

EVALUATING TRANSDERMAL ALCOHOL MEASURING DEVICES

Final Report

November 2007



1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Evaluating Transdermal Alcohol Measuring Devices		5. Report Date November 2007	6. Performing Organization Code
7. Author(s) Paul R. Marques and A. Scott McKnight		8. Performing Organization Report No.	
9. Performing Organization Name and Address Pacific Institute for Research and Evaluation 11720 Beltsville Drive, Suite 900, Calverton, MD 20705 Phone: 301-755-2700 Fax: 301-755-2799		10. Work Unit No. (TRAIS)	11. Contract or Grant No. DTNH22-02-D-95121
12. Sponsoring Agency Name and Address National Highway Traffic Safety Administration 1200 New Jersey Avenue SE., Washington, DC 20590		13. Type of Report and Period Covered Final Report (8/31/04 – 2/28/06)	14. Sponsoring Agency Code
15. Supplementary Notes James F. Frank, initial COTR; DeCarlo Ciccel, final COTR			
16. Abstract This report is an evaluation study of two types of transdermal devices that detect alcohol at the skin surface representing two types of electrochemical sensing technology. The AMS SCRAM™ ankle device and the Giner WrisTAS™ wrist device were worn concurrently for the evaluation by 22 paid research subjects (15 males, 7 females), for a combined total of 96 weeks. Each subject participated in both laboratory drinking to .08 grams per deciliter (g/dL) BAC and normal drinking on their own. A total of 271 drinking episodes with BAC ≥ .02 g/dL were logged: 60 were from laboratory dosing, and 211 were from self-dosed drinking. Both devices detected alcohol at the skin surface. The SCRAM™ unit has security features and automated reporting protocols that make it suitable for the offender market, whereas the WrisTAS™ unit is a research prototype that has had trials as an aid to detection for alcohol treatment settings. Neither unit had false-positive problems when true BAC was <.02 g/dL. False negatives were defined as TAC (transdermal alcohol concentration) response <.02 g/dL when true BAC ≥ .02 g/dL. Overall, the true-positive hit rate detected by WrisTAS™ was 24 percent. The low detection rate for the WrisTAS™ was largely due to those devices' erratic output or not recording during nearly 67 percent of all episodes. SCRAM™ correctly detected 57 percent across all BAC events, with another 22 percent (total 79%) detected, but as <.02 g/dL. SCRAM™ devices were more accurate earlier than later in the trials and may have had problems with water accumulation that reduced sensitivity. When subjects dosed themselves to BAC ≥ .08 g/dL, SCRAM™ correctly detected 88 percent of these events. The report summarizes comments from research subjects, offenders, and vendors who manage transdermal detection programs.			
17. Key Words Alcohol, Transdermal, Detection, Laboratory, Field		18. Distribution Statement Copy available from the NHTSA Web page: www.nhtsa.dot.gov	
19 Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21 No. of Pages	22. Price

Table of Contents

Technical Report Documentation Page.....	i
Executive Summary	1
Objective.....	4
Background	5
Controlling Convicted DWI Drivers	5
Literature Search	5
Alternative Methods for Monitoring Alcohol Consumption.....	6
Measurement of Alcohol in Sweat	7
Electrochemical Alcohol Sensors	7
Methods	11
Subject Screening	11
Basic Data Elements	12
Alcohol Consumption Procedures.....	12
Laboratory Dosing	12
Self Dosing	13
Device Acceptability	13
Real World Emulation	13
Coding Notes.....	14
Results	16
A Priori Considerations.....	16
Categories of Findings Reported	17
Drinking Data for Analysis.....	18
False Positives.....	20
Simple Detection: Sensitivity When $BAC \geq .02$ g/dL	20
WrisTAS™ Judged Performance Detecting $BAC \geq .02$ g/dL.....	21
SCRAM™ Judged Performance Detecting $BAC \geq .02$ g/dL.....	22
SCRAM™ Automated Alerts	22
Plotted Example Charts	24
Other Measures of Coded Hits and Misses: SCRAM™ and WrisTAS™.....	24
Magnitude of Peak Differences	24
Time Delays in BAC to TAC Peaks	26
Performance of SCRAM™ Before and After Check Valve Removal	28
Sex of Subject- SCRAM™	28
Sex of Subjects - WrisTAS™.....	29
Within Subjects Change in Accuracy	30
SCRAM™ within Subjects	30
WrisTAS™ within Subjects	30
SCRAM™ Accuracy over Time.....	31
Rule-Based Sensitivity Analysis	33
Setting up the Rules	33
Dealing with Spikes and Other Noise Parameters.....	35
Dealing with Shifting Baselines.....	35
Signal Detection Statistics: d' and A'	35
ROC (Receiver Operating Characteristic) Curves.....	36

Circumvention Detection.....	38
Environmental Challenges	39
Wearability and Acceptability	41
Focus Groups of SCRAM™ Offenders	41
Interview and Contacts with SCRAM™ Providers.....	42
Discussion.....	43
Giner WrisTAS™	43
AMS SCRAM™	45
Accuracy and Sensitivity	45
Communications	46
Other Considerations	47
Signal Detection Analysis	48
Circumvention Detection	48
Perceptions: Offenders and Vendors	49
Final Note and Recommendations.....	50
References	51

Appendices

Appendix A – Literature Search Sources	54
Appendix B – AMS Letter Regarding Check Valve.....	56
Appendix C – Participant Agreements and Instructions.....	58
Appendix D – Wearability of the SCRAM™ and WrisTAS™ Devices	68
Appendix E – SCRAM™ Focus Group and Users Feedback.....	70
Appendix F – One Subject’s Four-Week Chart.....	73
Appendix G – Coded Judgment Examples	79
Appendix H – SCRAM™ Vendor Feedback.....	85

Figures

Figure 1. The SCRAM™ Ankle Unit	8
Figure 2. The WrisTAS™ Wrist Unit.....	9
Figure 3a. BAC Maximum Recorded with Laboratory Dosing.....	19
Figure 3b. BAC Maximum Recorded with Self-Dosing.....	19
Figure 4. SCRAM™ Alerts (True Positive Detection) by BAC Cutoffs when Self-Dosed	23
Figure 5. Alerts (True Positive Detection) by BAC Cutoffs when Dosed in the Laboratory.....	24
Figure 6. Bland Altman Type Data Layout to Compare BAC and TAC for WrisTAS™ (6a left), and SCRAM™ (6b right) over a Range of Measured Values	25
Figure 7. Mean Percent Difference \pm 1 Standard Error of TAC from BAC by Sex and Type of SCRAM™ Device	29
Figure 8. Mean Percent Difference \pm 1 Standard Error of TAC from BAC by Sex for WrisTAS™ Device.....	30
Figure 9a. SCRAM™ TAC Percent Deviation from BAC by Days Wear Time \pm 1 Standard Error.....	32

Figure 9b. SCRAM™ Hit Rate (Y1) by Judgment or Alcohol Alerts and Accuracy (Y2) by Continuous Days of Wear	33
Figure 10. Signal Detection SCRAM™ at 4 BAC Values	37
Figure 11. Signal Detection WrisTAS™ at 4 BAC Values.....	37
Figure 12. AMS SCRAM™ Temperature Sensor Data Over 13 Days for Subject MS10a (1/29/05-2/10/05)	38
Figure 13. Skin Temperature on Arm and Leg of Subject While in the Cold Room.....	40

Tables

Table 1. BAC Characteristics Lab-Dosing and Self-Dosing Episodes.....	18
Table 2. WrisTAS™ Performance Detecting BAC \geq .02 g/dL	22
Table 3. SCRAM™ Performance Detecting BAC \geq .02 g/dL.....	22
Table 4. Agreement of Two Methods for Estimating SCRAM™ Hits \geq .02 g/dL	23
Table 5a. Characteristics of BAC Maxima and SCRAM™ TAC Maxima for True Positives (SCRAM™ HITS) by Type of Dosing.....	27
Table 5b. Characteristics of BAC Maxima and WrisTAS™ Maxima for True Positives (Wrist HITS) by Type of Dosing	27
Table 6. Algorithm-based Signal Detection Statistics for SCRAM™ and WrisTAS	36

Executive Summary

The objective of this research was to evaluate the accuracy and precision of two types of electrochemical transdermal alcohol sensors.

Introduction

There is only a very sparse research literature directly relevant to electrochemical transdermal alcohol detection. This includes papers published between 1992 and 2003 by project consultant Robert Swift, M.D., Ph.D., and his colleagues at Brown University reporting on the WrisTAS™ device; and a 2006 paper by Joseph Sakai, M.D., and his colleagues at the University of Colorado, who evaluated the SCRAM™ device. In addition, a biophysical model of transdermal alcohol measurement was published in 2006 by Anderson and Hlastala, biophysicists from University of Washington. These papers plus a few additional abstracts from recent scientific meeting presentations constitute the published literature on transdermal alcohol devices. In addition to these directly relevant papers, this report also provides some context for transdermal alcohol estimation by summarizing some background literature related to other nontraditional means of estimating alcohol exposure such as alcohol biomarkers, sweat patches, and noninvasive approaches to estimating blood alcohol concentration (BAC) such as, near infrared spectroscopy.

With respect to transdermal alcohol, Swift (2003) reported that approximately 1 percent of consumed alcohol is lost through the skin as a vapor. The concentration of alcohol at the skin surface reflects the concentration of alcohol in the blood (BAC), but curves plotted to represent the change at the skin surface show a delay of 2 or more hours on the ascending side and often somewhat more on the descending side relative to BAC.

Methods

Two devices, the Alcohol Monitoring Systems (AMS) Secure Continuous Remote Alcohol Monitor (SCRAM™), and the Giner Inc Wrist Transdermal Alcohol Sensor (WrisTAS™) were used in combined laboratory and field trials for 96 total weeks of wear (an average of 4.3 weeks per subject) by 22 subjects (15 males, 7 females). The SCRAM™ device locks onto the ankle and is worn 24/7 for the full duration of the study, including showering, and cannot be removed by the subjects without activating an alert condition. The WrisTAS™ device is a research prototype that affixes with a Velcro strap to the wrist and must be removed for showering. Neither device can be fully immersed.

In the laboratory, subjects were dosed in the morning based on weight and sex to a BAC calculated to reach .08 grams per deciliter (g/dL); in 60 such dosing trials, the average BAC attained was .083 g/dL. In the subjects' own field-initiated drinking, the mean BAC attained was .077 g/dL during 211 trials when the minimal BAC was $\geq .02$ g/dL. The 271 episodes with BAC $\geq .02$ g/dL formed the "signal" for detection analyses. All subjects had to provide daily drinking and eating logs, and each was given a handheld portable breath-test device to use for the study duration that enabled them to record BACs when drinking on their own.

Transdermal alcohol detection evaluation proceeded in three ways: (1) coded judgments of response magnitude based on visual inspection of the device data, (2) alerts issued by the AMS Scramnetwork server denoting that an alcohol-positive event had occurred (SCRAM™ only), and (3) through use of an automated algorithm that smoothed spikes from the data and accommodated to shifting baselines. Alerts issued by the Scramnetwork showed a 93.5 percent concordance with the judged

coding of human investigators for true positives, and a 91.5 percent concordance with false-negative judgments. This strong agreement represents a $\kappa=.85$ ($p=.000$). Although there was no comparable alert system for WrisTAS™, this degree of concordance between judged events and automated alerts for SCRAM™ serves to endorse the accuracy of the coded judgments against an external referent. The judged detection of alcohol by the transdermal devices relative to BAC forms the primary outcome data in this evaluation.

Results

The results demonstrated neither device has problems with false positives, but both had problems with false negatives and/or with unreadable data. The SCRAM™ false-negative rate due to complete response failure was 15 percent, and across all BACs, the SCRAM™ overall true-positive rate was 57 percent. The difference between the sum of those numbers and 100 percent represents some coding uncertainty explained in the report (another 22.5% was detected but as less than .02 g/dL BAC). Overall, the true-positive detection rate increased as the BAC increased from .02 to .08 g/dL. BAC episodes of .08 g/dL or greater that were attained during normal drinking were detected at a true-positive rate of .88 by SCRAM™. The WrisTAS™ sensor had a false-negative rate due to complete response failure of 8 percent (defined as on and working but not responding to ethanol), and an overall true-positive rate of 24 percent. The difference between the sum of those two numbers and 100 percent represent WrisTAS's™ missing or erratic data; 67 percent of the positive BAC episodes were either missing or unreadable from the WrisTAS™ data. This aspect of WrisTAS™ is the largest concern, and it has been suggested by the manufacturer that the problem is a consequence of a faulty chipset that controls data I/O functions. The WrisTAS™ device tested, version 5, has now been replaced with version 6. We have no evaluation data on version 6.

The SCRAM™ system's sensitivity and accuracy declined over the duration of wear; an aggregate near-perfect accuracy and high rates of sensitivity during the first period of wear declined as a function of time in service. This finding emerged as the largest concern with SCRAM™. The most likely cause of this problem is a consequence of water accumulation inside the sensor housing: as water accumulates the sensor's ability to detect ethanol is reduced. The SCRAM™ device that was tested has now been replaced by a device with less dead airspace for holding water, and this has reportedly solved the problem of water accumulation. We have no evaluation data on this newer version of SCRAM™.

Results showed that laboratory studies in which the calculated dose of alcohol was consumed in a 30-minute period yielded lower transdermal responses than when subjects dosed themselves (in normal self-initiated drinking). This was more of a problem with SCRAM™, which samples every 30 to 60 minutes, than with WrisTAS™, which samples continuously. In self-paced normal drinking, (self-dosed) subjects' consumption ordinarily proceeded for several hours and this manner of intake provided for a more sustained BAC signal detectable by SCRAM™ than was possible with a brief spike following rapid dosing.

Transdermal signals of female subjects were generally measured as lower than those of males relative to the BAC attained. This was the case for both types of transdermal sensors. Anderson and Hlastala (2006) have shown that the thickness of the stratum corneum, the outermost layer of the epidermis, and the hydration state of the subject are factors in the movement of ethanol across dermal barriers to the skin surface. The proportional body water content differs between the sexes, and this may partially explain this finding.

Discussion

In evaluation of circumvention protection, the SCRAM™ system performed well. It may be possible for a highly motivated offender who is familiar with the SCRAM™ design to devise a procedure to temporarily block alcohol without blocking the infrared sensor that detects obstructions or the temperature sensor that monitors temperature near the skin surface. However, it seems unlikely that circumvention by obstruction can constitute a real threat to the integrity of this system while drinking because it would require constant vigilance by the offender. The communication protocols built into SCRAM™ that combine daily automated upload of data and the issuance of daily alerts to a program monitor will likely prevent most offenders from beating this system. The Scramnetwork server works well and proved to be a sophisticated and stable authorization and data-tracking system.

