Published in final edited form as:

Alcohol. 2019 December ; 81: 101–110. doi:10.1016/j.alcohol.2018.08.014.

Processing Transdermal Alcohol Concentration (TAC) data to Detect Low-Level Drinking

John D. Roache, Ph.D.^{1,3,*}, Tara E. Karns-Wright, Ph.D.¹, Martin Goros, M.S.², Nathalie Hill-Kapturczak, Ph.D.¹, Charles W. Mathias, Ph.D.¹, Donald M. Dougherty, Ph.D.^{1,3}

¹Department of Psychiatry, The University of Texas Health Science Center at San Antonio San Antonio, TX

²Department of Epidemiology and Biostatistics, The University of Texas Health Science Center at San Antonio San Antonio, TX

³Department of Pharmacology, The University of Texas Health Science Center at San Antonio San Antonio, TX

Abstract

Background.—Several studies have objectively quantified drinking through the use of Alcohol Monitoring System's (AMS) transdermal alcohol concentration (TAC) device known as SCRAM CAM. Criteria that AMS uses to detect drinking are known to be conservative and only reliably detect heavy drinking equivalent to 5 or more standard drinks. Our group has developed Research Rules used to process TAC data in a manner that will detect low-level and moderate drinking even though it is below the AMS criteria for detection.

Methods.—16 male and 14 female paid research volunteers wore TAC monitors for 28 days in their natural environment and responded daily to text message prompts to self-report yesterday's drinking. Current analyses describe the Research Rules that we developed and how use of those rules impact the detection of self-reported drinking treated as a Gold Standard in Sensitivity/Specificity analysis.

Results.—We observed 606 occurrences of positive TAC events over a total of 867 days and processed the TAC data to retain 345 as possible drinking events even though AMS criteria confirmed drinking for only 163 of these events. The kinds of TAC events removed or retained by our rules are illustrated as are cases of low and moderate drinking days that were detected by our rules but not by the conservative AMS criteria. AMS Confirmed TAC events have a high specificity (99.8%) to detect primarily heavy drinking, but have a poor sensitivity to detect lower level drinking and a poor specificity as an indicator of alcohol abstinence. In contrast, our Research Rules detected 100% of TAC events detected by AMS but also detected 31% of the lower-level drinking events not detected by AMS with 91% specificity.

*Corresponding author: John D. Roache, Ph.D. Professor, Departments of Psychiatry and Pharmacology, Psychiatry Division of Alcohol and Drug Addiction, and Neurobehavioral Research Laboratory & Clinic (NRLC), The University of Texas Health Science Center at San Antonio, 7703 Floyd Curl Drive, San Antonio, TX 78229-3900, USA, roache@uthscsa.edu, Phone (210) 562-5406; Fax (210) 562-5403.

None of the authors have conflict of interests concerning this manuscript.

Conclusions.—Reliance upon the AMS criteria for alcohol detection affords a high specificity for detection of heavy drinking but is a poor indicator of abstinence rates. In contrast use of our Research Rules provides more sensitive means to quantify either any drinking or low-moderate levels of drinking while still maintaining good specificity.

Keywords

Alcohol Consumption; Transdermal Alcohol Monitoring; Transdermal Alcohol Concentration (TAC); Heavy Drinking; low-level Drinking; sensitivity; specificity

Introduction

The National Institute on Alcohol Abuse and Alcoholism (NIAAA) has challenged entrepreneurs to develop a wearable alcohol biosensor (Challenge.gov, 2015; 2016). Though NIAAA prioritized devices that would non-invasively measure alcohol in the blood or other bodily fluids, the winners of these challenges (BACtrack and Quantac, in 2015 and 2016, respectively) designed devices that measure transdermal alcohol concentrations (TAC) using wearable wrist monitors. However, these devices are still in the prototype stages and have not been validated. One of the earliest and the most validated devices on the commercial market today to measure TAC (Marques and McKnight, 2007, 2009; McKnight et al., 2012) is the Secure Continuous Remote Alcohol Monitor (SCRAM), produced by Alcohol Monitoring Systems (AMS) Inc. (Highlands Ranch, CO). SCRAM automatically records TAC data every 30 minutes, 24-hr/day, 7 days/week in a digital format which is uploaded to an AMS web server. Shortly after upload, TAC data is available for download for research analysis. Additionally, AMS will review higher-level TAC events to determine whether or not drinking has been detected. SCRAM is widely used in the criminal justice system to verify alcohol abstinence in driving while intoxicated (DWI) offenders or other types of offenders who are court-ordered to remain abstinent during pretrial or probation supervision periods. Judges and pretrial/probation/parole officers use TAC monitoring in conjunction with alcohol treatment services to minimize incarceration while protecting the public from repeat offenders (Huddleston et al., 2008; Voas et al., 2011). AMS provides reports to criminal justice officials of all suspected drinking events that have been confirmed by proprietary processing of the TAC data. To achieve high evidentiary standards of confidence, AMS criteria for processing TAC data are deliberately conservative resulting in reliable drinking detection only when approximately five standard drinks are consumed within two to three hours (Barnett et al., 2014; McKnight et al., 2012; Roache et al., 2015). A consequence of this conservative AMS standard is that lower levels of drinking will go undetected (Roache et al., 2015; Karns-Wright et al., 2018).

The exact criteria AMS uses to report a positive drinking event are proprietary, but to our knowledge required 3 or more TAC readings above the 0.02 g/dl threshold and within defined absorption and elimination rate standards (Barnett et al., 2011; 2014). Though conservative criteria preserve the high specificity of a positive drinking event (i.e., avoiding false positive readings) for criminal justice applications, the reduced sensitivity of these criteria mean that investigators using the monitors for other means need to adapt the AMS rules to detect drinking events that do not meet the AMS criteria (Barnett et al., 2011, 2014).

Barnett and colleagues were the first to adapt the criteria to require only a single TAC reading above 0.02 g/dl and only one of the absorption or elimination criteria to be met. Though they acknowledge that even those adapted criteria are still conservative, lower level TAC data below that cut-off were treated as if it represented “abstinence.”

