinforcing attendance in alcohol treatment populations increases attendance effectively [9], but there are very few studies that have reinforced reduced alcohol consumption among patients in treatment [10]. The advent of alcohol sensors and novel alcohol biomarkers (including phosphatidylethanol) should allow for greater evaluation of CM for reduced drinking in patients with alcohol use disorders, but fully powered RCTs with alcohol treatment populations are needed.

Transdermal sensors provide a semi-quantitative drinking measure that can be used to support reductions in drinking and/or to shape behavior towards abstinence. Our peak TAC variable had a lower effect size than other variables, and this may be because when participants chose to drink they did so at their typical level. Therefore,

reinforcing lower TACs (e.g. an absolute value or a percentage of one's peak) might be a promising harm-reduction intervention. Additionally, observing TAC values during treatment could help to identify different response trajectories, which offers the potential for detecting different treatment needs early. Alcohol biosensors can provide a valuable indicator of treatment-induced behavior change and may be reasonable to use as an outcome measure [13,50].

CONCLUSION

Using continuous alcohol monitoring as an objective measure, clinically meaningful behavior change was detected among heavy drinking participants who received contingent reinforcement for showing no alcohol use. Effect sizes were most commonly in the medium-to-large range during intervention, suggesting that the potential for using sensor technology to support behavior change is high. RCTs among treatment populations with alcohol use disorders are needed.

Declaration of interests

R.M.S. has received research funding from Farmaceutico CT and travel and honoraria from Lundbeck. N.P.B., M.A.C., J.G.M., S.M.C. and J.W.T. have none to declare.

Acknowledgements

The authors wish to thank Polly Gobin, John Hustad, Michelle Loxley, Elizabeth Meade, Nadine Mastroleo and Lindsay Orchowski who assisted in this research, Tim Souza and Cheryl Eaton who prepared data, Paul Marques who provided initial consultation on study design, Thad Leffingwell who provided consultation for measures and data processing, Robert Stout for analytic advice, and Jeff Hawthorne at Alcohol Monitoring Systems. This investigation was supported in part by research grants R21 015980, R21 020943 and T32 AA07459 from the National Institute on Alcohol Abuse and Alcoholism. 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.

References

1. Dutra L., Stathopoulou G., Basden S. L., Leyro T. M., Powers M. B., Otto M. W. A meta-analytic review of psychosocial interventions for substance use disorders. *Am J Psychiatry* 2008; **165**: 179–87.
2. Lussier J. P., Heil S. H., Mongeon J. A., Badger G. J., Higgins S. T. A meta-analysis of voucher-based reinforcement therapy for substance use disorders. *Addiction* 2006; **101**: 192–203.
3. Prendergast M., Podus D., Finney J., Greenwell L., Roll J. Contingency management for treatment of substance use disorders: a meta-analysis. *Addiction* 2006; **101**: 1546–60.
4. Higgins S. T., Budney A. J., Bickel W. K., Foerg F. E., Donham R., Badger G. J. Incentives improve outcome in outpatient behavioral treatment of cocaine dependence. *Arch Gen Psychiatry* 1994; **51**: 576–86.
5. Budney A. J., Higgins S. T. *A Community Reinforcement and Vouchers Approach: Treating Cocaine Addiction*. National Institute on Drug Abuse therapy manuals for drug addiction: Manual 2 (NIH Publication no. 98–4309). Washington, DC: US Department of Health and Human Services; 1998.
6. Higgins S. T., Budney A. J., Bickel W. K. Applying behavioral concepts and principles to the treatment of cocaine dependence. *Drug Alcohol Depend* 1994; **34**: 87–97.
7. Petry N. M., Martin B., Finocche C. Contingency management in group treatment: a demonstration project in an HIV drop-in center. *J Subst Abuse Treat* 2001; **21**: 89–96.
8. Alessi S. M., Hanson T., Wieners M., Petry N. M. Low-cost contingency management in community clinics: delivering incentives partially in group therapy. *Exp Clin Psychopharmacol* 2007; **15**: 293–300.
9. Helmus T. C., Saules K. K., Schoener E. P., Roll J. M. Reinforcement of counseling attendance and alcohol abstinence in a community-based dual-diagnosis treatment program: a feasibility study. *Psychol Addict Behav* 2003; **17**: 249–51.
10. Petry N. M., Martin B., Cooney J. L., Kranzler H. R. Give them prizes and they will come: contingency management for treatment of alcohol dependence. *J Consult Clin Psychol* 2000; **68**: 250–7.
11. Warner E. A., Sharma N. Laboratory diagnosis. In: Ries R. K., Miller S. C., Fiellen D. A., Saitz R., editors. *Principles of Addiction Medicine*, 4th edn. Philadelphia: Lippincott Williams & Wilkins; 2009, pp. 295–304.
12. Alessi S. M., Petry N. M. A randomized study of cellphone technology to reinforce alcohol abstinence in the natural environment. *Addiction* 2013; **108**: 900–9.
13. Litten R. Z., Bradley A. M., Moss H. B. Alcohol biomarkers in applied settings: recent advances and future research opportunities. *Alcohol Clin Exp Res* 2010; **34**: 955–67.
14. McDonell M. G., Howell D. N., McPherson S., Cameron J. M., Srebnik D., Roll J. M. *et al.* Voucher-based reinforcement for alcohol abstinence using the ethyl-glucuronide alcohol biomarker. *J Appl Behav Anal* 2012; **45**: 161–5.
15. Swift R. M. Direct measurement of alcohol and its metabolites. *Addiction* 2003; **98**: 73–80.
16. Marques P. R., McKnight A. S. *Evaluating transdermal alcohol measuring devices (DOT HS 810 875)*. Washington, DC: National Highway Traffic Safety Administration; 2007.
17. Sakai J. T., Mikulich-Gilbertson S. K., Long R. J., Crowley T. J. Validity of transdermal alcohol monitoring: fixed and self-regulated dosing. *Alcohol Clin Exp Res* 2006; **30**: 26–33.
18. Barnett N. P., Tidey J., Murphy J. G., Swift R., Colby S. M. Contingency management for alcohol use reduction: a pilot study using a transdermal alcohol sensor. *Drug Alcohol Depend* 2011; **118**: 391–9.
19. Dougherty D. M., Hill-Kapturczak N., Liang Y., Karns T. E., Cates S. E., Lake S. L. *et al.* Use of continuous transdermal alcohol monitoring during a contingency management procedure to reduce excessive alcohol use. *Drug Alcohol Depend* 2014; **142**: 301–6.
20. Dougherty D. M., Karns T. E., Mullen J., Liang Y., Lake S. L., Roache J. D. *et al.* Transdermal alcohol concentration data collected during a contingency management program to