User comments allude to some discomfort, especially among females, and one female research subject showed evidence of bruising after a week of wear. Court-ordered users (including women) who were part of a focus group found SCRAM™ to be occasionally annoying but acceptable, and a tolerable alternative to jail time. Two commented that it helped goad them toward sobriety in a way that other motivators were unable to do. Vendors and others who manage SCRAM™ programs were generally positive about their experiences with it. Alcohol Monitoring Systems (AMS) staff commented that about 20 percent of the offenders seemed unable to control their drinking and had to be removed from the SCRAM™ program. However, we can provide no external corroboration of this estimate.

Overall, these devices performed more poorly than we expected with respect to sensitivity and accuracy; however, with independent evaluations, the manufacturers can improve their products. The attainable accuracy, however, may only be an approximation of BAC due to subject-specific factors that influence ethanol gas concentration at the skin surface. There is no doubt that the transdermal concept is valid as long as expectations of quantitative parity with BAC are moderated.

There is a parallel in these early findings about the accuracy of transdermal devices that is reminiscent of the early accuracy of alcohol ignition interlock devices. First generation interlock devices were often criticized for failing to match the performance characteristics of more conventional breath-test devices, despite interlocks having to operate in an often hostile automotive environment of heat, cold, dust, and vibration. Similarly, TAC is not BAC, and the expectation of parity is an impractical expectation to place on this nascent technology. Both interlocks and transdermal sensing need to be judged first on their potential contributions to public safety. Moreover, just as interlock devices have improved in the 20 years since their first adoption, it is reasonable to expect that the transdermal-sensing equipment will also improve. These devices warrant further development and further study.

Objective

The objective of this study was to obtain laboratory data on the precision and accuracy of commercially available transdermal alcohol-detecting and monitoring devices.

Before initiating this research study, we conducted a preliminary search of the literature. This resulted in information on two devices that, based on the research literature or their use in correctional settings, were sufficiently far along in their development to warrant evaluation. These two devices are from Alcohol Monitoring Systems Inc.--known as Secure Continuous Remote Alcohol Monitoring (SCRAM™) – and from Giner Inc.--known as the WrisTAS™, which denotes Wrist Transdermal Alcohol Sensor.

To more completely evaluate performance of these devices, the research design was extended to a study of their performance with self-initiated drinking in the field. This comparison of performance in the laboratory and under field conditions constitutes the primary evaluation in this study. The addition of a field element to the evaluation was considered important because field self-dosing with alcohol overcomes some of the contrivances of laboratory study and more closely emulates regular drinking in the real world. This emulation extends to both the pattern of drinking and the level of drinking. Research ethics necessarily imposes dosing limitations that are usually more conservative than self-dosing. In the laboratory, we calculated a dose of alcohol for each subject that would bring the BAC up to .08 g/dL (80 mg/dL) within 30 to 60 minutes of initiating drinking. For field evaluation of BAC levels, subjects were entrusted with and trained to use handheld evidential breath-test devices and requested to maintain a drinking log.

Beyond evaluating the primary variables, we also evaluated the effects of the participant sex, duration of wear, the effect of cold exposure, and exercise-induced sweating. Attempts to circumvent the devices (to consume alcohol without detection) were conducted, as were evaluations of wear acceptability and appraisal of the programs both from the perspectives of users and program administrators.

This report provides background on the nature of transdermal alcohol sensors. It describes the experimental methods used to conduct the evaluation, reports on the findings of the evaluation, discusses those findings, and offers conclusions and recommendations based upon the research.

Background

Controlling Convicted DWI Drivers

A major objective of driving while intoxicated (DWI) offender sanctioning programs is to protect the public by preventing future impaired driving. Logically, there are three ways to accomplish this objective: (1) prevent all driving, (2) prevent all driving after drinking, or (3) prevent all drinking. The most widely used approach has been an attempt to prevent all driving by DWI offenders by suspending their licenses. Many studies have demonstrated that this effectively reduces DWI recidivism in addition to serving as a general deterrent to drinking and driving. However, the large growth in vehicle travel in recent years has contributed to the decline in enforcement of the laws against driving while suspended (DWS), so the effectiveness of this approach has been compromised. A study sponsored by the National Highway Traffic Safety Administration (NHTSA) suggests that anywhere from 36 to 88 percent of suspended DWI offenders continue to drive (McCartt, Geary, & Nissen, 2002). One response to this has been the vehicle or license plate impoundment sanctions that deprive offenders' use of their cars. Such laws also have been effective but are limited as many offenders drive vehicles registered to others.

Alcohol ignition interlocks on vehicles offer another straightforward approach to protecting the public by preventing impaired driving. Interlocks allow offenders to continue to drive while sober for employment and family support purposes. However, it has proven difficult to motivate offenders to install interlocks on their vehicles, and the use of a non-interlock-equipped vehicle remains a threat to the effectiveness of interlock programs (Marques et al., 2001). Although the growth of interlock programs as a proportion of all DWI arrestees in the United States is promising, there will always be some segment of the population who will not comply with the requirements of those programs.

The third approach, preventing drinking, offers a strategy that not only would protect the public from alcohol-related crashes, but also would reduce other alcohol-related problems such as violence, non traffic injuries, and alcohol addiction. In the past, courts have attempted to control drinking by requiring the administration of Antabuse™ (disulfiram) or by intensive supervision probation programs involving random, surprise breath tests for alcohol use. For the most part, those efforts have not been adequately evaluated. Somewhat more efficient monitoring of alcohol use is provided by house arrest and interlock programs. Typically, house arrest BAC¹ test units are stationary devices that use video images or voice recognition to identify the user and transmit the breath-test result over a telephone line. They do not, however, provide such information when the offender is at work or away from home. Similarly, interlock data loggers provide information on drinking but only when an attempt is made to start the offender's vehicle, and the vehicle operator is not guaranteed to be the offender of interest.

Literature Search

The advent of continuous, passive methods for monitoring alcohol use provides an opportunity to determine whether such methodology holds promise for controlling the impaired driving of DWI offenders with a minimum effect on their ability to maintain a normal lifestyle and support their

¹ Throughout this report, BAC, whether represented as g/dL or mg/dL, indicates BrAC (breath alcohol concentration measurement or grams of alcohol /210L breath).

families. Accordingly, we conducted a thorough search of the literature in an effort to identify transdermal alcohol detection literature and products that were otherwise unknown to the investigators. A comprehensive scan was undertaken using computer searchable databases, including Medline, Dialog (with hundreds of databases), ERIC, Lexis-Nexis, and other databases/catalogs.

To accomplish this objective, we used various combinations of key words such as transdermal alcohol/ethanol and measuring, monitoring, skin, sweat, and vapor. We examined Dialog databases in the areas of law enforcement, social science, psychology, medicine, transportation, and safety (see Appendix A for a list of screened databases). In searching the literature, we were most attentive to descriptions of the precision and accuracy of such measurements and the physiological and environmental conditions that affect transdermal alcohol measurement. Although there are a few papers on the transdermal alcohol measuring devices and more on environmental conditions that affect alcohol measurement, it is rare to find papers on factors affecting transdermal alcohol measurement. There is not yet a mainstream of research on transdermal alcohol measurement in the physiology, pharmacology, medical, or safety literature other than as represented by the citations reviewed here.

Alternative Methods for Monitoring Alcohol Consumption

Alcohol's presence in the body can be determined through the use of bodily specimens, most typically blood or breath, but also urine and oral fluid from saliva. Most ethanol is transformed to acetaldehyde and acetate through liver enzymes in a well-known detoxification process that has a predictable time course and is described by the Widmark equation (Widmark, 1932). This is not the only pathway by which alcohol is biotransformed into more inert substances, but it is the most widely known.

The presence and measurement of alcohol markers (such as carbohydrate deficient transferrin, gamma glutamyl transferase, ethyl glucuronide, or fatty acid ethyl esters) is possible due to minor metabolic pathways that are consequential to ethanol exposure or play a role in degrading the ethanol. These alternate pathways yield measurable products that persist for many days or weeks after the ethanol per se has been changed or lost. The surveillance window, the period following exposure to a drug when traces can be detected, is substantially extended for alcohol by examination of these ethanol biomarker or metabolites. Gilg, Buchholtz, and Huth (2000); Gjerde and Morland (1987); Gjerde, Sakshaug, and Morland (1986); and Bjerre (2003) have all reported the extent to which these substances can aid in the profiling of high-risk alcohol-involved drivers.

Those measurement options aside, the forensic measurement of alcohol has most often depended upon direct measurement of ethanol before it is transformed. At present, there are still very few forensic applications of the alcohol markers that persist beyond the brief hours it takes to transform and degrade the ethanol consumed in a drinking session. The most widely used ethanol measurements in blood or breath are closely similar due to the equilibrium between capillary ethanol concentration and ethanol vapor on opposite sides of the membranes in the lung alveoli. Breath ethanol measurement uses a conservative partition coefficient to relate the BAC and BrAC (breath alcohol concentration) as sources of information. The efficient and predictable disposal of ethanol by the dehydrogenase enzymes in the liver requires that BAC or BrAC ordinarily be sampled several times per day to reconstruct an alcohol consumption curve. During the time transformation and deactivation of ethanol is proceeding in the bloodstream, a small proportion (less than 1%) of the

ethanol is lost through the skin as vapor through two processes: passive diffusion through the skin and active excretion from the eccrine or sweat glands (Swift, 2003).

Because alcohol eventually becomes distributed freely in all body compartments, it can be measured in any body fluid, including compartments as remote as the vitreous in the eyes and cerebrospinal fluid in the brain and as large as the interstitial fluid that bathes cells throughout the body.

A device under development by Spectrx, Inc., samples alcohol in interstitial fluid under the skin by using a laser to shoot miniature holes in the skin and then samples the concentration of alcohol in the fluid of this extra-cellular interstitial compartment. This company has been aided by Small Business Innovative Research (SBIR) funds from the National Institutes of Health (NIH) and may someday have an alcohol-sensing product ready for commercial application. Spectrx's interest in alcohol as a target substance follows from its primary product development research in glucose monitoring and determination of therapeutic medication levels. All its products are and will be subject to Federal Drug Administration (FDA) approval because the laser approach to sampling crosses the dermal barrier. Another company of interest is TruTouch Inc., staffed by former defense physicists who are skilled in the uses of near-infrared spectroscopy. They have developed a device that can passively evaluate alcohol concentration in the interstitium without breaking the skin surface (Brown & Ridder, 2005; Ridder, Brown, & Ver Steeg, 2005; Ridder, Hendee, & Brown, 2005).

Measurement of Alcohol in Sweat

Among the most accessible of the nontraditional sampling media for alcohol concentration determination is sweat. Several methods have been used to estimate alcohol consumption by detection and attempts at measurement of ethanol in sweat, which will hereafter be referred to generically as transdermal alcohol concentration (TAC). These include the sweat patch that accumulates alcohol over several days (Phillips, Greenberg, & Andrzejewski, 1995), a little known ethanol band-aid that uses colorimetric technology (Roizen, Lichtor, & Lane, 1990), and the electrochemical methods that convert alcohol into an electrical signal proportional to concentration. The two electrochemical devices that exist today—and the two evaluated in this report—are based on different technologies. At their current stages of development, they are suited to different problem populations.

Menssana Research, Inc., is the company that produced the alcohol sweat patch and its chief executive officer, Michael Phillips, M.D., holds the patent. Menssana has developed several products that have grown from an interest in the detection of volatile organic compounds in breath and sweat. This company's interest in alcohol appears to have been a tangential development, as its primary business is in making detectors of minute, picomolar quantities of exhaled or expired volatiles that are often indicative of disease states. The company's business model has historic lines dating back to the Hippocratic era when diagnosticians were trained to use their noses to help in the detection of disease states. Breath volatiles may indicate cancer, diabetes, transplant rejection, and other diseases. Despite an active research program in alcohol detection that lasted 15 years (1980-1995), published literature about the Alcopatch, with Phillips usually the lead author, ended in 1995. The original work on the Alcopatch was supported by NIH development funds, and once the SBIR support ended, so too did further development.

Electrochemical Alcohol Sensors

The main transdermal sensing systems available today are shown in Figures 1 and 2. The AMS device, the SCRAM™ (see Figure 1), introduced into the forensic market a few years ago, puffs air at

specific intervals to volatilize ethanol and then samples the ethanol vapor in the space between the skin and the fuel-cell sensor. Sampling intervals are variable, but in most applications, the device draws air into the fuel cell hourly. If alcohol is detected, the sampling frequency increases to every 30 minutes. The Giner device, the WristAS™ (see Figure 2), developed during the 1990s, samples alcohol vapor almost continuously and then records averaged data once every 30 seconds to 10 minutes. The data are later downloaded via a serial port on the device to a computer. In most applications, 5-minute averages of data are stored for later download.

SCRAM™

The SCRAM™ consists of three components: (1) a small SCRAM™ bracelet worn on the ankle, (2) a SCRAM™ modem, and (3) a remote server for aggregating data from offenders and for reporting these data to the monitoring staff. The unit is locked on and worn 24 hours a day, 7 days a week, for up to 6 months based on current practice. The AMS device uses an automated daily uplink feature for data transfer to the remote server via the modem; this product was designed for security and remote reporting to minimize circumvention and to render data usable by courts or corrections. In most applications, the SCRAM™



Figure 1. The SCRAM™ Ankle Unit

modem is scheduled to transfer data from the ankle bracelet during normal sleeping hours. The AMS ankle bracelet weighs about 8 ounces (Figure 1) and is approximately the size of two demitasse coffee cups linked by a data cable embedded in strong straps. One side houses the sensors, and the other side, the digital signal processing hardware. The sensor side of SCRAM™ includes an air pump that actively draws in ethanol vapor from the skin surface into the fuel cell. In addition to alcohol, sensors detect changes in the temperature and infrared signals near the skin. The two non-alcohol sensors are important parts of the circumvention detection protocols. In addition to the sensors that aid in detecting circumvention, if the strap is cut or the offender ceases to accommodate the requirements of the program, a probation officer or program administrator would be alerted within 24 hours that a break in continuity had occurred or that the modem had failed to uplink. When properly installed on an ankle, a lock-in retainer makes it very difficult for an offender to remove the SCRAM™ without cutting the strap. The device cannot be immersed but it can be worn while showering and while maintaining personal hygiene inside the rubber muffs.