To better understand the true capabilities of SCRAM for use outside the criminal justice system, we performed three experimental studies (Dougherty et al., 2012; Hill-Kapturczak et al., 2014; 2015) wherein known amounts of alcohol (1 to 5 beers) were administered to participants wearing the monitor in a controlled laboratory environment. We showed that lower level drinking of 1–3 beers produced dose-related and time-related TAC readings that went completely undetected by AMS. Further, we observed occasions where AMS reviewed higher level TAC recordings from these known drinking events and reported them as “unconfirmed” and possible “environmental contaminants”. We also noted several instances of TAC events following controlled drinking that were outside of the AMS absorption or elimination criteria. We showed that these lower level positive TAC readings clearly can and do reflect low-level drinking and cannot be assumed to reflect “abstinence” per se (Roache et al., 2015). Thus, we began the process of developing a series of nine “TAC Research Rules” that we have increasingly used to automatically process TAC data for the purpose of detecting lower-level drinking and to remove spurious TAC readings due to possible environmental contamination. The rules were informed by the human laboratory study data as well as our own observations of TAC recordings that AMS was “Unable to Confirm” or determined to be environmental contamination.

Previously, we used SCRAM in contingency management studies (Dougherty et al., 2014; 2015b) designed to reduce heavy drinking in which participants reduced their heavy drinking, but continued to drink at low or moderate levels. However, due to the time delay between TAC data upload and final AMS resolution, we began to develop our own Research Rules for processing TAC data in real time to determine whether or not drinking occurred. This has allowed us to deliver contingencies in a timely fashion with more confidence. Recently, we used those rules in an outpatient contingency management clinic designed to suppress drinking among DWI offenders (Mathias et al., 2018). In that quality improvement project, data was collected and analyzed across four cycles of quality improvement to find the best TAC criteria and contingent payment strategy for this population.

Recently, we also used our TAC Research Rules to examine the correspondence between processed TAC data and daily self-reported drinking (Karns-Wright et al., 2018). Assuming that self-reported drinking is the appropriate “gold-standard,” that study showed that there is an increased sensitivity and only a slight decrease in the specificity to detect non-heavy, lower level drinking using the TAC Research Rules for drinking detection. Though specificity may be critical in criminal justice applications of this technology (Barnett et al., 2011; 2014; Marques and McKnight, 2009), many research and perhaps even treatment programs may be interested in the detection of lower-level drinking. Similarly, abstinence-based treatments would be well-advised to understand that TAC readings which are below the AMS threshold to confirm a drinking event do not indicate abstinence. Given NIAAA initiatives to study the health benefits and risks of non-heavy, low-moderate drinking

(NIAAA, 2000), the field clearly needs to understand how to interpret TAC readings associated with such drinking.

The current study describes the TAC Research Rules developed by our group to process TAC data as well as how their use affects the determination that a drinking event has occurred in a study wherein paid research volunteers were instructed to engage in their normal drinking behaviors while wearing SCRAM monitors for 28 days.

Methods

Study design.

This study actually is an analysis of data collected from a previous study (Karns-Wright et al., 2018) conducted between March 2015 and December 2015. In that study, research volunteers recruited through community advertising were paid to: i) wear the SCRAM monitors for 28 days, ii) received a link to a website daily to self-report their drinking from the previous day, and iii) come to the laboratory once per week for four weeks to upload TAC data via SCRAM Direct Connect™. Participants were instructed to drink as they normally would over the 28-day study period and that the purpose of the SCRAM monitors was to evaluate natural patterns of alcohol drinking. The previous study (Karns-Wright et al., 2018) described the methods of data acquisition in greater detail but below is a summary of pertinent procedures.

Participants.

Participants were 30 adult (>21 years of age) men (16) and women (14) who were physically healthy (by medical history) and did not have psychiatric disorders as determined by interview using the Structured Clinical Interview for DSM-IV-TR Axis I Disorders research version (SCID-I/NP; First et al., 2001) and Modified SCID for Substance Use (mSCID; Martin et al., 1995). Study inclusion required self-reports of regular patterns of drinking at least once per week with at least 1 day per month of “heavy” drinking (defined as 4 standard drinks for females and 5 for males), as determined by Timeline Follow-back Interview (TLFB; Sobell & Sobell, 1992). The drinking patterns of individual participants assessed in this study are shown in supplement table S1. Study eligibility also required participants to have Internet access and email accounts so that they could take part in a web-based online survey of daily drinking. Participants were excluded if they exhibited alcohol withdrawal symptoms (≥ 10 on the Clinical Institute Withdrawal Assessment for Alcohol-revised; CIWA-Ar; Sullivan et al., 1989), urine positive drug tests for cocaine, opiates, methamphetamines, barbiturates, or benzodiazepines (Panel/Dip Drugs of Abuse Testing Device, Redwood Biotech, Santa Rosa, CA), or a positive pregnancy test (Redi Screen™ Pregnancy Test Strips, Redwood Biotech, Santa Rosa, CA) or breastfeeding for women (medical history). The Institutional Review Board at The University of Texas Health Science Center at San Antonio approved the experimental protocol and all participants signed a written informed consent.

Monitoring Transdermal Alcohol Concentration (TAC).

After consenting and screening, eligible participants were invited back for Day 2 where they were fitted with a newly calibrated SCRAM CAM™ (SCRAM, Alcohol Monitoring Systems Inc., Highlands Ranch, CO) ankle bracelet, which they wore continuously for the next 28 days in their natural environment. Procedures for TAC monitoring using the SCRAM have been described previously (Dougherty et al., 2014; 2015b; Karns-Wright et al., 2018) and TAC data were read from the monitors at once weekly laboratory visits.

AMS resolutions of TAC readings for drinking confirmation.

After the data are uploaded from the SCRAM monitor to the secure web server via Direct Connect®, AMS reviews only those positive TAC readings that have at least 3 consecutive TAC readings > 0.02 g/dl. AMS review resolves the TAC readings as either “Confirmed” or “Unable to Confirm,” indicating that drinking has occurred using their proprietary decisions taking into account factors related to the absorption and elimination rates, as well as evidence of environmental contamination or tampering with the device. (See Barnett, Meade, & Glynn, 2014; Roache et al., 2015; McKnight et al. (2012); for further description of AMS resolution criteria.)

Processing of TAC data using laboratory Research Rules.