- reduce at-risk drinking. *Drug Alcohol Depend* 2015; **148**: 77–84.
21. Dougherty D. M., Lake S. L., Hill-Kapturczak N., Liang Y., Karns T. E., Mullen J. *et al.* Using contingency management procedures to reduce at-risk drinking in heavy drinkers. *Alcohol Clin Exp Res* 2015; **39**: 743–51.
 22. US Department of Agriculture and US Department of Health and Human Services. *Dietary Guidelines for Americans*, 7th edn. Washington, DC: US Government Printing Office; 2010.
 23. Pittman B., Gueorguieva R., Krupitsky E., Rudenko A. A., Flannery B. A., Krystal J. H. 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**: 612–8.
 24. Sullivan J. T., Sykora K., Schneiderman J., Naranjo C. A., Sellers E. M. Assessment of alcohol withdrawal: the revised clinical institute withdrawal assessment for alcohol scale (CIWA-Ar). *Br J Addict Alcohol Other Drugs* 1989; **84**: 1353–7.
 25. First M. B., Gibbon M., Spitzer R. L., Williams J. B. W. *User's Guide for the Structured Clinical Interview for DSM-IV Axis I Disorders—Research Version*. New York: Biometrics Research Department, New York State Psychiatric Institute; 1996.
 26. McLellan A. T., Alterman A. I., Cacciola J., Metzger D., O'Brien C. P. A new measure of substance abuse treatment. Initial studies of the treatment services review. *J Nerv Ment Dis* 1992; **180**: 101–10.
 27. Barnett N. P., Meade E. B., Glynn T. R. Predictors of detection of alcohol use episodes using a transdermal alcohol sensor. *J Exp Clin Psychopharmacol* 2014; **22**: 86–96.
 28. Sobell L. C., Sobell M. B. Timeline follow-back: a technique for assessing self-reported alcohol consumption. In: Litten R. Z., Allen J. P., editors. *Measuring Alcohol Consumption: Psychosocial and Biochemical Methods*. Totowa, NJ: Humana Press; 1992, pp. 41–72.
 29. Sobell L. C., Sobell M. B. *Alcohol Timeline Followback Users' Manual*. Toronto, Canada: Addiction Research Foundation; 1995.
 30. Matthews D. B., Miller W. R. Estimating blood alcohol concentration: two computer programs and their applications in therapy and research. *Addict Behav* 1979; **4**: 55–60.
 31. Petry N. M. A comprehensive guide to the application of contingency management procedures in clinical settings. *Drug Alcohol Depend* 2000; **58**: 9–25.
 32. Roll J. M., Shoptaw S. Contingency management: schedule effects. *Psychiatry Res* 2006; **144**: 91–3.
 33. National Institute on Alcohol Abuse and Alcoholism. Drinking Levels Defined; 2015. Available at: <http://www.niaaa.nih.gov/alcohol-health/overview-alcohol-consumption/moderate-binge-drinking> (accessed 20 January 2016) (Archived at <http://www.webcitation.org/6egVyXbGK>).
 34. Jatlow P. I., Agro A., Wu R., Nadim H., Toll B. A., Ralevski E. *et al.* Ethyl glucuronide and ethyl sulfate assays in clinical trials, interpretation, and limitations: results of a dose ranging alcohol challenge study and 2 clinical trials. *Alcohol Clin Exp Res* 2014; **38**: 2056–65.
 35. Anton R. F. Commentary on: ethyl glucuronide and ethyl sulfate assays in clinical trials, interpretation, and limitations: results of a dose ranging alcohol challenge study and 2 clinical trials. *Alcohol Clin Exp Res* 2014; **38**: 1826–8.
 36. Festinger D. S., Marlowe D. B., Croft J. R., Dugosh K. L., Mastro N. K., Lee P. A. *et al.* Do research payments precipitate drug use or coerce participation? *Drug Alcohol Depend* 2005; **78**: 275–81.