AMS Inc. is a technology company privately held by the principals who also hold the senior management positions. Many of the senior staff at AMS had previously worked in the computer industry and have expertise in data management. AMS has funded its own unpublished developmental research on the SCRAM™ device (Zettl, 2002). In the *Journal of Offender Monitoring*, Steven Bock (2003) from Michigan's Electronic Monitoring Center reported on characteristics of the SCRAM™ device on 19 offenders under study. The offenders were reportedly satisfied with the device, but Bock reported AMS modified the strap following efforts by some in this group to tamper with the strap.

In 2003, AMS funded a study by Drs. Thomas Crowley and Joseph Sakai at the University of Colorado Health Sciences Center. The Colorado study design had two parts. Part 1 included a one-day laboratory analysis in which subjects arrived, were hooked up, drank, and had the bracelet removed. Part 2 was a 7-day wear study in which subjects (alcohol and nonalcohol dependent) logged drinking while wearing their SCRAM™ bracelets. Published results (Sakai, Mikulich-

Gilbertson, Long, & Crowley, 2006) reported no episodes of false positives based on either lab dosing or reported consumption. BAC and TAC findings could not be considered quantitatively equivalent, but there was qualitative parity between reported drinking and SCRAM™ results. The devices were reported to be reasonably comfortable, and they capably discriminated between alcohol-dependent and social drinkers.

Because the SCRAM™ device is used with a court-ordered or a corrections population, it is important to anticipate and have a means to detect efforts to circumvent the alcohol sensing and reporting protocols built into the device. AMS has evaluated the characteristics of several types of materials and reports that all known materials capable of blocking alcohol can be detected due to the infrared sensor. The AMS principals note that data interpretation protocols can distinguish a wide range of attempts to block the air sampling with plastic, Mylar, and other interferents (i.e., sources of interference). The algorithms to detect circumvention are not capable of full automation, so the software flags suspicious results and this sends an alert to an expert human monitor who must view these questionable cases to distinguish between true positives, various classes of circumvention, and true negatives. The circumvention detection logic is dependent upon the signals from all three types of sensors in the device. Basic circumvention attempts are easily detected when the skin temperature or infrared sensor readings suggest the proximity of the sensor to the skin has changed.

WrisTAS™

The Giner device, WrisTAS™, which was developed primarily for use in medical settings with more compliant subjects, lacks the protocols for detecting tampering and has not yet been adapted for court use.

Nonetheless, the device does have a skin resistance/conductance sensor and a temperature sensor. These sensors, when operative, can aid in determining if a person removed or blocked the device. When in service, data from the device are periodically downloaded to a computer via a serial port interface. A new version (WrisTAS™ version 6) can send radio frequency data without having to conduct a manual download to a computer. The Giner device has had research trials primarily aimed at alcohol treatment applications for monitoring patient compliance and for early detection of lapses in an abstinence program. The Giner device (Figure 2), shaped and worn like a wristwatch, is much smaller than SCRAM™, and its location is very convenient. Were it to be used with an offender population, it would need protocols to prevent unauthorized removal because, currently, it is secured to the wrist with a Velcro strap. It cannot get wet and must be removed for showering.



Figure 2. The WrisTAS™ Wrist Unit

The Giner device has been reported to be linear within physiologic (normal pharmacologic ranges) of ethanol dosing. There has been somewhat more scientific investigation of the Giner than of the AMS device due to developmental support from the National Institute on Alcohol Abuse and Alcoholism (NIAAA). Swift, Martin, Swette, LaConti, and Kackley (1992) and Swift (2000) reported that the WrisTAS™ linearity extends from 5 to 500 mg/dL (.005 to .50 g/dL). Transdermal devices output a TAC that reportedly parallels the more familiar BAC curves but is shifted to the right with a 1- to 2-hour delay. Swift (2003) reports that the area under the curve for TAC and BAC of Giner devices correlate with $r=.8$. It should be noted that Dr. Swift has served as a consultant to this NHTSA evaluation. He informs us that published evaluations of the WrisTAS™ are based on devices that have been specifically selected for a high-fidelity response to allow a proper appreciation of the potential of the sensor. However, many WrisTAS™ devices have a relatively high failure rate, resulting in data loss. Loss of data may be caused by problems in voltage regulation or chipset

failure. A conference report in June of 2005 by Greenfield, Tujague, Bond, and Kerr (2005) found that version 5 WrisTAS™ devices when not pre-selected for a high-fidelity response could be considered high quality only on about 20 percent of the evaluation days. Newer versions of the WrisTAS™ devices that were not available for this evaluation project are reportedly more reliable.

The underlying technology that allows the WrisTAS™ to rapidly sample ethanol vapor is different from conventional fuel cell sensors that require some delay after sampling while the alcohol fuel is oxidized. Where a traditional fuel cell could not easily be cleared quickly after sampling, the WrisTAS™ sensor is not affected by rapid sampling because it measures a continuous oxidation current. The SCRAM™ and WrisTAS™ devices measure transdermal ethanol by different approaches to the problem, and each have stronger and weaker points, but both are actively being modified in an attempt to improve the technology.

Both of these transdermal devices are high technology in the best sense. The SCRAM™ device uses conventional fuel-cell sensors but marries them to sophisticated computer algorithms and data transfer technologies in a secure device that is already used in forensic applications. The WrisTAS™ device uses a constant hydrated platinum electrode maintained at a controlled potential and bathed in aqueous electrolyte held in a reservoir. In the WrisTAS™, an electrode oxidizes the ethanol and forms acetic acid that diffuses into the reservoir. The current is converted to a digital signal that is averaged and stored at preset time intervals from 30 seconds to 10 minutes. The data capacity is 3 weeks at 10-minute intervals. Data are downloaded to a computer serial port. The reservoir must be replenished about once a month with a few drops of de-ionized water.

Based on this literature search described above, we have determined that the two transdermal devices identified are the only ones that are available for study, though only one of these, the SCRAM™, is being leased commercially. Nonetheless, we included both the SCRAM™ and WrisTAS™ devices as part of the precision and accuracy evaluation as they represent two different approaches to electrochemical alcohol detection at the skin surface.

As part of this evaluation, we made an effort to determine the ease with which a user could circumvent the device without detection. We evaluated the effects of chilled air (and therefore reduced skin temperatures) and mild exercise-induced sweating on the precision and accuracy of transdermal devices relative to known BrAC levels from a fuel-cell based handheld breath tester.

Methods

Subject Screening

Subjects between the ages of 21 and 35 were invited to give signed consent and then participate in a screening process for subject selection. The consent procedure was reviewed by the PIRE East IRB, a human subjects review body that operates under Federalwide Assurance Number FWA # 00007038. Subjects were selected based on the following criteria: (1) they must report drinking regularly but must not drink to extreme excess and must not have drug- or alcohol-related health or criminal problems, and (2) they should not use contraindicated medications or be pregnant. Subjects were asked to provide a urine sample upon entry into the subject pool to confirm that they showed no positives for benzodiazepines (e.g., Valium/Xanax/Ativan type drugs) or the NIDA 5 panel drugs (opiates, amphetamines, cocaine, phencyclidine, or marijuana). The urine toxicology screening devices were Biosites Triage Tox Screens, and the results were machine read from the Biosites Triage Meter Plus. Pregnancy tests were performed with over-the-counter, e.p.t.™, early pregnancy tests for home use.

In addition, a prospective subject's normal level of alcohol consumption was assessed by requiring that he or she fill in the AUDIT (Alcohol Use Disorders Inventory), and respond to a subset of questions from the AUDADIS (Alcohol Use Disorders and Associated Disabilities Schedule). The former assessment is a high-sensitivity screen for potential problem-level drinking, whereas the latter is more attuned to diagnosable DSM-IV criteria for alcohol abuse or alcohol dependence.

Of the 55 contacts (mean age 26.2, 56% male) made with prospective subjects during the 7 months of active testing, 32 contacts were eventually screened for consideration according to the described criteria. Of those screened, 10 were excluded. Two people tested positive for marijuana, one tested positive for benzodiazepines (prescription anti-anxiety medicine), and eight reported risk levels of alcohol consumption.

There were two phases of transdermal device evaluation, a 4-week wear phase and a 2-week wear phase. The two week trials were added as a supplement following the basic 4-week phase in order to answer an emergent question about SCRAM™ related to a check valve (described later). Eight of the original 18 subjects from the 4-week phase (studied in 3 waves of 6 subjects) agreed to serve again as subjects in the 2-week phase to address the emergent questions. For that purpose, this group was supplemented by four new subjects. The repeat assessment was desirable for two reasons: (1) it allowed for a within-subjects pre-post evaluation of the SCRAM™ device after the check valve design change (see Appendix B), and (2) using the same subjects reduced the amount of preliminary explanation and training required for participants. There was nothing about repeat participation that could adversely affect device sensitivity or any of the dependent measures collected, and the repeat subjects attained comparable BACs and TACs as the four new subjects with similar SCRAM™ devices. All together, the final subject pool participating in the research included 15 males (5 repeated for phase 2) and 7 females (3 repeated for phase 2) for a total of 30 trials (mean age 26.7, 68% male) covering 96 weeks of wear among 22 subjects. A trial in this case is defined as the duration that a subject was attached continuously to the transdermal device.

Participant Consent Forms, Participant Agreement, and Instructions for Participants are found in Appendix C. Subjects were paid \$100 a week for participation, plus a bonus paid at the end of the study if they stayed for the full duration of 4 weeks (\$400) or 2 weeks (\$200). The bonus was deemed

warranted due to the disruptive effects of subjects dropping out of a study with a tight timeframe. No one dropped out.

Basic Data Elements

This study required subjects to log all alcohol drinks or drink equivalents during the study period and to periodically self-test their BACs with a handheld portable breath tester (PBT), a CMI Intoxylizer SD-400. Whenever a subject consumed more than one drink, he or she was instructed to follow standard procedure of either performing a mouth rinse with water and/or waiting 15 minutes after the last beverage sip before testing with the PBT. This was accomplished outside the office by entrusting all subjects with a breath-test device that they were to bring with them whenever they might be drinking away from home. Accordingly, with two transdermal sensors (the WrisTAS™ and the SCRAM™), a fuel-cell PBT, and a log of drinking, we have four data elements that represent ethanol consumption during the study.

Alcohol Consumption Procedures

There were two types of alcohol consumption in this study: (1) laboratory dosing in which subjects came to the PIRE offices and drank in the morning on an empty stomach, and (2) self-dosing or free-form alcohol consumption of their own choosing. Each episode of alcohol consumption is coded and logged for each subject. Lab-dosed and self-dosed drinking events are tracked separately since the method of dosing is quite different.

Laboratory Dosing

In the lab, subjects were given an amount of distilled spirits calculated from Widmark equations to bring their BACs to 80 mg/dL (.08 g/dL) when consumed over a 30-minute period. The theoretical dose for males was .64 gm/kg and for females .56 gm/kg. A spreadsheet was programmed for this calculation. The program recorded information about subject sex, bodyweight, percentage of alcohol of preferred beverage, and duration of the drinking (to account for concurrent metabolism). The program returned a dose suitable for achieving a target BAC of .08 g/dL after 30 minutes. In practice, subjects selected vodka, rum or gin; all were 80 proof (40% alcohol). Subjects were permitted to add soft drinks or juice to their drinks. Subjects were instructed to space their consumption over the full 30-minute period. We waited an additional 15 minutes after drinks were finished before the start of breath testing. BAC readings in the laboratory phase represent the average of two PBT fuel-cell testers used successively within the same 1-minute period. The agreement between the two readings was very high. Whenever subjects were dosed in the laboratory condition, a nurse or EMT was present in the event of any adverse reaction, such as aspiration of vomit, severe flushing reaction, or other reactions (there were none). BAC levels in lab dosing were expected to rise to a peak within 30 to 60 minutes of beginning consumption. Subjects were instructed to arrive for laboratory dosing with no measurable BAC, and to limit drinking the night before so that TAC would have a chance to return to zero before lab dosing. They were also asked to avoid drinking for several hours after laboratory dosing so that TAC could return to .00 g/dL before any further self-dosing. There were a few occasions when laboratory and self-dosing TAC levels overlapped. There were no cases where laboratory dosing began with measurable BACs. On two occasions, subjects arrived for dosing with measurable BACs from the night before. In these cases, dosing was rescheduled for a later time. There were 2 lab-dosed drinking events for each of the 30 subject trials for a total of 60 dosing events.

Self-Dosing

When self-dosing, subjects were instructed to keep a log of all food and alcohol consumed on any drinking episode when they had two or more alcohol drinks, and to record breath-test readings after consuming alcohol. Subjects were encouraged to include whatever they felt noteworthy, but at a minimum, they were told the logs should contain the following:

- The time at which a meal was consumed and whether the meal was considered to be a small snack or a small, medium, or large meal.
- The time at which alcohol was consumed and the number of drinks consumed, where a drink equals one 12-ounce beer, one 4-ounce glass of wine, or a mixed drink containing one shot of liquor. Subjects were required to note when drinks consumed were larger or smaller than standard size (e.g., large beers or double shots of liquor) or stronger than usual (e.g., shots of 150-proof rum).
- The time at which breath tests were taken and the reading. Subjects were instructed to take the first breath test 30 minutes after the first drink and once an hour afterward until their BACs reached .00 g/dL or until the subject went to bed for the night.

Subjects were instructed to e-mail drink logs to project staff once a day. On drinking days, the logs would contain eating, drinking, and breath-test data. On non-drinking days, subjects were to send an e-mail stating that they had not had alcohol that day. In practice, subjects sometimes sent logs less frequently than once a day, but drinking logs were ultimately submitted for all but five of 249 self-dosed drinking events. The information about meal size was collected to help resolve any discrepancies between TAC and BAC levels.

Device Acceptability

In both the 4-week and the 2-week study phases, the procedures were identical. The longer duration study period satisfactorily answered the question about acceptability and wearability. For the most part, subjects easily complied with the requirements of the 4 weeks of wear, despite some initial discomfort. Following removal of the device, subjects were debriefed on discomfort issues and ways in which the devices interfered with everyday living. Wear-related issues are discussed in Results. Some subjects were more adversely affected than others, but no subjects dropped from the trial. Information about device acceptability for research purposes can be found in Appendices D and E that review focus group discussions with offenders.

In addition to comfort/discomfort issues, another important aspect of the extended wear evaluation was accuracy and precision change over time, an aspect that has been reported upon previously. This topic is also reviewed in Results.

Real World Emulation

Because research subjects must be given a monetary incentive to participate and real world offenders are under some threat by the courts or corrections to participate, the compliance motivation of these two groups is different. More germane perhaps, research subjects were selected because they drink and are expected to drink during the evaluation, whereas offenders are explicitly told not to drink. This difference means that accuracy and precision findings in this study need to be qualified before

attempting to generalize results as to how the devices would normally perform with subjects who were expected to remain abstinent.