We developed 9 TAC Research Rules, which are listed in Table 1 and described in more detail in the results section of this paper. These rules were developed to automatically remove TAC events exhibiting profiles that previous experience deemed likely due to environmental contamination of the device monitor. Rules were informed principally by observations of TAC curves generated in our laboratory studies in which known doses of alcohol (1 to 5 beers) were administered under controlled conditions (Dougherty et al., 2012; Hill-Kapturczak et al., 2014; 2015). However, we also had multiple conversations with AMS regarding our laboratory observations in an attempt to understand their proprietary criteria to identify environmental contamination. In the absence of any exposure to any alcohol, TAC readings will read zero every 30 minutes for 24 hours per day, 7 days per week. However, in the presence of consumed alcohol, or due to environmental contaminants or device tampering, positive TAC readings will occur. For our analysis, we defined a possible TAC event (Rule 1) as any non-zero (positive) TAC reading that was preceded by and followed by at least two zero TAC readings. Each and every possible TAC event was successively numbered for each participant during the 28-day observation period and all of these TAC events were subjected to the analysis of the TAC Research Rules as described herein. With this analysis, we first examined the point-to-point slopes (Rules 2–3 apply to change in TAC reading divided by the change in time in hours) to remove TAC reading points that cause implausibly large point-to-point variation within a possible TAC event (i.e., impossibly large spikes in TAC values within a series of positive TAC readings). After individual points are removed, Rules 4–9 then are then successively applied to consider whether to remove or retain a possible TAC event.

Because drinking typically takes place during the nighttime hours and there is a 2.5–4.0 hour time delay in the onset of TAC data relative to the start of drinking (Barnett et al., 2017; Karns-Wright, et al., 2017), all data were analyzed based on a noon-to-noon 24 hour period

rather than a midnight-to-midnight calendar day. In our previous analyses this attribution has proven reasonable, because even with calendar-based timeline follow-back methods, drinking is typically attributed to a particular day or “night”, even when drinking occurred past midnight. Because drinking binges may span several days, we adopted the convention of attributing any positive TAC or self-reported drinking to the noon-to-noon day on which the drinking/TAC readings began.

Online Self-report of Alcohol Consumption.

Daily, at approximately 9:00 am each day, participants received an e-mail link uniquely for them to complete a survey. Participants were given instructions on how to complete their daily alcohol consumption on a web-based survey (i.e., Survey Gizmo; Boulder, CO). Participants were asked if they drank yesterday. If they selected “Yes,” they were asked how many standard-size cans/bottles (graphically displayed) of beer (12 oz), wine (5 oz), shots (1.5 oz), and mixed or straight drinks (with 1.5 oz hard liquor) as well as the brand of each type of alcohol consumed. In addition, they were asked the start and finish time of the drinking event. From these data, daily self-reports of drinking were classified as belonging to one of three categories: (1) **None**: no drinking; (2) **Moderate**: greater than 0 drinks but < 5 self-reported drinks for men or < 4 drinks for women; and (3) **Heavy**: 5 self-reported drinks for men or 4 self-reported drinks for women. In order for the web-based self-report to be considered valid, participants had to login and complete their survey within the calendar date of the email sending them a link.

Participant Payment.

Participants received \$45 for each weekly laboratory visit, \$5 per day for wearing the transdermal alcohol monitor, \$5 per day they reported their drinking (amount, type, time) on the online self-report survey, and a \$25 weekly bonus for completing all the days of the week.

Data Analyses.

One participant completed only 2 weeks of monitoring before he discontinued the protocol for personal reasons. In order to insure 28 complete days of monitoring, most participants wore the SCRAM device for 29–30 days with the days of putting on or removing the device being only partial days. The final analysis included a total of 870 participant-days for analysis. Descriptive statistics (counts, frequencies, means, standard errors) were used to demonstrate how the Research Rules impacted the processing of TAC event data in our laboratory. For the estimation of frequencies of occurrence, the total number of participant-days includes all observed TAC events, or all self-reports of alcohol consumption, and all participant days on which no TAC recordings were observed or no self-reports of drinking occurred.

Each of the 9 Research Rules were successively applied to: 1) define the occurrence of a TAC event; 2) remove potentially aberrant points from within a TAC event; and finally 3) remove a TAC event from further consideration as a possible drinking event. In order to evaluate the significance of using the Research Rules, we compared TAC event determinations from our rules with the resolutions emanating from AMS, the device

manufacturer. The sensitivity and specificity analysis used the daily drinking self-report data as a “Gold Standard” to designate true positives (TP) and true negatives (TN) and identify the false positives (FP) and false negatives (FN) incorrectly designated by AMS or our Research Rules. Calculations assumed that Sensitivity % = $TP/(TP+FN)*100$ and Specificity % = $TN/(TN+FP)*100$.

Results

Participant Characteristics.

Participants ranged in age from 21–54 (mean = 28.3) years. Consistent with San Antonio demographics, the sample was predominately white (63.3%) of Hispanic origin (56.7%). The majority (80%) had at least some college and drank 8.3 ± 7.9 standard drinks per week pre-study entry. A total of 30 participants completed a total of 870 days on protocol, and completed timely alcohol self-reports of yesterday’s drinking on 83% of those days.

TAC Research Rules

Each of the 9 TAC Research Rules were successively applied to TAC data collected during a 28-day observation period in which we also had daily self-reports of drinking. Below are descriptions of the TAC Research Rules and their impact on examples where pertinent.

Rule 1: A possible Event begins with a positive TAC value (TAC > 0) proceeded by and followed by at least two consecutive TAC = 0 values.—Rule 1 is used to define a possible TAC event, and in this study, a total of 606 TAC events were identified. Of these, successive application of the rules ultimately removed 261 TAC events leaving 345 of them considered as possible drinking events. Fig. 1 shows the distribution of possible TAC events for each individual participant.

Though we code each event as beginning and ending with two TAC = 0 values, there could be any number of positive TAC readings ranging from one to many dozens of positive TAC readings which can happen when a binge episode may span 2–3 days, etc. Because we distinguish between individual TAC events by requiring at least two successive TAC = 0 readings, it also is possible to have more than one TAC event within a 24 hour period. Fig. 2 illustrates different kinds of TAC events, all of which were associated with self-reported drinking. Fig. 2a shows a short, low-level TAC event that was never reviewed or resolved by AMS. Fig. 2b shows a more substantial TAC event that AMS did resolve but was “Unable to Confirm” as a drinking event. Figs. 2c and 2d show examples of multiple TAC events occurring within the same day. Fig. 2c shows a small event (#35) on Thursday night and then a second more substantial event (#36) associated with late night early Friday morning drinking for a female participant. AMS reviewed both events, but only “Confirmed” one of them. Fig. 2d shows two lower level events for a woman who reported drinking a total of 6.9 drinks spread out during a Saturday afternoon and evening – neither of these events was reviewed by AMS.