 37. Rosado J., Sigmon S. C., Jones H. E., Stitzer M. L. Cash value of voucher reinforcers in pregnant drug-dependent women. *Exp Clin Psychopharmacol* 2005; **13**: 41–7.
 38. Festinger D. S., Dugosh K. L., Kirby K. C., Seymour B. L. Contingency management for cocaine treatment: cash vs. vouchers. *J Subst Abuse Treat* 2014; **47**: 168–74.
 39. Petry N. M., Peirce J. M., Stitzer M. L., Blaine J., Roll J. M., Cohen A. *et al.* Effect of prize-based incentives on outcomes in stimulant abusers in outpatient psychosocial treatment programs: a national drug abuse treatment clinical trials network study. *Arch Gen Psychiatry* 2005; **62**: 1148–56.
 40. Petry N. M., Alessi S. M., Hanson T., Sierra S. Randomized trial of contingent prizes versus vouchers in cocaine-using methadone patients. *J Consult Clin Psychol* 2007; **75**: 983–91.
 41. Peirce J. M., Petry N. M., Stitzer M. L., Blaine J., Kellogg S., Satterfield F. *et al.* Effects of lower-cost incentives on stimulant abstinence in methadone maintenance treatment: a National Drug Abuse Treatment Clinical Trials Network study. *Arch Gen Psychiatry* 2006; **63**: 201–8.
 42. Olmstead T. A., Petry N. M. The cost-effectiveness of prize-based and voucher-based contingency management in a population of cocaine- or opioid-dependent outpatients. *Drug Alcohol Depend* 2009; **102**: 108–15.
 43. Dumett M. A., Rosen I. G., Sabat J., Shaman A., Tempelman L., Wang C. *et al.* Deconvolving an estimate of breath measured blood alcohol concentration from biosensor collected transdermal ethanol data. *Appl Math Comput* 2008; **196**: 724–43.
 44. Rosen I. G., Luczak S. E., Weiss J. Blind deconvolution for distributed parameter systems with unbounded input and output and determining blood alcohol concentration from transdermal biosensor data. *Appl Math Comput* 2014; **231**: 357–76.
 45. Hill-Kapturczak N., Lake S. L., Roache J. D., Cates S. E., Liang Y., Dougherty D. M. Do variable rates of alcohol drinking alter the ability to use transdermal alcohol monitors to estimate peak breath alcohol and total number of drinks? *Alcohol Clin Exp Res* 2014; **38**: 2517–22.
 46. Hill-Kapturczak N., Roache J. D., Liang Y., Karns T. E., Cates S. E., Dougherty D. M. Accounting for sex-related differences in the estimation of breath alcohol concentrations using transdermal alcohol monitoring. *Psychopharmacology (Berl)* 2015; **232**: 115–23.
 47. Dai Z., Rosen I. G., Wang C., Barnett N. P., Luczak S. E. Using drinking data and pharmacokinetic modeling to calibrate transport model and blind deconvolution based data analysis software for transdermal alcohol biosensors. *Math Biosci Eng* 2016; **13**: 911–34.
 48. Barnett N. P., Souza T., Rosen I. G., Luczak S. E., Glynn T. R., Swift R. Transdermal Alcohol Sensor Data Macro, version 1.3 [software]. Providence, RI: Brown University; 2015.
 49. Carroll K. M., Rounsaville B. J. A perfect platform: combining contingency management with medications for drug abuse. *Am J Drug Alcohol Abuse* 2007; **33**: 343–65.
 50. Donovan D. M., Bigelow G. E., Brigham G. S., Carroll K. M., Cohen A. J., Gardin J. G. *et al.* Primary outcome indices in illicit drug dependence treatment research: systematic approach to selection and measurement of drug use endpoints in clinical trials. *Addiction* 2012; **107**: 694–708.