Coding Notes

Having defined drinking episodes as having a peak BAC $\geq .02$ g/dL, the next task was to determine if the transdermal devices could detect or follow a BAC elevation by registering a positive TAC value. For each subject, all BAC and TAC data were compiled into a spreadsheet, and a composite graph was created that included self-administered dosing, lab-dosing, SCRAM™ TACs, and WrisTAS™ TACs. Converting this information into discrete categories is a judgment task. Two coders independently reviewed 41 episodes and achieved a high degree of agreement (>96%) on the definition of episodes, maximal BAC, maximal SCRAM™ TACs, maximal WrisTAS™ TACs, and time of these maxima. Because there were no fundamental disagreements between the two coders, the subsequent judgments of a single coder were used. Using the spreadsheet, the following data were recorded for each drinking episode:

- The time and level of the maximum BAC recorded.
- The time of and maximum SCRAM™ TAC associated with the drinking episode following the maximal BAC occurrence.
- The time of and maximum WrisTAS™ TAC associated with the drinking episode following the maximal BAC occurrence.
- The placement into categories of those TAC responses that were not clear true-positive hits but not easily coded as simple false negatives. The definitions for these “gray” categories are described below.

When there were two or more maximum TAC or BAC readings of the same value for a drinking event, the earliest was used, but if a later reading exceeded an earlier one by as little as .001 g/dL, the later higher reading was taken as the maximum. The data were recorded for all drinking events noted by the subject, whether or not the maximum BAC reached .02 g/dL. In a few cases, a new drinking episode occurred so soon after a prior drinking episode that it was unclear whether a TAC was a late response to the first episode or a response to the second. When this occurred, the time of the first drink of the second episode was set as the end time of the first episode, and all other data coded accordingly. Occasionally, spikes in TAC readings would result in a reading significantly higher than surrounding readings. These were deemed to be “outliers,” and were not recorded as maximum TAC readings for an event. In a later section, we describe automated smoothing algorithms used to exclude outlier spikes from the data series.

For each drinking episode and for each transdermal device, the coder categorized the TAC readings as one of the following:

True Positives

- *A hit* - The drinking episode was clearly visible with a TAC of .02 g/dL or more.

Subtypes of False Negatives

- *< .02 g/dL* - The drinking episode was evident but the TAC response was < .02 g/dL.
- *Low-confidence* - A change in TAC was apparent, though it would have been difficult to identify it as reflecting a drinking episode without knowing that the BAC was

elevated, mostly due to high variance in TAC before and after drinking episodes.

- *Too noisy* - The TAC readings were patternless, too noisy, or too variable to clearly distinguish a drinking episode.
- *Missing data* - There were no TAC data produced or retrieved during or after a drinking episode that could allow an estimation of sensor accuracy.
- *False negative with response failure* - The transdermal device was on and recording, but based on TAC reading, it seems to have completely missed a positive BAC ≥ 0.02 g/dL.

Results

A Priori Considerations

The objective of this evaluation was primarily to acquire information on the accuracy and precision of two types of transdermal ethanol sensors. Accordingly, it is useful to review the concept of accuracy and precision.

Accuracy, as used here, is an estimate of the extent to which a transdermal sensor evaluates a sample of ethanol vapor expired at the skin surface as equivalent to a known concentration of ethanol dissolved in body fluids and in general circulation. The standard (BAC) by which we evaluate the accuracy of transdermal sensors can be precisely known by estimating it from BAC. Although ethanol must move through several body water compartments to appear at the skin surface, the peak transdermal ethanol level (voltage or current signals that can be converted to equivalent units of ethanol concentration) can be set to represent the known BAC standards. The estimated amount of alcohol detected at the skin surface is referenced to BAC values typically associated with such levels of transdermal alcohol. Accuracy estimations are based on how the transdermal sensors' voltage calibration (SCRAM™) or oxidation current (WrisTAS™) can be stably referenced to a level of alcohol in the blood. When the mean of the transdermal signal is very close to the mean of the BAC, then the device will be considered highly accurate.

The precision of transdermal sensors is estimated by the ability of the sensors to regularly detect the same levels of ethanol when confronted with the same degree of ethanol signal. Precision has to do with the replicability of a finding and is best estimated by the standard deviation of results if the mean ethanol signal it is attempting to estimate is constant. When the standard deviation is small, the device is precise.

The precision criterion for transdermal sensors may be more difficult to estimate than accuracy. The time course of a BAC clearance curve is well known and described by the Widmark formula; however, the ethanol concentration estimated by transdermal sensors is just a small fraction (~1%) of the ethanol cleared from circulation. Passage through body water compartments and skin surfaces introduces complex kinetics into the problem of estimating the reduction in ethanol at the skin surface over time after a peak BAC level begins to decline. Accordingly, the BAC level at any point in time, other than the peak level, may not be a useful standard against which to estimate the precision of a transdermal device. If the devices are calibrated at the peak, their readings may diverge from BAC levels on both ascending and descending phases. If BAC and TAC were always the same, there would be less of a problem.

The two devices under evaluation differ on several dimensions. Perhaps the most significant difference is that the AMS SCRAM™ is available for the criminal justice marketplace and is in service now, whereas WrisTAS™ is a research prototype and Giner has no device on the market. Accordingly, these are not competitive products, and although both are electrochemical devices, they are not based on the same underlying technology. SCRAM™ is based on alcohol fuel-cell technology, whereas WrisTAS™ is based on a hydrated proton exchange membrane. Each has certain strengths and weaknesses. This evaluation is not an effort to find a winner but rather to characterize their precision and accuracy under specifiable conditions.

The AMS SCRAM™ device, available for use by courts or corrections, restricts the estimated BAC report to .08 g/dL, even when the actual estimated BAC exceeds that value. The flat peaks at .08

g/dL, seen on AMS SCRAM™ specimen charts, is a convention the company has adopted to reduce unnecessary detail because any TAC of .08 g/dL is well above the “actionable” level. However, for this evaluation, AMS has made available the underlying raw data that is not usually available to its customers. The ability to track the full range of BAC information, including those extending higher than .08 g/dL, adds detail that has allowed PIRE to more fully characterize the behavior of the SCRAM™ sensors.

Although the WrisTAS™ device has a native output that writes a file in mg/dL, for reader convenience to maintain compatibility with forensic conventions of BAC, we converted all those values to g/dL.

In addition, the y-axis on the output charts of Giner WrisTAS’s own data retrieval and display utility shows a zero point that may not always represent zero TAC. In most of the analyses reported here, we regard the zeros returned by the Giner data utility as true zero TACs and regard all positive values as positive TACs. In practice, however, the Giner head technician advised that he regards the true zero TAC on the WrisTAS™ output to be more accurately defined by the baseline from which the sensor output departs when it detects alcohol and to which it returns when a TAC signal falls. Stated differently, the WrisTAS™ baseline drifts, and although they are writing algorithms to correct for that drift, the version 5 that was made available for study does not have a native correction for baseline drift. Accordingly, in the sensitivity analyses later in this document and the Receiver Operating Characteristic (ROC) curves that are based on algorithmic reading of the data, we have included adjustments to the WrisTAS™ data so a functional zero baseline is inferred from a series of points before an alcohol episode. Before that section much later in the “Results,” however, we interpret zero WrisTAS™ TAC as representing zero alcohol.

Categories of Findings Reported

The following topics are reported in this “Results” section:

- Definition of episodes of drinking
 - Laboratory drinking
 - Self-dosed drinking
- Simple coded judgments of drinking detection
 - WrisTAS™ transdermal sensor
 - SCRAM™ transdermal sensor
 - AMS alerts (SCRAM™ algorithm detected hit)
 - Graphical examples of data
- Absolute deviation of the transdermal device peak from the BAC peak
- Time difference between the transdermal device peak and the BAC peak
- Percentage of deviation of the transdermal device peak from the BAC peak
- SCRAM™ check valve replacement/removal (see text)
 - Before removal of faulty check valve
 - After removal of the check valve

- Effects by subject sex
- Accuracy of SCRAM™ ankle bracelet as a function of duration of wear
- Definition of nonevents, smoothing algorithms, calculation of ROC curves, d' prime (d' , a measure of accuracy), and sensitivity analysis
- Effect of cold skin and sweating on TAC
- Wearability issues for 4 weeks
- Circumvention detection of SCRAM™ devices
- Offender focus groups – views of SCRAM™ wearers (content on CD)
- Provider (electronic monitoring providers) views of SCRAM

Drinking Data for Analysis

With a few known exceptions, each episode of alcohol consumption within each of the 30 distinct experimental trials when a subject wore the transdermal sensors in this study (i.e., the 22 unique subjects plus the 8 who served a second time) was marked by one or more PBT samples. During our evaluation, 309 episodes of drinking were logged either in the laboratory ($n=60$) or by the subjects conducting self-dosing ($n=249$). Of those 309 drinking episodes, 271 achieved a PBT-measured BAC value equal to or in excess of .02 g/dL. The range of BAC for these 271 episodes was from .020 g/dL to .230 g/dL with a mean of .078 g/dL. The data in Table 1 characterize the drinking episodes available for analysis. The two figures (3a and 3b) demonstrate the differences by type of dosing. In the laboratory setting (Figure 3a), the BAC results are symmetrical around .08 g/dL, the BAC level that we attempted to achieve with the dosing calculator based on the Widmark equation. All subjects were told to eat very lightly before coming to the laboratory to minimize differences in absorption time.

Table 1. BAC Characteristics Lab-Dosing and Self-Dosing Episodes

Type		N	Minimum	Maximum	Mean	Std. Deviation
Lab dosing	BAC max	60	.041	.123	.08287	.015771
Self dosing	BAC max	211	.020	.230	.07675	.044295
Total	BAC max	271	.020	.230	.07811	.039835

Figure 3b characterizes the BACs obtained from the self-dosing protocol. Self-dosing is simply normal drinking. Although evident from Table 1 that the means of the two types of dosing are similar, the self-dosing protocol provides some experience in evaluating the transdermal sensors when used by heavy drinkers.

For the 38 self-dosed drinking episodes that were $<.02$ g/dL (not shown), the mean BAC was .011 g/dL with a standard deviation of .005 g/dL. For all further analyses unless otherwise noted, these 38 low-BAC drinking episodes were excluded; the 271 episodes of $\geq .02$ g/dL formed the basic data elements for this analysis. This convention of ignoring values lower than .02 g/dL is the same as is done by AMS SCRAM™ with their offender monitoring algorithms.

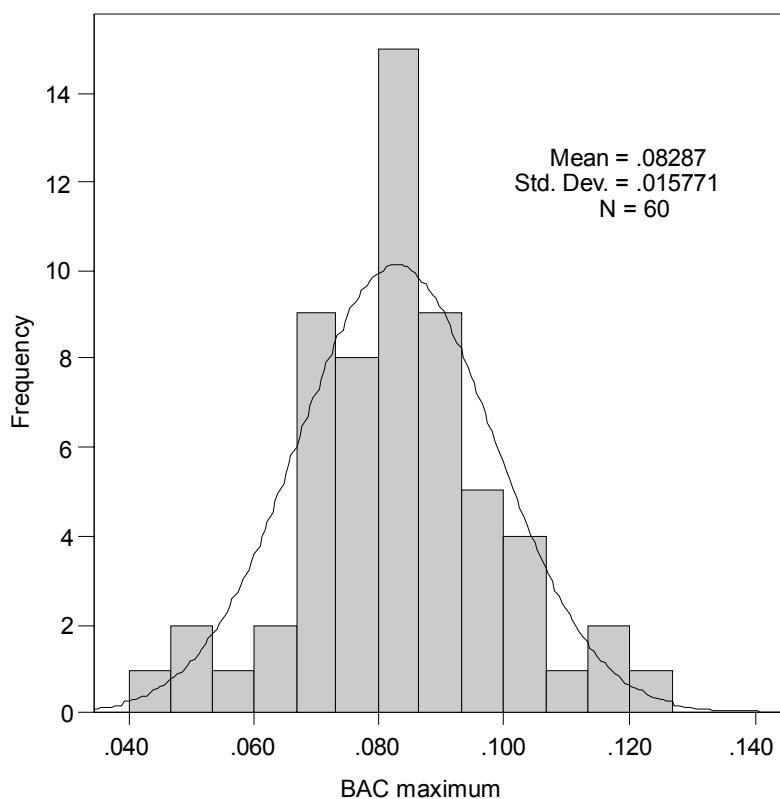


Figure 3a. BAC Maximum Levels Recorded with Laboratory Dosing

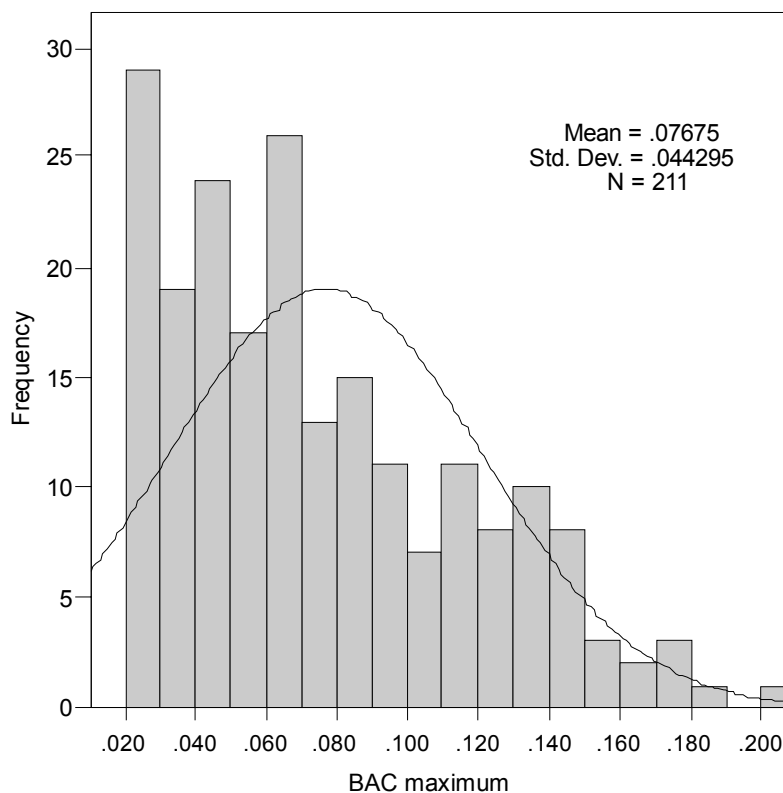


Figure 3b. BAC Maximum Levels Recorded with Self-Dosing

False Positives

In addition to the episodes just described with known BAC tests, there were 14 occasions of self-dosed drinking when no BAC test information was provided by the subjects. We know from the drinking logs that these were not genuine false-positive transdermal responses; drinking did occur, but we do not have the corresponding BAC values. Rather than log these as false positives, it was more valid to exclude these from the sample. In the majority of these cases, there was drinking, but the subject did not use the PBT to evaluate the BAC level. Reasons included forgetting to bring the PBT, did not have a breath tube, or as per our instructions, claimed to have had only one drink (the instructions called for logging BAC anytime two or more drinks were consumed in temporal proximity). In a few of those cases, elevated SCRAM™ TACs were an artifact of the initialization process, where voltages may result in artificially high TAC readings until the SCRAM™ ankle bracelet establishes a baseline reading for a non-drinking subject. Actual false positives among the subjects we studied were rare, and when false positives did occur, it was attributable to an undetected external source of alcohol. The concept of false positives has two different meanings with transdermal detection: (1) we can say with some confidence that there are few or no events that the devices created that look like drinking but were really not drinking because with drinking logs of well-compensated subjects and BAC results we knew definitively when real drinking occurred, but (2) we also know that there are external sources of ethanol signal that are unrelated to drinking that can cause a transdermal response. We detected one clear TAC event related to shaving cream containing an ethanol product, and there was one event that may have been related to perfume in the environment but we could not definitively identify the source. The shaving cream example did not look like a drinking event since it decayed too rapidly, whereas the other did look like a drinking event and had no clear cause.