Rule 2: Remove repeated high frequency TAC readings with < 20 min inter-reading interval.—Though SCRAM is supposed to auto-sample TAC levels every 30

minutes, there are occasions of more or less frequent readings. In the current analysis, 95.3% of the readings occurred within 30–32 minutes since the last reading with another 3.8% of the readings occurring between 32–36 minutes since the last reading. As we developed programming rules, we noticed high frequency repetitions of monitor readings in 2–6 minute intervals sometimes occurred on the first or second day of monitoring. We created Rule 2 to remove any repetitive point readings of less than 20 min inter-reading intervals. The current analysis found 167 such readings observed in 20 participants. Most of those (72%) did occur in the first few hours of wearing the monitor on the first day, but they sporadically happened on other days as well. For all but two of these high frequency readings, the TAC readings were zero. The two non-zero TAC readings occurred in the midst of a drinking event of a single participant and recorded TAC readings within the range of other proximal readings, but at one-minute inter-reading intervals. Thus, Rule 2 appears mostly unnecessary as it had little impact on determination of any TAC event.

Rule 3: Remove steep point-to-point slope rises > 0.182 g/dl/hour or drops < -0.126 g/dl/hr unless the < -0.126 g/dl/hr is the last TAC reading leading to a zero TAC of an event.—Rule 3 was developed to remove extreme and irregular spikes and dips in the TAC curves. Although these slopes are greater than those stated by AMS as “physiological”, the slope values of $+0.182$ g/dl/hr or -0.126 g/dl/hr were based upon the maximum point-to-point slope of changes in TAC value (g/dl/hr) we observed in our previous human laboratory studies where known amounts of alcohol were consumed under controlled conditions (Dougherty et al., 2012, 2015a; Hill-Kapturczak et al., 2014). In the current study, a total of 22 participants had 54 positive TAC events from which 120 individual TAC readings (points) were removed by Rule 3. Notably, only 10 (18.5%) of these were points coming from events that were removed by subsequent rules, and 33 (61.1%) were from events confirmed by AMS as a drinking event. This indicates that most of the TAC points removed by Rule 3 came from events that were either retained by our rules or confirmed by AMS as drinking events. Finally, there were 18 events (33.3%) where one of the removed TAC points was greater than the eventual peak TAC value, meaning that we removed an extreme or high point, which resulted in lower peak TAC values. A complete count of points removed by Rule 3 is shown in supplement table S2. Most participants had only a few points removed from a few events, however two participants (M3001, shown in Fig 3C and M3030, data not shown) had many more points removed than other participants suggesting there are individual differences in the likelihood of steeply sloping changes in the concentration-time curves for TAC.

Fig. 3 shows examples of points being removed from TAC events. In Fig. 3A an extreme point was removed from an AMS confirmed event which substantially changed the peak TAC estimate. Fig. 3B shows another extreme point removed from an event that ultimately was removed by Rule 7 (described below). Fig. 3C shows a single TAC event which spanned three days and from which 10 points were removed by Rule 3. The event reflects heavy drinking Friday night and TAC never fell to zero before drinking resumed Saturday night. The participant failed to complete a valid self-report on Saturday morning, but on Sunday morning reported drinking 22 standard drinks “yesterday.” This AMS Confirmed event had

more points removed than any other TAC event though point removal had little impact on either the shape of the curves or the peak TAC observation.

Removal of events with Rules 4–9.—Rules 2 and 3 only removed individual points from a TAC event but Rules 4–9 were used to remove entire TAC events from consideration as possible drinking events. The distribution of events removed for all participants is also shown in Fig. 1. Note that every participant had at least 2 TAC events removed and as many as 18 events were removed from a single participant. Rule 4 removed only one event in each of only four participants (only 2%). Its purpose was to remove events when the very first TAC reading was implausibly high given the known pharmacokinetics of consumed alcohol. Rule 5 removed the most events (61%) because those events had only a single positive TAC reading. Rule 6 was developed to be more stringent than Rules 4 and 5. However, in the current study it did not remove any event because programming algorithms executed rules one at a time in sequence so that when lower numbered rules removed an event, the event was no longer subjected to the additional Rules. The next most likely kind of event to be removed (28%) occurred when the first TAC reading was positive and all subsequent readings were negatively sloping (i.e., Rule 7). An example of this kind of event is also shown in Fig. 3B.

Rule 8 did not remove any events in the current study, but it was intended to remove short TAC span events having only 2–3 observations with implausible slopes given the known pharmacokinetics of consumed alcohol. Rule 9 removed 10% of events due to persistent, low level TAC readings (peak TAC < 0.01 and TAC event duration (“TAC span”) > 240 minutes). Rules 4–9 were all based upon feedback from AMS but modified to be consistent with observations from our laboratory and naturalistic studies as well.

Figs. 2 and 3 also illustrate different types of TAC events that our TAC Research Rules retained even though most of these were never considered or were rejected by AMS review. Fig. 2A shows low level TAC readings (peak TAC < 0.01) associated with the self-report of 3 drinks in a female participant. This event is well below the threshold for AMS review. Fig. 2B shows a high TAC-level (peak TAC = 0.099) event that AMS was “Unable to Confirm.” AMS decision algorithms could not rule out “possible environmental contamination” but our TAC Research Rules retained the event and M3045 reported heavy drinking (5.5 drinks) that day. Fig. 2c shows two TAC events occurring within the same noon-to-noon day in F3059 who reported drinking heavily (6 drinks) Thursday evening and into the early morning Friday. AMS considered both of these TAC events, but was “Unable to Confirm” Event #35 even though they did “Confirm” Event #36. Fig. 2d shows two lower-level TAC events occurring on the same day in F3061 who reported consuming 6.9 drinks spread out during Saturday afternoon and evening. Again, neither of these TAC events achieved AMS criteria for review.

Correspondence between Research Rules and AMS determinations of TAC events—Table 2 reveals all 606 positive TAC events and their disposition according to whether our Research Rules “Removed” or “Retained” the event, and how that compares with the AMS Resolutions. Even though our rules were enacted without regard to AMS determinations, none of the 163 AMS confirmed events were removed by our rules. There

were 53 Events that AMS was “Unable to Confirm,” but our rules retained 45 of those and only removed 8. Finally, our rules removed 253 of the remaining 390 positive TAC events, none of which were reviewed by AMS; that left 137 TAC events retained by our rules to potentially detect lower-level TAC events such as were illustrated in Fig. 2a and 2c.

Table 3 shows the peak TAC values and TAC span (# minutes between the first and last positive TAC reading) for AMS resolutions of TAC events and for the remaining TAC events retained or removed by the rules. Because only the more substantial TAC events are reviewed by AMS in the first place, both the “Confirmed” and “Unable to Confirm” AMS resolutions generally had higher peak TAC values and longer TAC span durations than the remaining events that were considered only by our rules. Of the 137 lower-level events retained and the 261 events removed by our rules, removed events were more likely to be lower peak and shorter span TAC events than those that were retained.

Correspondence between self-reported drinking and TAC event dispositions—

Table 4 shows the extent to which self-reported drinking corresponded with various criterion definitions of a TAC event. This table includes all participants and days and events as defined by TAC event classification. TAC events spanning more than one day were considered for each day of self-reported drinking. For days containing more than one TAC event within the same day, the analysis prioritized AMS “Confirmed” events over “Unable to Confirm” events and events retained by Research Rules were prioritized over events removed by the rules, etc. This process resulted in a total of 917 day-TAC events included in the analysis. The majority (55%) of AMS Confirmed TAC events were associated with self-reports of heavy drinking (only 22% were days reported as less than heavy drinking). The 51 Day-events that AMS reviewed but did not confirm were uniformly distributed across self-report categories with 25% being heavy drinking days. Of the 182 occasions where TAC events were retained by our rules but were not confirmed by AMS, the largest proportion (42%) were days of non-heavy drinking, and only 21% were days reported as not drinking. Finally, of the 170 days where TAC events were removed by the Research Rules and there was no other kind of TAC event, the majority (56%) were days reported as not drinking and only 2% of them were days of heavy drinking. There were 392 days in which no TAC event of any kind occurred and 71% of those were associated with self-reports of not drinking, though 17% of days did have reported drinking. The majority (67%) of the 562 days that contained no AMS Confirmed or Rules-retained events were self-reported as not drinking.

Table 4 also shows the sensitivity and specificity of AMS and Rules-retained TAC events to detect “Any” drinking (i.e., Heavy + Non-heavy) treating drinking self-reports as the Gold Standard and the absence of AMS or rules-retained TAC events as a negative Test Detection. AMS confirmed drinking had a very high specificity (99.8%) but a low sensitivity (39.9%) for detection. We also performed the same analysis for the rules-retained events that were not already AMS confirmed and obtained similar levels of sensitivity and specificity to detect other drinking events – none of which were detected by AMS confirmation. If we had used the rules only without parsing the AMS Confirmed events, the Rules alone (not shown in table) would yield a Specificity of 90.4% and Sensitivity of 68.5%. Finally, because of the common interest in using TAC to confirm abstinence, we treated the last column of 562 days in which no TAC was retained by our rules or detected by AMS as if it were an objective test

of alcohol abstinence. As a test to detect no drinking after the application of our TAC Research Rules, we would obtain a sensitivity of 68.5% and specificity of 90.4% to detect abstinence. If we had used the absence of AMS confirmation to detect abstinence (not shown in table), there was a 62.1% sensitivity to detect abstinence, but only a 2% specificity of correct detection of abstinence.

Discussion

This study supports the value of using a set of Research Rules to process TAC event data to study drinking behavior in an outpatient sample of drinkers wearing the SCRAM monitor in their natural environments. Specifically we have shown an ability to remove or retain TAC events using more inclusive criteria than the conservative standards employed by the manufacturer (AMS). We have: 1) characterized rules that are programmed to automatically process TAC event recordings and make on site decisions to retain or remove TAC events as possible indicators of drinking without relying upon AMS determinations; 2) illustrated the kinds of TAC events retained by our rules as possible drinking events but excluded by the AMS resolution and confirmation process; 3) shown that Research Rules can be used to remove from consideration TAC events that are likely to be false positive readings not associated with drinking; 4) demonstrated that by removing certain TAC events and retaining others, we are able to gain sensitivity to detect lower-level, non-heavy drinking that AMS criteria would not detect and that our Rules detected with good (90.6%) Specificity; and 5) showed clearly that the absence of an AMS confirmed drinking event does not indicate abstinence from drinking.

For evidentiary standards, it is imperative to avoid false positive test outcomes. Because the “on label” use of SCRAM is for alcohol detection in criminal justice applications, AMS has deliberately set conservative thresholds for the detection of a drinking event to preserve the specificity of their test as a detector of drinking events. The current data as well as previous work (Barnett et al., 2014; McKnight et al., 2012; Roache et al., 2015; Karns-Wright et al., 2018) support a very high specificity of positive drinking detections using AMS. Indeed, we found a 99.8% specificity of detection of “any” drinking with most (55%) of those being episodes of “heavy” drinking. This outcome, in context with the previous literature, lends further documentation supporting AMS criteria meeting the principals set forth in *Frye vs. the United States* (293 F. 1013; D.C. Cir., 1923) and *Daubert v. Merrel Dow Pharmaceuticals, Inc* (43 F.3d 1311; 9th Cir., 1995) regarding evidentiary standards. However, perfect specificity comes at a cost of sensitivity. A consequence of the conservative AMS standard is that lower levels of drinking will go undetected. The current study illustrated drinking-related TAC events that were not detected by AMS and quantified the prevalence of “non-heavy” drinking among these events clearly showing that the absence of AMS detection does not indicate abstinence from alcohol. We estimated that if the absence of AMS confirmation were used as evidence of abstinence, there would be only 2% specificity of SCRAM as an abstinence detector. Therefore we conclude that the absence of AMS confirmation should not be interpreted as abstinence from alcohol. This has important implications for researchers or clinicians who may want to detect “any” drinking and not just heavy drinking as well as for the criminal justice system who should understand that SCRAM devices cannot be used to confirm abstinence in DWI Offenders.