External sources of alcohol are all around. Many personal hygiene, home, and automotive products contain ethanol, and people who use or are around these products may show positive for exposure to ethanol even though they consumed no alcohol. For example, aerosol Lysol has between 79 and 85 percent ethanol, and dozens of body sprays, such as Avon Naturals, have between 60 and 98 percent ethanol. The National Library of Medicine has a searchable database with product ingredient information (<http://householdproducts.nlm.nih.gov>). Restricting a search to ethanol-containing ingredients yields 447 consumer products that fall into eight categories, but of these products, 265 (59%) fall into one category: personal care products that are used on or near the skin. The issue of false positives is revisited in a later section dealing with the calculation of ROC curves. The recent paper by Sakai et al. (2006) reported no evidence of false positives in their SCRAM™ research. However, external sources of alcohol can and do result in positive transdermal responses because ethanol or ethanol-like ingredients are in many consumer products. It is important for vendors of this technology to adequately train their staff and customers to appreciate this.

Simple Detection: Sensitivity When BAC ≥ .02 g/dL

Each type of device had strong points and weak points. The following section reviews some of the evidence for simple detection by each type. The definition of a “hit” for the purposes of describing simple detection was the occurrence of any elevated transdermal alcohol equal to or higher than .02 g/dL in proximity with a time when the BAC was also higher than .02 g/dL. Defining a transdermal response as a true-positive “hit” was made through observation of the data series and curves for both WristAS™ and SCRAM™; the coding criteria are shown in “Methods.” In addition, for SCRAM™, a

supplemental method for judging a positive transdermal was available through use of the SCRAM™ alcohol detection algorithm. This is the protocol used by AMS for alcohol monitoring of offenders.

WrisTAS™ Judged Performance Detecting BAC \geq .02 g/dL

Comprehensive evaluation of the WrisTAS™ alcohol sensors was impeded by an apparently unreliable chipset (as understood from communications with the biomedical research director of Giner Inc.) that is unique to WrisTAS™ version 5. The chipset manages the storage and retrieval of data from the device version 5 that was under evaluation in this study. Of the 271 episodes of drinking with known maximal BAC \geq .02 g/dL, only 64 or 23.6 percent were hits, correctly detected by WrisTAS™ to be \geq .02 g/dL. A more accurate way to state that is whether or not the BAC was correctly detected by the Giner sensor; on only 23.6 percent of those occasions could correct data be coaxed out of the device for display and evaluation. It should be noted that a similar data loss was reported at the 2005 meeting of the Research Society on Alcoholism (RSA) by Greenfield et al. (2005), who also evaluated the WrisTAS™ version 5. They found that “of the 456 person-days of data collected, the WrisTAS™ unit failed on 194 person-days (43%).” They further distinguished between “all-person days” (n=246) that could be studied and “high-quality person days” (n=89). Accordingly, the findings we present here concerning data loss are not unique.

Using the protocol described in the “Coding” section, we determined that the problems with the WrisTAS™ fell into the categories shown in Table 2. The most common false negative subtype (37.6%) was missing data (no data captured because the device shut down), followed by data that were too erratic or noisy to be associated with a BAC (16.2%), or data that could not be judged a positive hit due to the rater’s low confidence (13.3%) whether it could be coded as a hit. False negatives due to response failure occurred in 8.1 percent of cases. These negative findings could be due to any of the following reasons:

- Recording of data would cease at some point between starting data collection and subsequent downloading.
- No data could be found in the device at all (the data record was blank or an old data record was not overwritten).
- Data record was erratic with no pattern such that there was little relationship between samples on successive 5-minute intervals.
- The TAC baseline reading varied over time, rendering it difficult to tell changes in TAC from changes in baseline. This happened for different reasons, including a change in the devices’ TAC calculation formula that caused TAC measurements to vary as a function of skin temperature. A calculated correction for this category of errors is introduced in a subsequent section.

Table 2. WrisTAS™ Judged Performance Detecting BAC ≥.02 g/dL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	hit	64	23.6	23.6	23.6
	<.02	3	1.1	1.1	24.7
	low-confidence	36	13.3	13.3	38.0
	too noisy	44	16.2	16.2	54.2
	missing data	102	37.6	37.6	91.9
	non-response	22	8.1	8.1	100.0
	Total	271	100.0	100.0	

A notable aspect of the results in Table 2 is the relative rarity (1.1%) of WrisTAS™ underdetecting a positive BAC (row <.02). This row contains only three cases, and in two of those cases, the maximal BAC attained was only .020 g/dL. This suggests that when the alcohol information is recorded properly, it appears to work well; however, the failure rate, as represented by version 5, is clearly unacceptable with over 75 percent of trials either not detected or not logged.

SCRAM™ Judged Performance Detecting BAC ≥.02 g/dL

Table 3 shows the SCRAM™ device correctly detected BACs equal to or higher than .02 g/dL in 155 of the 271 positive BAC events, for a valid hit rate of 57.2 percent overall. The SCRAM™ devices did not have missing data problems. Sixty (22.1%) of the positive BAC events were detected as a positive TAC but less than .02 g/dL. False negatives in which no response was found occurred in 14.8 percent of the events.

Table 3. SCRAM™ Performance Detecting BAC ≥.02 g/dL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	hit	155	57.2	57.2	57.2
	<.02	60	22.1	22.1	79.3
	low-confidence	7	2.6	2.6	81.9
	too noisy	8	3.0	3.0	84.9
	missing data	1	.4	.4	85.2
	non-response	40	14.8	14.8	100.0
	Total	271	100.0	100.0	

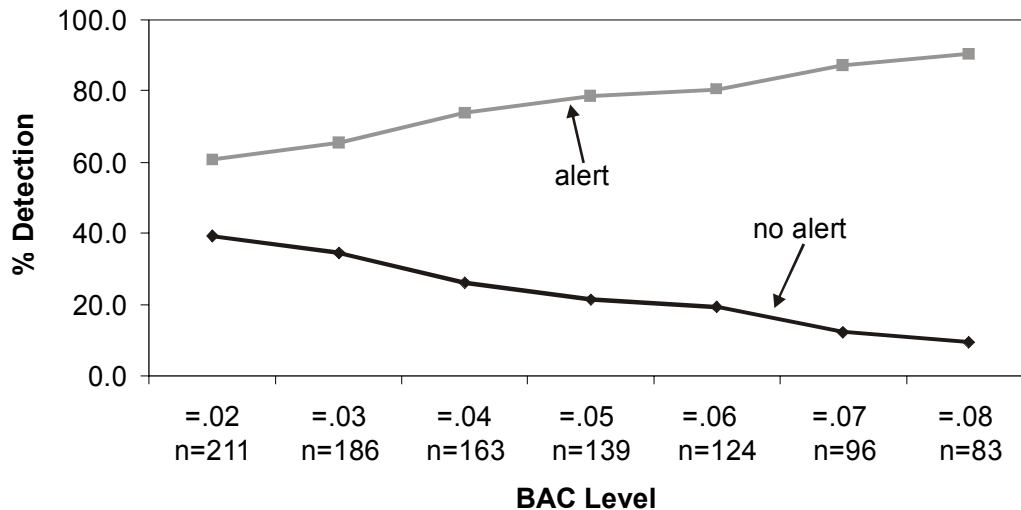
SCRAM™ Automated Alerts

With SCRAM™, there are two different methods to assess the detection rate, via coded rater judgments described in “Methods” as shown in Table 3 and through the “Alcohol Alerts” found on the AMS server. The PIRE-coded judgments and the AMS Alerts were independently scored; Table 4 shows a cross-tabulation between them. The Alcohol Alert derives from an AMS proprietary algorithm that serves to call a finding to the attention of an offender monitor or supervisor. The agreement for true-positive detection by these two methods was 93.5 percent and the agreement for true negatives was 91.5 percent ($\kappa = .85, p = .000$). The 155 coded hits (total row value) in Table 4 are the same events as the 155 hits in Table 3. Each method agreed on 145 events and differed on 10 events. Upon examination, all disputed events were near the thresholds for detection.

Table 4. Agreement of Two Methods for Estimating SCRAM™ Hits $\geq .02$ g/dL

Count		AMS alc alert		Total
		no alert	alert	
Coded Hit	.00	106	10	116
	1.00	10	145	155
Total		116	155	271

Because we have found no false positives of any note, the key question for a simple detection evaluation concerns the false-negative rate. The alerts provided by SCRAM™ can be used to portray how well SCRAM™ detects true positives in the range of BAC $\geq .02$ g/dL to BAC $\geq .08$ g/dL, the range in which SCRAM™ normally flags a TAC. The data in Figure 4 show the hit rates (true-positive rates) and misses (false-negative rates) at seven cutoff levels when subjects dosed themselves. Below the cutoff criterion on the x-axis is shown the number of drinking episodes represented by the plotted points. The data in Figure 5 show the same type of plot when subjects are dosed in the laboratory to a known BAC targeted at .08 g/dL (80 mg/dL).

**Figure 4. SCRAM™ Alerts (True Positive Detection) and No Alerts (Misses) by BAC (g/dL) Cutoffs When Self-Dosed**

When subjects are dosed in the laboratory, the alert or true-positive detection rate is much lower. These findings are summarized in Figure 5. The possible reasons for this are discussed later. In Figure 5, the lines remain parallel through the first three categories because there is no difference in the underlying data ($n=60$). There is no difference because the laboratory dosing of the subjects to the target BAC level of .08 g/dL was fairly consistent (and this is represented by the central [lepto] kurtotic distribution of the lab BAC data of Figure 3a). However, the contrived dosing regimen we used in the lab made it more difficult for the SCRAM™ devices to register elevated BACs. When the doses were higher, and especially when the duration of elevated BAC was longer, then the SCRAM™ alert was more likely to occur. This is well represented by the difference between the self-dosing and the lab-dosing evidence in Figures 4 and 5 and the slight growth in detection of TAC at higher BAC for lab-dosing.

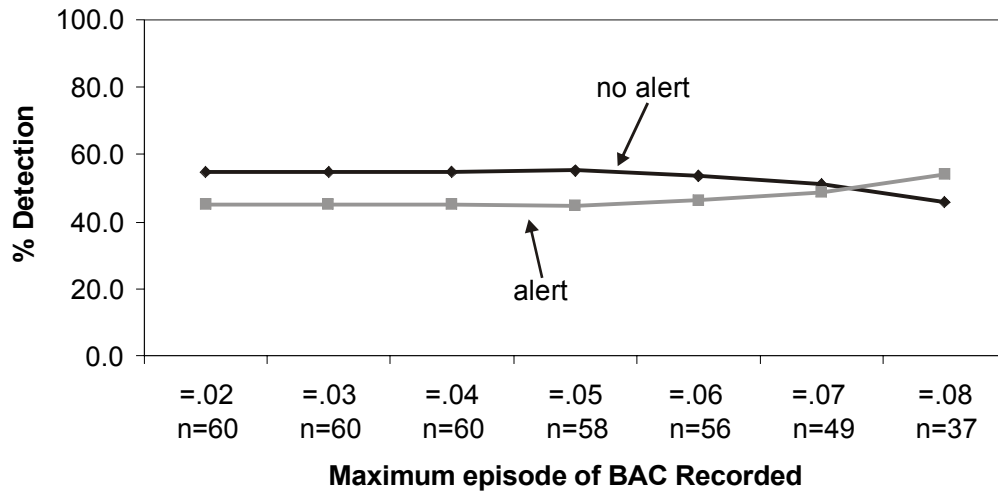


Figure 5. . SCRAM™ Alerts (True Positive Detection) and No Alerts (Misses) by BAC (g/dL) Cutoffs When Dosed in the Laboratory

Plotted Example Charts

Appendix F shows a complete 4-week chart for one subject. In successive pages, the date is represented on the x-axis, and the charts show data from SCRAM™ TAC, WristAS™ TAC, lab-measured BACs, and self-measured BACs. The y-axis scale varies from 0 to .3 g/dL (300 mg/dL).

Appendix G shows an example of each coded judgment type for each device. The judgment types were hits, TAC <.02 g/dL, low confidence, too noisy, and false negative. The category “missing data” is not represented as there is nothing to show.

Other Measures of Coded Hits and Misses: SCRAM™ and WristAS™

Magnitude of Peak Differences

There is a commonly reported 2-hour lag between a peak BAC and a peak TAC under normal circumstances. This lag has been described in the peer-reviewed research literature and in the materials from the private companies. The delay likely represents transit time of ethanol from body core out to the skin surface. On the descending limb of a BAC curve, the TAC curve usually lags behind the BAC curve more so, such that the absorption and elimination curves are asymmetrical.

When subjects dosed themselves with ethanol, the majority of the drinking episodes that were judged to be “hits” had a mean peak SCRAM™ TAC lower by .014 g/dL than BAC in paired comparisons of mean peaks in BAC (.095 g/dL) and SCRAM™ TAC (.081 g/dL) during the same drinking episode (this represents a mean difference of less than 12%). By contrast, in the laboratory dosing studies where the TAC response was muted, paired comparisons within a drinking episode found the mean SCRAM™ TAC peak (.047 g/dL) to be 44 percent lower than mean peak BAC (.085 g/dL). That is to say, the lab mean peak difference is three times lower than the self-dose peak difference for episodes of drinking that were detected as true positives, even when the BAC peak was approximately the same.