Outside of forensic evidentiary requirements, other testing contexts may desire greater emphasis on the sensitivity component of test detection. Therefore, researchers have adapted the AMS criteria to liberalize TAC detection thresholds to increase the sensitivity for detection of possible drinking. For instance, Barnett and colleagues (2011) required only a single TAC reading above 0.02 g/dl and required only that either the absorption rate from 0 to peak TAC was <0.05 g/dl/hr or that the elimination rate (from peak to zero) was and <0.025 g/dl/hr if peak TAC <0.15 g/dl or <0.035 g/dl/hr if peak TAC >0.15 g/dl. More recently, Barnett and colleagues have made available a program to process TAC data using their rules and approaches for the interpretation of TAC events and translating them into estimated blood alcohol concentrations (Barnett et al., 2015). However, those rules continue to use the 0.02 g/dl cut-off for consideration of possible drinking events. In contrast, the current results combined with our controlled laboratory studies (Roache et al., 2015) and the recently reported study (Karns-Wright et al., 2018), showed that many lower-level TAC readings are indeed associated with lower level drinking. The current results and previous work (Roache et al., 2015; Karns-Wright et al., 2018) strongly argue that some clinical and research uses of TAC monitors require adoption of less conservative TAC detection thresholds in order to gain sensitivity to detect lower level or moderate, but non-abstinent drinking.

In order to preserve sensitivity for drinking detection from SCRAM-based TAC data, our strategy has been to consider all positive TAC events as possible drinking and to develop rules to remove those events which are mostly likely to be false positives. Using TAC data observed in laboratory studies, we developed algorithms to estimate breath alcohol concentrations (Dougherty et al., 2012; Hill-Kapturczak et al., 2014; 2015) and the number of standard drinks consumed (Dougherty et al., 2015a). We also reported using contingencies to suppress heavy drinking days, but observed that participants increased the prevalence of non-heavy drinking days (Dougherty et al., 2014; 2015b). Then, as we began to develop rules that could remove suspect TAC events and allow retention of other lower-level TAC events, we piloted the use of these rules in a clinical program evaluation of contingency management to promote abstinence in DWI offenders (Mathias et al., 2018). A significant finding of the current study is the demonstration that our Research Rules removed a total of 261 of 606 non-zero TAC events (43%) most of which were associated with days of self-reported abstinence. The Sensitivity and Specificity analysis of TAC events retained by our rules, but not confirmed by AMS, showed that 57% were associated with any drinking, most of which was non-heavy drinking, resulting in an overall 90.6% specificity for detection of any drinking which is very good for detection of mostly moderate drinking completely missed by AMS criteria.

Given the potential for environmental contamination, it is not surprising that TAC monitors record non-zero TAC events that should not be confirmed as drinking events. Environmental contamination from topical or aerosolized alcohol can cause positive TAC readings and these may be commonly encountered in environments including beauty/hair salons, bars, and probably even bakeries (AMS consultation + personal observations). Environmental contamination exhibit certain characteristic profiles that many of our rules were designed to detect and remove as likely false positive TAC events. These characteristics include persistent low-level TAC readings and sudden spikes or drops in TAC readings that are

physiologically implausible given the known pharmacokinetics of alcohol absorption or elimination. The validity of using rules to remove these type of TAC events is supported by current findings that the majority of events removed by our rules were associated with self-reports of not drinking and only rarely (2%) occurred on days of reported heavy drinking.

This study has several limitations in the generality of our findings. First, we have a limited sample size in a population of non-treatment seeking drinkers who mostly drank heavily only several days per month. We do not know if the same rules would have worked equally well in a population of heavier drinkers. Second, this analysis evaluated the Research Rules together, but did not systematically validate each individual rule or its incremental value to impact the specificity of TAC test results. It may be fruitful to pursue improving and refining these rules in future studies. Third, our analysis of correspondence with self-reported drinking had to make certain assumptions as to what TAC data to include or not include given that there can be multiple TAC events on the same day of self-report or multiple days of self-report during a binge drinking episode that shows up as a single TAC event spanning multiple days. Fourth, though the use of requiring daily self-reports of drinking should improve the validity of self-report (e.g., Hoeppe et al., 2010; Searles et al., 2002), our particular methods have not been validated elsewhere and we do not know how or if that level of self-monitoring might have altered the pattern of drinking in our participants. A final limitation of our Rules is that we do not know how they will apply to some of the new wearable TAC sensors under development. Many of our rules are based upon point-to-point analyses of slopes and most of the new technologies sample at a much faster rate than the once per 30 minutes of the SCRAM device.

Other investigators are using TAC data to quantify levels of drinking. For example, some have use the SCRAM device, and some used the WristAS™ (Wrist Transdermal Alcohol Sensor; Giner, Inc., Newton, MA), and some have compared the two (e.g., Bond et al., 2014; Luczak and Rosen, 2014; Dai et al. 2016). This line of work is very promising as a way to individualize the TAC profile characteristics for known individuals drinking known amounts. We are unaware of how that work or process will manage false positive TAC recordings due to environmental contamination. To our knowledge, ours is the first study to systematically consider all positive TAC events to detect and remove only suspected false positive readings due to environmental contamination. As new technologies for TAC recordings become available due to the NIAAA challenge initiatives for wearable device monitoring (Challenge.gov, 2015; 2016), we suspect many of these same issues will be important and require research to validate how TAC data are processed to draw conclusions regarding possible drinking events. Finally, though SCRAM devices from AMS, Inc. are well-validated for use in the criminal justice system as a specific detector mostly of heavy drinking, we strongly urge all users of SCRAM Technology to be aware that the absence of AMS confirmed drinking is a poor detector of abstinence and misses most instances of less than heavy drinking.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Research reported in this publication was supported by the National Institute on Alcohol Abuse and Alcoholism [R01AA14988 and R01AA022361 to DMD] and the National Institute of Drug Abuse [T32DA031115 for TEK-W] of the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. DMD also gratefully acknowledges support from a research endowment, the William and Marguerite Wurzbach Distinguished Professorship.

References

- Barnett NP, Meade T, and Glynn R (2014) Predictors of Detection of Alcohol Use Episodes Using a Transdermal Alcohol Sensor. *Exp Clin Psychopharm* 22:86–96
- Barnett NP, Souza T, Rosen IG, Luczak SE, Glynn TR, Swift R (2015) Transdermal Alcohol Sensor Data Macro (Version 1.3) [software]. Brown University.
- Barnett NP, Tidey J, Murphy JG, Swift R, Colby SM (2011) Contingency management for alcohol use reduction: a pilot study using a transdermal alcohol sensor, *Drug Alcohol Depend* 118:391–399. [PubMed: 21665385]
- Bond JC, Greenfield TK, Patterson D, Kerr WC (2014) Adjustments for drink size and ethanol content: New results from a self-report diary and transdermal sensor validation study. *Alcohol Clin Exp Res* 38:3060–3067. [PubMed: 25581661]
- Challenge.gov (2015) Design and produce a wearable device to monitor blood alcohol levels in real time. It should be inconspicuous, low profile, and appealing to the wearer. <https://www.challenge.gov/challenge/a-wearable-alcohol-biosensor/>
- Challenge.gov (2016) A wearable alcohol biosensor: A second challenge. <https://www.challenge.gov/challenge/wearable-alcohol-biosensor/>
- Dai A, Rosen IG, Wang C, Barnett NP, Luczak SE (2016) Using drinking data and pharmacokinetic modeling to calibrate transport model and blind deconvolution based data analysis software for transdermal alcohol biosensors. *Mathematical Biosciences and Engineering* 13(5), 911–935. [PubMed: 27775390]
- Dougherty DM, Charles NE, Acheson A, John S, Furr MR, Hill-Kapturczak N (2012) Comparing the detection of transdermal and breath alcohol concentrations during periods of alcohol consumption ranging from moderate drinking to binge drinking. *Exp Clin Psychopharm* 20:373–81.
- Dougherty DM, Hill-Kapturczak N, Liang Y, Karns TE, Cates SE, Lake SL, Mullen J, Roache JD (2014) Use of Continuous Transdermal Alcohol Monitoring during a Contingency Management Procedure to Reduce Excessive Alcohol Use. *Drug Alc Depend* 142:301–306.
- Dougherty DM, Hill-Kapturczak N, Liang Y, Karns TE, Lake SL, Cates SE, Roache JD (2015a) The Potential Clinical Utility of Transdermal Alcohol Monitoring Data to Estimate the Number of Alcohol Drinks Consumed. *Addict Disord & Their Treat* 14:124–130.
- Dougherty DM, Karns TE, Mullen J, Liang Y, Lake SL, Roache JD, Hill-Kapturczak N (2015b) Transdermal Alcohol Concentration Data Collected During a Contingency Management Program to Reduce At-Risk Drinking. *Drug Alc Depend* 148:77–84.
- First MB, Spitzer RL, Gibbon M, Williams JBW (2001) Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition (SCID-I/NP). New York: Biometrics Research, New York State Psychiatric Institute.
- Hill-Kapturczak N, Lake SL, Roache JD, Cates SE, Liang Y, Dougherty DM (2014) Do Variable Rates of Alcohol Drinking Alter the Ability to Use Transdermal Alcohol Monitors to Estimate Peak Breath Alcohol and Total Number of Drinks? *Alc Clin Exp Res* 38:2517–2522
- Hill-Kapturczak N, Roache JD, Liang Y, Karns TE, Cates SE, Dougherty DM (2015) Accounting for Sex-Related Differences in the Estimation of Breath Alcohol Levels using Transdermal Alcohol Monitoring. *Psychopharmacology* 232:115–123. [PubMed: 24923985]
- Hoepfner BB, Stout RL, Jackson KMM, Barnett NP (2010) Giw giid us fube-grained Timeline Follow-back data? Comparing 30-day TLFB and repeated 7-day TLFB alcohol consumption reports on the person and daily level. *Addict Behav* 35:1138–1143. [PubMed: 20822852]

- Huddleston CW, Marlowe DB, Casebolt R (2008, 5) Painting the current picture: A national report card on drug courts and other problem solving court programs in the United States. (Vol II, No. 1). Alexandria, VA: National Drug Court Institute.
- Karns-Wright TE, Dougherty DM, Hill-Kapturczak N, Mathias CW, Roache JD (2018) The correspondence between transdermal alcohol monitoring and daily self-reported alcohol consumption. *Addictive Behaviors* 85:147–152. [PubMed: 29910035]
- Karns-Wright TE, Roache JD, Hill-Kapturczak N, Liang Y, Mullen J, Dougherty DM (2017) Time delays in transdermal alcohol concentrations relative to breath alcohol concentrations. *Alcohol* 52:35–41. [PubMed: 27522029]
- Luczak SE, Rosen IG (2014) Estimating BrAC from transdermal alcohol concentration data using the BrAC estimator software program. *Alcohol Clin Exp Res* 38:2243–2252. [PubMed: 25156615]
- Marques PR, McKnight AS (2007) National Highway Traffic Safety Administration. Evaluating transdermal alcohol measuring devices (Report No. DOT HS 810 875). Washington, DC: U.S. Government
- Marques PR, McKnight AS (2009) Field and laboratory alcohol detection with 2 types of transdermal devices. *Alc Clin Exp Res* 33:703–711. DOI: 10.1111/j.1530-0277.2008.00887.
- Martin CS, Kaczynski NA, Maisto SA, Bukstein OM, Moss HB (1995) Patterns of DSM-IV alcohol abuse and dependence symptoms in adolescent drinkers. *Journal of Studies on Alcohol and Drugs*, 56, 672–80.
- Mathias CW, Hill-Kapturczak N, Karns-Wright TE, Mullen J, Roache JD, Fell JC, Dougherty DM. (2018) Translating transdermal alcohol monitoring procedures for contingency management among adults with DWI. *Addictive Behaviors* 83:56–63. [PubMed: 29397211]
- McKnight AS, Fell JC, Auld-Owens A (2012) Transdermal alcohol monitoring: Case studies. (Report No. DOT HS 811 603). Washington, DC: National Highway Traffic Safety Administration.
- NIAAA: National Institute on Alcohol Abuse and Alcoholism (2000). Health risks and benefits of alcohol consumption. *Alcohol Research & Health* 24:5–11.
- Roache JD, Karns TE, Hill-Kapturczak N, Mullen J, Liang Y, Lamb RJ, Dougherty DM (2015) Using transdermal alcohol monitoring to detect low-level drinking. *Alcohol Clin Exp Res* 39:1120–1127. [PubMed: 25988708]
- Searles JS, Helzer JE, Rose GL, Badger GJ (2002) Concurrent and retrospective reports of alcohol consumption across 30, 90, and 366 days: Interactive voice response compared with timeline follow back. *J Studies Alcohol* 63:352–362.
- Sobell LC, Sobell MB (1992) Timeline follow-back: A technique for assessing self-reported ethanol consumption In Allen J & Litten RX (Eds.), *Measure Alcohol Consumption: Psychosocial and Biological Methods* (pp. 41–72). Human Press: Totowa, NJ.
- Sullivan JT, Sykora K, Schneiderman J, Naranjo CA, Sellers EM (1989) Assessment of alcohol withdrawal: The revised clinical institute withdrawal assessment for alcohol scale (CIWA-Ar). *British Journal of Addiction*, 84:1353–1357. [PubMed: 2597811]
- Voas RB, DuPont RL, Talpins SK, Shea CL (2011) Towards a national model for managing impaired driving offenders. *Addiction* 106:1221–1227. [PubMed: 21205054]

Figure 1:
non-zero TAC events observed. Each bar reflects one of 30 participants. Stippled bars = # events removed, Black bars = # events confirmed by AMS, and Grey bars = # events retained by TAC Research Rules so that the total stack reflects the total # non-zero TAC events observed in study.