TAC output from the WrisTAS™ generally overestimates BAC levels in the self-dose situation if no correction is made for variable WrisTAS™ baselines. A sensitivity analysis in which baseline adjustments are made will follow in a later section. However, when we regard a zero value as a zero TAC, we find the paired comparisons of BAC and TAC for WrisTAS™ hits, estimated from mean peak levels to differ by .073 g/dL; this represents a WrisTAS™ TAC higher than peak BACs by 86 percent. By contrast in the laboratory situation, for true-positive hits, the WrisTAS™ devices logged TAC peaks only .019 g/dL lower than mean peak BAC (TAC 21% lower than BAC). The difference of mean peak TAC from BAC broken by type of drinking event is large considering the mean peak BAC in both situations was in the .08-.09 g/dL range. The SCRAM™ and WrisTAS™ devices appear to have different optimal BACs when accuracy is highest. SCRAM™ more accurately estimated self-dosing, and WrisTAS™ more accurately estimated lab-dosing. Neither device accurately estimated both types of drinking.

The Bland-Altman (1986) method was devised as a way to compare two types of measurement to examine similarities or differences, an approach used by Sakai et al. (2006) in their evaluation. When laid out for this method, deviations of the linear fit line from a horizontal line off the Y axis suggest there are underlying differences in two methods. For comparison purposes between BAC and TAC, data are from all results where TAC was judged a “hit.” The average of paired TAC and BAC maximum values are plotted on the x-axis and the difference between them on the y-axis. TAC results that are stable across all BAC values would line up with a slope of zero; when perfectly identical the mean difference would be zero on the y-axis. The charts in Figure 6 are plotted this way and show a fitted line along with 95 percent confidence intervals. The overall measured mean difference between BAC and TAC is represented by a lightly shaded dashed line from the y-axis. The left panel (Figure 6a) is for WrisTAS™ and the right panel for SCRAM™ (Figure 6b), both y-axes have tick mark steps scaled at .10 g/dL. The evidence suggests that WrisTAS™ TAC shows a very high systematic error such that higher TAC values linearly exceed BAC values. The SCRAM™ results capture zero within the 95 percent confidence intervals.

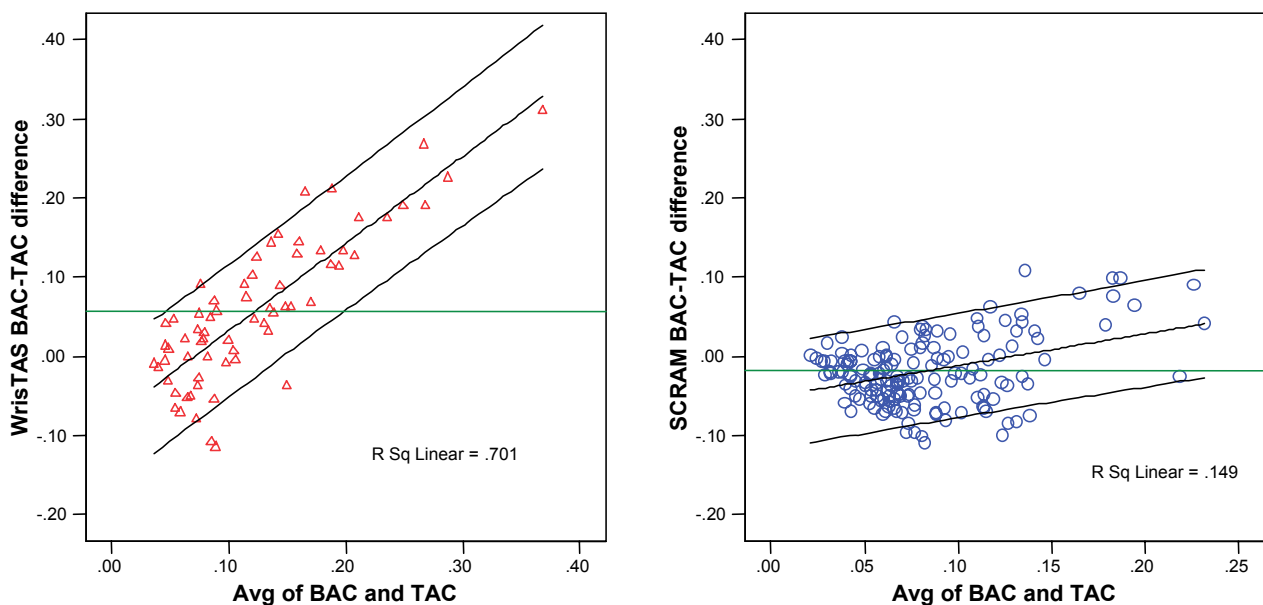


Figure 6. Bland-Altman Type Data Layout to Compare BAC and TAC for WrisTAS™ (6a left), and SCRAM™ (6b right) over a Range of Measured Values

A correction for WrisTAS™ baseline would reduce the upper end of its high levels, but if the adjustment were linear, it might also reduce detection at lower BAC levels. The difference in estimation may represent the different types of events each device is calibrated to detect (e.g., SCRAM™ self-dosed illicit drinking in an offender population and WrisTAS™ laboratory-dosing situations as a treatment research device). Tables of data for SCRAM™ and WrisTAS™ coded judgments are shown in Tables 5a and 5b.

For both devices, Tables 5a and 5b show that the rapid rise and fall of laboratory-dosed BAC resulted in a lower peak transdermal signal than was found with the more protracted self-dosed drinking episodes. In the laboratory trials subjects were instructed to consume their dose within 30 minutes. The dose in all cases was sufficient to elevate BAC to the target range (as shown in Table 1 and Figure 3a), and the form of ethanol used was distilled spirits diluted with some mixer (orange, cranberry, or tomato juices, or a soft drink). The rapid ascent and descent of BAC appears to prevent the transdermal alcohol concentration from following as closely as when individual subjects more typically self-dose with spaced, longer-duration drinking. It should be noted that there is little about the laboratory-dosing sessions that emulate real-world drinking (four to six drinks in 30 minutes and then no more). The time required for the ethanol to move across body water compartments appears to have the effect of flattening the overall transdermal curve for the SCRAM™ device, more so than for the WrisTAS™ device.

Time Delays in BAC to TAC Peaks

When there were true-positive hits with WrisTAS™, the timeframe of the results was less variable than when there were true-positive hits with SCRAM™. Estimating the performance of either device regarding the time lag is best done with the lab-dosed studies where we had control over the amount of alcohol under consideration. It can be seen in Table 5b that the mean time delay in a lab-dosed peak TAC was 2.40 ± 1.5 hours later than mean BAC for the WrisTAS™. This is within the expected range based on the published literature. In Table 5a, the SCRAM™ device had a mean peak delay of TAC in lab-dosing relative to BAC peaks of 4.5 ± 2.9 hours. This lengthier delay with SCRAM™ exceeds by about 2 hours the theoretically expected delay in TAC maximum. An explanation for this that appears consistent with our observations and our conversations with technical staff at AMS reflects a water management problem inside the SCRAM™ unit that may cause dilution and delay in recording TAC.

In the initial series of studies (the first 11 subject trials of the 4-week wear studies), AMS determined that a safety check valve was impeding moisture exit from the SCRAM™ unit due to faulty materials provided by its supplier and subsequently used in construction of the valve (see AMS letter Appendix B). After the first 11 subject trials (22 dosing studies), newer AMS devices were provided to the evaluators with the check valve removed. It may be that the delayed peak responses between BAC and TAC reflected residual alcohol laden moisture inside the SCRAM™ unit, leading to a continuing positive alcohol signal and/or a dilution effect of any new alcohol-laden sweat vapor coming in from the skin. This might explain the time delay found with the SCRAM™ units.

Table 5a. Characteristics of BAC Maxima and SCRAM™ TAC Maxima for True Positives (SCRAM™ HITS) by Type of Dosing**Descriptive Statistics - SCRAM™**

TYPE		N	Minimum	Maximum	Mean	Std. Deviation
lab	BAC maximum	31	.041	.123	.08519	.016125
	SCRAM max TAC	31	.010	.130	.04639	.025566
	TAC/BAC % diff	31	-85.29	42.86	-44.7248	29.71988
	TAC/BAC peak diff	31	-.08	.04	-.0388	.02701
	TAC/BAC peak time diff (hr)	31	.00	13.33	4.5108	2.91417
	SCRAM HIT	31	1.00	1.00	1.0000	.00000
	Valid N (listwise)	31				
self	BAC maximum	124	.020	.230	.09474	.041914
	SCRAM max TAC	124	.008	.271	.08086	.056369
	TAC/BAC % diff	124	-89.61	134.57	-12.3700	44.87133
	TAC/BAC peak diff	124	-.11	.11	-.0139	.04502
	TAC/BAC peak time diff (hr)	123	-1.08	26.80	5.1355	4.17534
	SCRAM HIT	124	1.00	1.00	1.0000	.00000
	Valid N (listwise)	123				

Note: Zero or negative time differences in TAC/BAC peak reflect cases when future episodes of drinking overlapped with past episodes. For SCRAM™, there were six such cases (four self-dosed, two lab-dosed).

Table 5b. Characteristics of BAC Maxima and WristAS™ Maxima for True Positives (Wrist HITS) by Type of Dosing**Descriptive Statistics - WristAS™**

TYPE		N	Minimum	Maximum	Mean	Std. Deviation
lab	BAC max any event	13	.041	.101	.08008	.018500
	wrist max TAC	13	.021	.145	.06131	.035326
	% wrist dev	13	-75.86	46.46	-21.5417	39.94847
	wrist peak diff	13	-.07	.05	-.0188	.03475
	wrist peak time diff	13	-.07	4.68	2.4013	1.50053
	wristhit	13	1.00	1.00	1.0000	.00000
	Valid N (listwise)	13				
self	BAC max any event	51	.026	.213	.09339	.042050
	wrist max TAC	51	.031	.523	.16682	.109821
	% wrist dev	51	-78.77	333.87	86.1260	93.50698
	wrist peak diff	51	-.12	.31	.0734	.09011
	wrist peak time diff	51	-.80	25.75	3.8807	4.00458
	wristhit	51	1.00	1.00	1.0000	.00000
	Valid N (listwise)	51				

Note: Zero or negative time differences in TAC/BAC peak reflect cases when future episodes of drinking overlapped with past episodes. For WrisTAS™, there were three such cases (two self-dosed, one lab-dosed).

Performance of SCRAM™ Before and After Check Valve Removal

A materials problem with the check valve, which is supposed to prevent moisture from entering the device, may have also prevented moisture from exiting. When removed, it raised the SCRAM™ true-positive hit rate in the laboratory from 8 of 22 (36.4%) to 24 of 38 (63.2%), a chance probability of .07; similar levels of improvement were found for males and females. Overall for laboratory detection, the mean TAC peak before the check valve removal was 50 percent lower than BAC peak, and after check valve removal, the peak BAC to peak TAC difference was 41 percent. These findings suggest that the valve was an impediment to correctly detecting alcohol in the laboratory, and removal of the check valve appears to have improved the device's detection capabilities. Although the gain in detection rate (i.e., hit rate) was large, the gain in accuracy (deviation between BAC and TAC) was small. Appendix B explains in more detail the nature of the check valve problem.

The check valve does not appear to explain the time delay in the peak difference between BAC and TAC maxima. Before check valve removal, the time difference between peaks in BAC and TAC was 3 hours, and after removal, it was 5 hours. We should caution, however, that these series may not be fully comparable as the hit rate was twice as high after removal of the faulty valve (difference values before valve removal are based on only 8 cases of lab dosing). The mean and standard deviation of peak time delay was similar before and after valve removal when looking at only the self-dosed events.

Sex of Subject – SCRAM™

As noted in the “Methods” section, we have 10 female subject trials (7 first and 3 repeats) and 20 male subject trials (15 first and 5 repeats). Across both types of dosing conditions and detection outcomes, females registered peak TACs that were 30.5 percent lower than their peak BACs, whereas males had peak TACs that were 14.6 percent lower than peak BACs.

Restricting the analysis to only those that were judged true-positive hits, the data in Figure 7 portray the percentage of difference between peak SCRAM™ TAC and peak BAC split by sex of subject. Figure 7 also portrays the percentage of difference before and after the check valve change. The sex differences are not statistically significant with the number of study cases reported here (hits only), but even when restricting to the true-positive hits, females trend toward a greater percentage of differences in peak TAC and peak BAC (the y-axis in Figure 7). It is well known that females have a lower percentage of body water than males, but whether this is related to the mean differences in Figure 7 is not known. This figure also suggests that among actual true positives, the check valve removal did not consistently influence the deviation of TAC from BAC; sex appears to be a more important influence on the percentage of deviation than was the check valve change. Across all episodes of drinking with BAC $\geq .02$ g/dL, SCRAM™ had 51 percent true-positive hits with females and 60 percent with males.

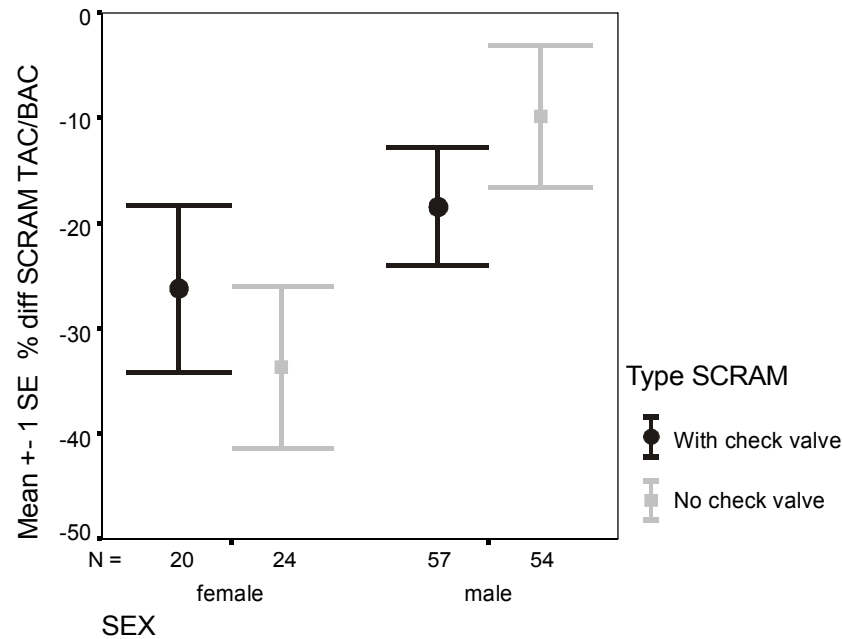


Figure 7. Mean Percent Difference \pm 1 Standard Error of TAC from BAC by Sex and Type of SCRAM™ Device

Sex of Subjects – WrisTAS™

The potentially lower transdermal response found for females using SCRAM™ technology was also found with WrisTAS™ technology. Although we know the true-positive hit rate of WrisTAS™ was low, there is a disproportionately lower WrisTAS™ hit rate for females than males. Of 185 drinking episodes with a BAC \geq .02 g/dL, WrisTAS™ on males correctly detected 31 percent of these, whereas WrisTAS™ on females detected only 8 percent of the 86 episodes with females. This difference is statistically significant ($p < .001$), as are measured differences in the peak values of WrisTAS™ TAC and BAC, and the percentage of difference in peak TAC and peak BAC. For males, the peak TAC measured 77 percent higher than the peak BAC, and for females, 48 percent lower than the peak BAC. This is shown graphically in Figure 8. It should be cautioned, however, that this large difference may be an artifact of only 7 cases of female drinking having been detected as a true positive with WrisTAS™. Also overall mean male BAC attained during true-positive events for WrisTAS™ was higher.