Figure 2:

Shown are serial TAC readings observed in 6 TAC events from four different participants. Each panel is labeled with M/F participant number, day of week, self-reported drinking, and AMS resolutions. (A) A low level TAC event observed in a female participant who self-reported 3 drinks on a Saturday evening. (B) A higher level TAC event observed in a male participant who reported drinking 5.5 drinks on a Wednesday - but AMS review was "Unable to Confirm" the event. (C) Two TAC events within the same day for a female

participant drinking 6 drinks. (D) Two TAC events on the same day for a female participant whose drinking was spread out over Saturday afternoon and evening.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 3:

Examples of 3 different TAC events from two participants where TAC points were removed (shown as red stars) by Rule 3. (A) An AMS confirmed event; (B) a TAC event ultimately removed by Rule 7; and (C) a single AMS confirmed event which spanned 3 days during a Friday - Saturday drinking binge. Note that TAC concentrations reached a device maximum at 0.437 g/dl during the early morning hours of Saturday morning.

Table 1:**Research Rules for Processing TAC data**

- Rule 1: A possible Event begins with one or more positive TAC ($TAC > 0$) values proceeded by and followed by at least two consecutive $TAC = 0$ values.
- Rule 2: Remove repeated high frequency TAC readings with < 20 min inter-reading interval.
- Rule 3: Remove steep point-point slope rises > 0.182 or drops < -0.126 unless the < -0.126 is the last TAC reading leading to zero TAC of an event.
- Rule 4: Remove Event if first positive TAC reading > 0.15
- Rule 5: Remove Event if it contains only a single non-zero TAC value.
- Rule 6: Remove TAC Events with only two non-zero points and slope > 0.08 between point 1 and point 2.
- Rule 7: Remove Event if first TAC reading is positive followed only by negative point-to-point slopes thereafter.
- Rule 8: For all Events with TAC Span < 70 , remove if Absorb-Rate > 0.077 or Elim-Rate > 0.035 .
- Rule 9: Remove Event if Peak TAC < 0.01 and TAC Span > 240 minutes.

All references to slope refer to g/dl/hr changes in TAC value. Absorb-Rate = peak TAC /time to peak; Elim-Rate = Peak TAC/time from peak to zero

Table 2.

Correspondence of Rule Outcomes and AMS Resolutions

<u>AMS Resolution</u>	<u>Research Rules Outcomes</u>		(Totals)
	Removed Events	Retained (Not Removed)	
“Confirmed”	0	163	163
“Unable to Confirm”	8	45	53
not reviewed	253	137	390
(Totals)	261	345	606

Table 3.

Description of TAC Events for different Event Determinations

	N	Peak TAC value (g/dl)				TAC span (minutes)			
		Mean	SEM	Min	Max	Mean	SEM	Min	Max
*AMS "Confirmed"	163	0.13	0.0 07	0.0 23	0.4 37	907.5	42 .6	183.5	4044.6
*AMS "Unable to Confirm"	45	0.11	0.0 15	0.0 22	0.4 37	626.4	41 .3	158.1	1286.7
<u>Rules Retained Events (& not reviewed by AMS)</u>									
Peak TAC < 0.01	43	0.007	0.0002	0.005	0.009	134.9	5.8	90.1	213.9
Peak TAC 0.01–0.02	61	0.014	0.0003	0.01	0.019	302.5	31.8	91.3	1435.3
Peak TAC 0.02	33	0.026	0.001	0.02	0.049	314	37.4	91.7	977.9
Total	137								
<u>Events Removed by Rules</u>									
Peak TAC < 0.01	190	0.006	0.0001	0	0.009	113.1	8.7	60.6	946.6
Peak TAC 0.01–0.02	50	0.013	0.0003	0.01	0.019	93.8	6.7	60.7	213.3
Peak TAC 0.02	21	0.105	0.027	0.021	0.044	192.9	31.9	61.1	519.02
Total	261								

Data are sample size (N), Mean, standard error (SEM), and the minimum (Min) and maximum (Max) values for the group. "TAC span" refers to the number of minutes between the first and last positive TAC reading within a defined event. "Peak TAC" values refer to the g/dl TAC reading at the peak within an event.

Table 4.

Correspondence of Self-Reported Drinking and various TAC Event Dispositions

	AMS “Confirmed”	AMS “Unable to Confirm”	Rules- Retained Events (not AMS confirmed)	Days containing only Events Removed by Rules	Days of no TAC Event	Days only having no TAC or Removed TAC events
# Day-TAC Observations	170	51	185	170	392	562
<u>Self-Reports of Drinking</u>						
not valid	39 (23%)	14 (27%)	41 (22%)	30 (18%)	49 (13%)	79 (14%)
not drinking	1 (1%)	12 (24%)	39 (21%)	96 (56%)	279 (71%)	375 (66.7)
Non-heavy drinking	37 (22%)	12 (24%)	78 (42%)	40 (24%)	53 (14%)	53 (14%)
Heavy drinking	93 (55%)	13 (25%)	27 (15%)	4 (2%)	11 (3%)	11 (3%)
Any drinking	130 (76%)	25 (49%)	100 (57%)	44 (26%)	64 (17%)	64 (17%)
	<u>Tests for Detection of “Any” Drinking</u>					<u>Detecting No Drinking</u>
Sensitivity	39.9%	-	30.6%	-	-	68.5%
Specificity	99.8%	-	90.6%	-	-	90.4%

The number of removed and retained events do not add to the same numbers as in Tables 2 and 3 because of days where multiple events occurred on the same day and removed event correspondence with drinking self-report can only be evaluated when no other event was retained by our rules or confirmed AMS. Days on which participants failed to complete a valid self-report are removed (listed as not valid), and completed self-reports are categorized as reports of not drinking (= 0 drinks), heavy drinking (4 for females or 5 for males), non-heavy drinking (> 0 drinks, but not heavy) or any drinking (heavy + non-heavy).