The consistency of the lower transdermal detection rates for female drinking may be more than an artifact; however, the available data do not allow us to speculate further about what may be causing these differences. There was no significant difference in the average BAC maxima for males and females ($.080 \pm .042$ for males and $.073 \pm .033$ for females).

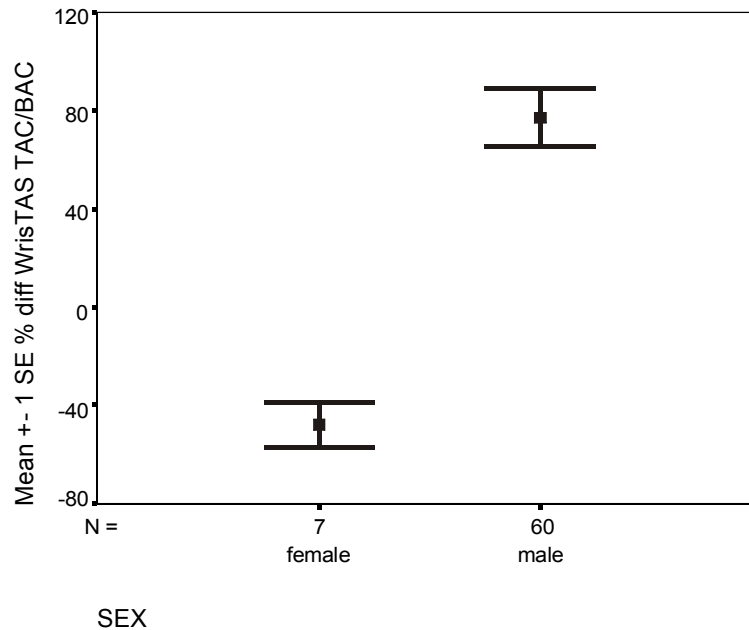


Figure 8. Mean Percent Difference \pm 1 Standard Error of TAC from BAC By Sex for WrisTAS™ Device

Within Subjects Change in Accuracy

SCRAM™ Within Subjects

When repeat wear episodes for the same subjects are compared, the first and second times wearing SCRAM™ bracelets permits an assessment of stability of performance across time. In the case of SCRAM™, this repeat assessment coincides with removal of the check valve, so the first and second laboratory dosing should show improvement if the valve was impeding proper performance of the devices. The mean alerts detected (using the AMS algorithm) in the first series with these subjects was just 25 percent, and after the valve removal, the alerts detected 62.5 percent of the lab-dosed BACs. For males, this detection was 30 percent in the first series and 80 percent in the second series. With the coded judgments, there were also large first- and second-round detection differences for males, with first-round hits equaling 40 percent and second-round hits equaling 80 percent. There were fewer females in the sample, and their detection was 33 percent for in both conditions.

WrisTAS™ Within Subjects

The WrisTAS™ devices appear to have detected more accurately in the first round (4-week study) than in the second round (2-week study). Judged WrisTAS™ hits were initially 50 percent for laboratory dosing, whereas with the same subjects, performance fell to an average of 12.5 percent in the second round. However, this likely reflects the large increase in missing data in the second round and should not be viewed as a change in accuracy separate from the growing problem with missing data (device non-response). As with SCRAM™, the performance with males was superior to females (60% dropping to 20% for males; 33% dropping to 0% for females). It is not clear if these differences by sex (and other sex differences already reported) are an artifact of sample size or are real consequences of differences in body water or other sex-specific physiological differences.

For example, Anderson and Hlastala (2006) modeled the kinetics of transdermal ethanol exchange and determined the stratum corneum, the outer skin layer, was the most determinative factor in ethanol concentration at the skin surface. Jacobi, Gautier, Sterry, and Lademann (2004) evaluated the stratum corneum of males and females and reported that females have a significantly higher skin pH than males but that other parameters (water loss, hydration, and sebum content) of this external layer do not differ by sex.

SCRAM™ Accuracy over Time

One of the hypotheses we formed while working with the SCRAM™ devices was the possibility that the longer subjects wore the devices, the less accurate the results. It seemed that the best trials were often the earliest ones.

Working from the coded judgments and the AMS alcohol alerts to indicate true-positive hit rates and using the most unbiased measure of peak accuracy – the percentage of deviation of the SCRAM™ peak TAC from the peak BAC, we evaluated the relationship between the percentage of deviation in accuracy and the duration of bracelet wear. Because the AMS server provides good detail on ankle bracelet history for each subject, we could plot the duration of bracelet wear so that any drinking episode and detection of it (true hits and percentage of deviation from BAC) could be related to the days of continuous wear.

Figure 9a has a time axis representing days of wear ranging from at least 2 through 16 or more, separated at 2-day intervals, until the final category labeled 16 days. At each point on the time axis, the chart shows the mean percentage of TAC relative to BAC ± 1 standard error. It should be cautioned that this chart represents a mix of both types of dosing (lab and self), both before and after the check valve change, and with both males and females (all are potential sources of variation). With the exception of the check valve change, the other variables should be spread fairly evenly within each category. The accuracy appears to drop off over the first 10 days and then settles into a period of higher variability, with a particularly large standard deviation at the 14-day category.

Plotting days of wear allows the episodes of drinking per day to vary (see Figure 9a). If the days of continuous SCRAM™ wear are ordered into a series and ranked, and then those ranks are broken out into 10 deciles, it distributes the number of drinking episodes evenly across categories. The deciles are shown on the category axis in Figure 9b ($n=27$ or 28 per decile to represent the 271 episodes of drinking). Plotted this way, it allows for a comparison of the mean true-positive hit rate with equivalent statistical power per decile. The real mean days in Figure 9b are shown below the 10 deciles of wear duration and from the 1st to the 10th decile, respectively; these range from .5 to 21.7 days. The figure shows the true-positive hits based on our coded judgments and the AMS alcohol alerts as overlapping lines that reference the y1-axis. On the right side (y2-axis), the mean percentage of deviation of the TAC from BAC is plotted (triangles). The percentage of deviation starts with decile 1 at zero on the y2-axis (a perfect match between BAC and TAC), and then TAC relative to BAC declines over time with increasing days of ankle bracelet wear. The hit rates (y1-axis) are based on a constant number of drinking episodes, but from decile 1 to decile 10, the number of episodes contributing to the deviation score on the y2-axis declines from 24 to 10 (this is because there is no TAC value to use in the calculation for false negatives). With fewer underlying cases, the variance increases, and this may contribute to the wide swing in the mean deviation score in decile 9. This is to a considerable extent reflecting the same data that caused the swing upward in Figure 9a at 14 days.

We considered whether a drop in hit rates and accuracy had some relation to the water/sweat vapor accumulation inside the devices. The water management issue was a topic that the AMS discussed

frankly with the evaluators during our periodic conversations. It should be noted that the trend toward lower true-positive hits over duration of wear did not appear to go away when selecting only those drinking episodes when subjects wore the SCRAM™ devices after the check valve removal. After check valve removal, the first 3 deciles combined achieved hit rates of .775 (.77 and .78 for the two methods for true-positive hits), whereas the last 4 deciles had hit rates of about .44 (.43 and .45).

It is important to consider that, although there is a decline in accuracy over time, this may be an artifact of testing subjects who consume alcohol regularly – that is, our subjects were required to drink alcohol, whereas AMS offenders are required to not drink alcohol. The nature of this evaluation study did not allow for an inference into accuracy following many days of continuous wear when a subject was not drinking. This would be an easy question to investigate by requiring abstinence for 2 weeks or more and then initiating self-dosed drinking. But that question was not anticipated in the work plan for this study, and regrettably we do not have an answer for it.

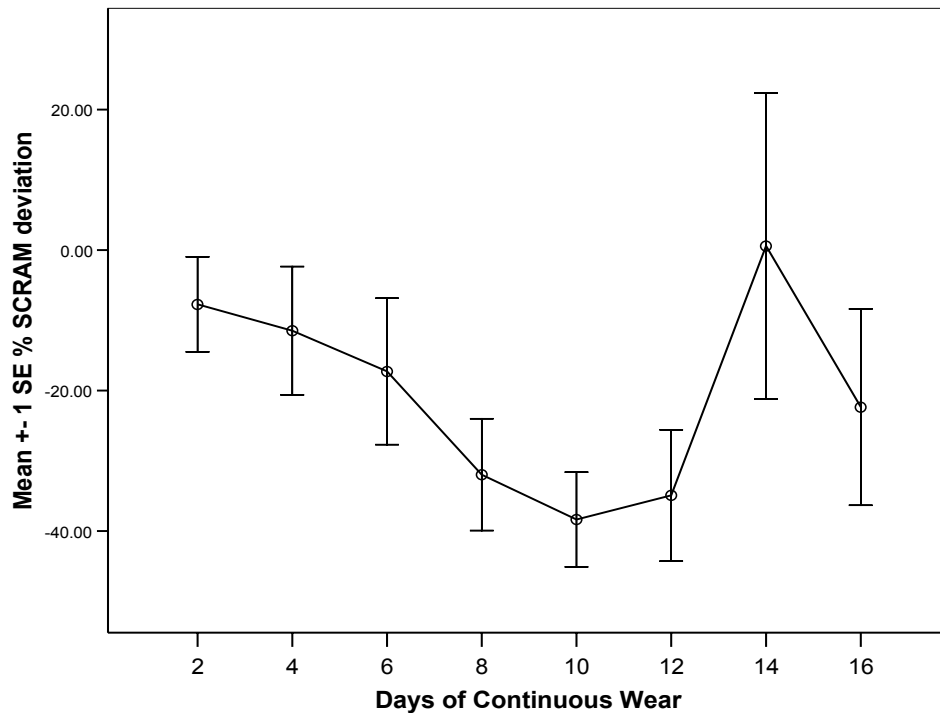


Figure 9a. SCRAM™ TAC Percentage of Deviation from BAC by Days Wear Time \pm 1 Standard Error

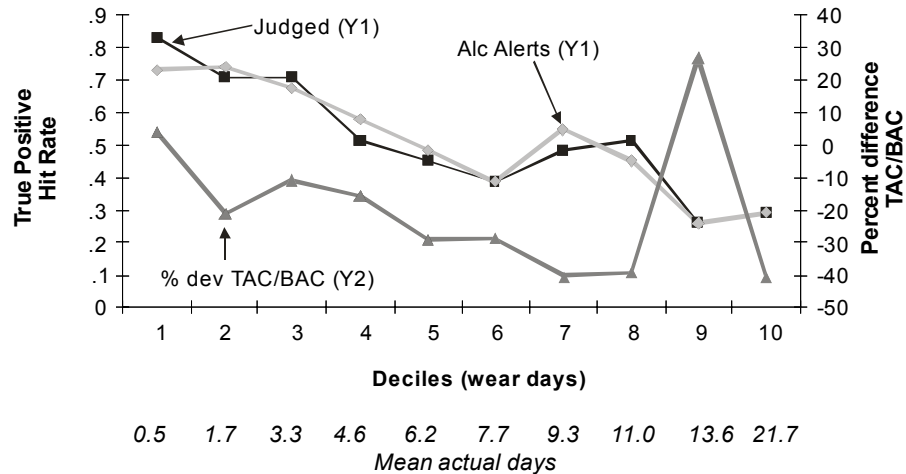


Figure 9b. SCRAM™ Hit Rate (y1) by Judgment or Alcohol Alerts and Accuracy (y2) by Continuous Days of Wear

Rule-Based Sensitivity Analysis

Up to this point, the evaluation has focused on the identification of simple decision criteria for judging whether a TAC response from the SCRAM™ or WrisTAS™ devices can be viewed as a hit or a miss based on all BAC episodes ≥ 0.02 g/dL. In addition to those coded judgments of whether a TAC response has occurred, we have also described use of the AMS alcohol alert algorithm as the criterion. However, the raw data (from which those judgments were made) can also be analyzed to build an algorithm from our own logic rules. This has the advantage of removing subjective judgments from coding and the disadvantage of needing to statistically remove known artifacts (e.g., such as transient spikes and floating baselines) that are often more easily found by just looking at a data record. The human observer is the most efficient pattern detector but is subject to various types of unintentional bias. The automated algorithmic analysis considers all the raw TAC and BAC data. As was noted in the Results Section, the judgment-based analysis excluded from the dataset 38 self-dosed drinking episodes that were < 0.02 g/dL BAC as well as positive TACs when BAC was missing for known reasons. In the raw data, all those events are included; the automated analysis is a fresh run at the raw numbers.

Setting up the Rules

The approach we took defines *four separate virtual experiments* for identification of BAC episodes. These correspond to screening for detection at four levels of BAC: $.02$ g/dL, $.04$ g/dL, $.06$ g/dL, and $.08$ g/dL. We calculated sensitivity statistics for each of these levels using methods derived from Signal Detection Theory (SDT), in which the true alcohol condition of the subjects (e.g., lower than $.02$ g/dL BAC vs. higher than $.02$ g/dL BAC) was predicted using a binary decision criterion based on the TAC reading, and then we noted the success of these predictions in the classical SDT 2x2 structure of true positives, false positives, false negatives, and true negatives. As before, however, we do not consider as false positive *measured* TAC events that exceeded a BAC once the BAC was actually higher than $.02$ g/dL. But when a BAC was less than $.02$ g/dL and the TAC was ≥ 0.02 g/dL, then that TAC is a false positive (in this way known episodes of missing BAC due to subject error will be flagged as false positives). If TAC exceeds $BAC \geq 0.02$ g/dL, then it has correctly detected a positive BAC. For example, a TAC of $.07$ g/dL is not considered a false positive for a BAC of $.04$ g/dL because the transdermal sensor would have done its forensic detection job and issued an alert

or flagged the positive. Accordingly, with these criteria it means that once the false-positive rate is established, then at increasing BAC test levels, there can be an increase in sensitivity but no change in the false-positive rate (1-specificity). In the virtual experiments, for each level of “real” BAC that we try to detect, we varied the criterion (or “cutoff”) value of the TAC through a range of thresholds (from .005 to .12 g/dL) to plot ROC curves.

To establish a false positive rate for TAC in the virtual experiments, it was necessary to have cases representing an actual negative condition for BAC — that is, occasions when there was a known zero BAC, which ideally should be recognized by the screening device as true negatives. We had a theoretical maximum of 60 such actual negative occasions when subjects came in for laboratory dosing. Our rules required that no one could be dosed unless their BAC was zero. So we calculated the false-positive rate upon those occasions. For SCRAM™, we used 58 zero BACs (two were excluded: one because there was no TAC produced and one because a subject’s TAC was on an obvious tail end of a prior drinking episode). With WrisTAS™, we used 42 zero BACs (one excluded was the same subject that SCRAM™ detected who was still elevated, and 17 were excluded due to missing data). When applying the screening cutoff level of .02 g/dL on the devices, the calculated false-positive rate for SCRAM™ was 12.3 percent, and for WrisTAS™, it was 25.9 percent. These are conservative (high) false-positive rates, representing 2 hours before the BAC test and 1 hour after the BAC test. WrisTAS™ responds more quickly to elevated alcohol than SCRAM™, and it is likely that some of the apparent elevation in the false-positive rate is due to a rapid transdermal response to the actual dosing that followed within 30 minutes of the zero BAC test. In addition, some of the unexpectedly high false-positive rates for both devices likely reflect the detection of some real elevated BACs from a prior night of drinking. Simply stated, the high false positive rates are to some extent artifacts of the algorithm setup.

The calculated rate of false positives therefore does not necessarily represent actual false-positive rates, as it was not possible to perfectly define and identify ideal instances of pure negative conditions. However, to automate the SDT analysis, there has to be some false-positive rate based on some known zero BAC value. The time intervals surrounding the laboratory dosing were the closest practical thing to “certain” zero BAC values that we had in the data set. Had we narrowed the window of observation, it not only would have reduced the false-positive rate, but also would have eliminated many of the data points available for making the estimate.

With respect to the true-positive rate (sensitivity), whether or not the transdermal signal correctly detected the BAC would depend upon the TAC result obtained *during* that drinking episode. For this algorithm, the drinking episode was defined as continuing from the earliest elevated BAC measure to an 8-hour interval immediately *after* the maximal BAC elevation. We modeled these data in three ways (e.g., allowing an episode to extend until the next episode or cutting off the period after 8 or 12 hours). The sensitivity results were generally higher when the interval was longer, but negligibly so, indicating that when successful, the devices would generally detect an elevated BAC early — within a few hours — if they were going to detect it at all.

The method we used, although it probably overestimated false-positive rates, was based on measured BAC values and allowed sufficient time after a maximal BAC for the transdermal device to respond. True positives, not false positives, are the more relevant measure for real-world applications of this technology. False-positive errors as a category have salience for people who work with industrial solvents or other alcohol-related environmental signals, so false-positive events should not be regarded as unimportant. But as a matter of normal performance and as discussed in an earlier section, there were few actual false-positive events found in the data when examined by

coded judgments. This is the same result reported by Sakai *et al* (2006) in their SCRAM™ evaluation of subjects who received no alcohol in one condition of their trial.

Dealing with Spikes and Other Noise Parameters

We evaluated these data unadjusted for noise and also with smoothing algorithms. Common time series smoothing algorithms (e.g., running medians) and other time-contingent logic rules of selection were used to reduce the influence of extreme spikes in the transdermal devices' data series. Without this adjustment, the rules-based approach might detect spurious "water spikes," transient interfering substances or non-alcohol-related positives to be true-positive events.

Dealing with Shifting Baselines

As noted earlier, the baseline that defines real zero on the WrisTAS™ is more variable than the nominal zero values provided on the raw data file and the charting protocol that Giner supplies. That is to say, even though the output from the WrisTAS™ may not read zero, we know from conversations with the Giner head technician that real zero drifts over time. Actual TAC values are more appropriately read as the difference from a stable low value to the high value during a drinking episode. To compensate for this drift, we calculated an algorithm that subtracts from the reported maximal TAC values a minimal "neo-baseline" level that we defined as based on the apparent low point observed in the 3 hours immediately before the drinking event began. For WrisTAS™, this represents 36 of the 5-minute samples (the device stores a result 12 times per hour). The 3-hour period preceding the rise caused by the alcohol signal was divided into three 1-hour intervals; the hours closest in time to the positive signal were weighted more strongly than the hour intervals 2 or 3 hours earlier. These 3 hours of WrisTAS™ signal then define the neo-baseline. These adjustments were made before determining the false-positive rates already described.

Signal Detection Statistics: d' and A'

For each of the four dichotomized levels of actual BAC, the TAC measures were also dichotomized at a level corresponding to .02 g/dL TAC, and then true-positive (TP) and false-positive (FP) rates were calculated, along with the SDT accuracy measure d' that ranges from zero (chance performance) to an effective practical limit of ~4.0. A d' of 4.0 would indicate a near-perfect ability to capture almost every actual positive without also mistakenly capturing any false negatives. This measure, d' , is calculated as the sum of the two z-scores corresponding to each of the sensitivity (TP) and specificity (1-FP) rates, and assumes that a sensitivity error has the same relative cost as a specificity error (an assumption which we realize is not true for those who would be mandated to use TAC devices).

These accuracy measures: sensitivity (TP percent), specificity (100-FP percent), and accuracy (d') are shown for each of four levels of BAC impairment in Table 6, all of which were calculated using the single-decision criterion of TAC \geq .02 g/dL as reflecting the practical cutoff value used in real life. That is to say, .02 g/dL is the level at which SCRAM™ units issue alerts, a BAC at which many interlocks lock out, and the lower end at which performance studies can begin to identify some impairment. It is the appropriate criterion.

The tabled values and the figures shown are prepared from those cases where the device under evaluation produced data – that is, the missing data described earlier as a problem with the WrisTAS™ are not represented here as false negatives. For both devices these sensitivity analyses evaluate the accuracy of the devices when they gave an alcohol output signal.

The sensitivity or true-positive rate for SCRAM™ (when it sees an alcohol signal of .02 g/dL at the skin) increases as the alcohol BAC signal increases. The detected values for SCRAM™ are within range of the numbers shown in Figure 4 for SCRAM™ hits. In Table 6, TP rates continue to rise from .02 g/dL at 65.3 percent through to .08 g/dL at 86.5 percent. In both analyses, the definition of a “hit” or true positive represents BAC detection when the TAC signal is .02 g/dL or greater.

For the WrisTAS™, the TP rates are nearly the same all the way across (from 84.4 and 82.3%) for BAC \geq .02 g/dL. This suggests that the device is working at a near optimal level once it detects alcohol – that is, the likelihood of this device detecting alcohol at .02 g/dL BAC is not very different from the likelihood of detection at .08 g/dL. By contrast, SCRAM™ sensitivity improves with higher BACs.

The SCRAM™ and WrisTAS™ sensitivity data are based on a varying number of drinking episodes at each BAC level and that number of events is shown in the row TPn in Table 6. As noted earlier, the false-positive rates are based on a fixed number of events and the FPn and FP rows do not vary with BAC level. Table 6 also shows area under the curve (A' which can be up to 1.0) and d' which reflects the degree of separation of the signal and noise distributions in standard deviation units. All the signal detection statistics shown are based on the smoothed and filtered data for which adjustments have been made for baseline drift and spurious spikes. The markers on each graph represent the point from which the d' statistic is evaluated. The closer the marker point is to the upper left corner of the graph, the higher the d' and therefore the more distinctly different the signal and noise distributions.

Table 6. Algorithm-Based Signal Detection Statistics for SCRAM™ and WrisTAS

	SCRAM™				WrisTAS™			
	.02 BAC	.04 BAC	.06 BAC	.08 BAC	.02 BAC	.04 BAC	.06 BAC	.08 BAC
TPn	265	217	174	117	179	146	121	79
TP	65.3%	73.3%	78.2%	86.5%	84.4%	82.9%	82.6%	82.3%
FP	12.34%	12.34%	12.34%	12.34%	25.90%	25.90%	25.90%	25.90%
FPn	58	58	58	58	42	42	42	42
A'	.821	.861	.890	.918	.840	.835	.836	.831
d'	1.392	1.580	1.702	1.982	1.464	1.469	1.436	1.392

ROC (Receiver Operating Characteristic) Curves

The theoretic ROC curves corresponding to the variable accuracy of the continuous TAC measures are shown in Figure 10 (SCRAM™) and Figure 11 (WrisTAS™). On the y-axes, perfect detection would have a true-positive rate of 1.0; on the x-axes, a detection protocol with no false positives would have a rate of zero. The position of each marker in Figure 10 is positioned at the intersection of false-positive rate and the true-positive rate for SCRAM™ (values from Table 6). The positions of the markers in Figure 11 overlay each other and are evidence that the WrisTAS™ sensitivity is not dependent on the BAC level at the levels evaluated in this study. This finding is in accord with the evidence in Table 2 showing only three cases of low TAC signal when a BAC episode was \geq .02 g/dL. The overlap of the curves suggests that the sensitivity of the WrisTAS™ devices has a ceiling effect at the BAC maxima evaluated and that its potential sensitivity is likely much lower than .02 g/dL.

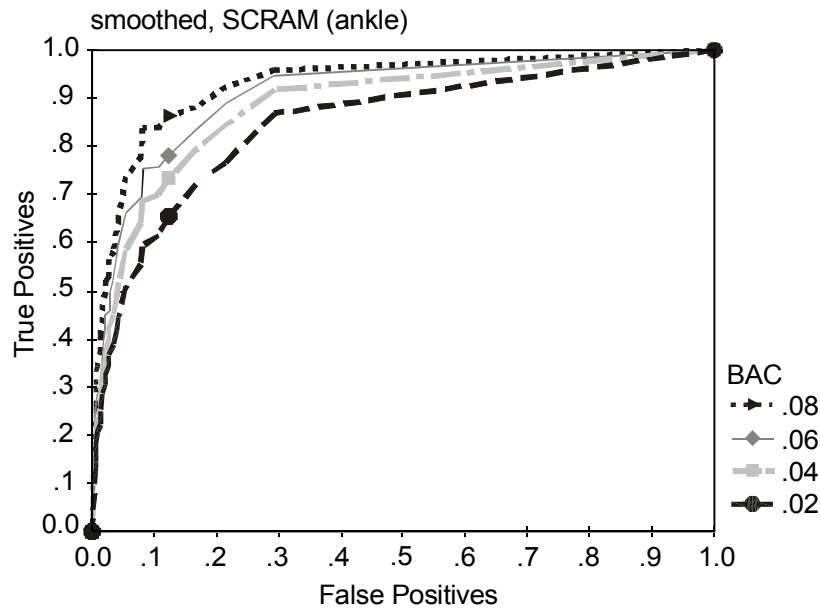


Figure 10. Signal Detection SCRAM™ at Four BAC Values

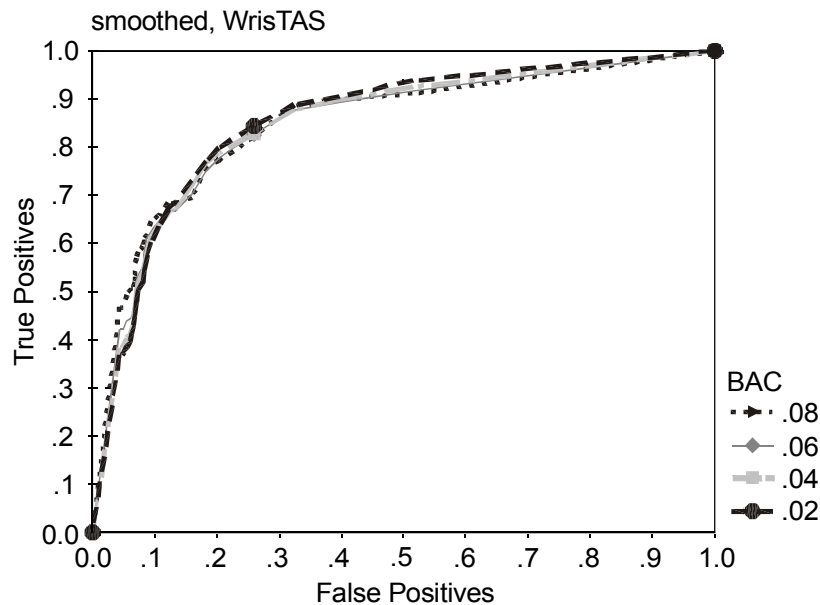


Figure 11. Signal Detection WristAS™ at Four BAC Values

These plots represent the inherent tradeoffs between sensitivity and specificity. We select .02 g/dL for reasons already cited. The results would be somewhat different if we were to use some other level of TAC as our dichotomizing decision cutoff. These curves are obtained by varying the TAC cutoff level (or, criterion) through a range of .005 to .120, producing a separate 2x2 table of “hits” and “errors,” for each cutoff value and then plotting the set of true-positive and false-positive rates from each 2x2 table. The bold plotted point shown on each graph is the presumptive decision criterion of $TAC \geq .02$ g/dL, corresponding to the rates and accuracy measures for the curves as shown in Table 6.

Circumvention Detection

All of the results provided and discussed to this point have concerned the alcohol sensors. Both the AMS SCRAM™ and Giner WrisTAS™ devices also have sensors for skin temperature (either in direct contact or near the skin), and sensors to detect either skin conductivity (Giner) or infrared signal (AMS). The latter sensors are designed to determine if the unit is no longer in proximity to the skin surface, or whether some blocking agent has been inserted between the skin and the alcohol sensor. Skin temperature can also be used to infer that a device is working normally. A benefit of the skin temperature sensor is that when examining multiple days of wear, the temperature sensor gives a good estimate of sleep cycles because while core body temperature declines during normal sleep hours, skin temperature ordinarily rises as heat leaves through the skin. Sleep cycles are seen in Figure 12 as 13 upward excursions.

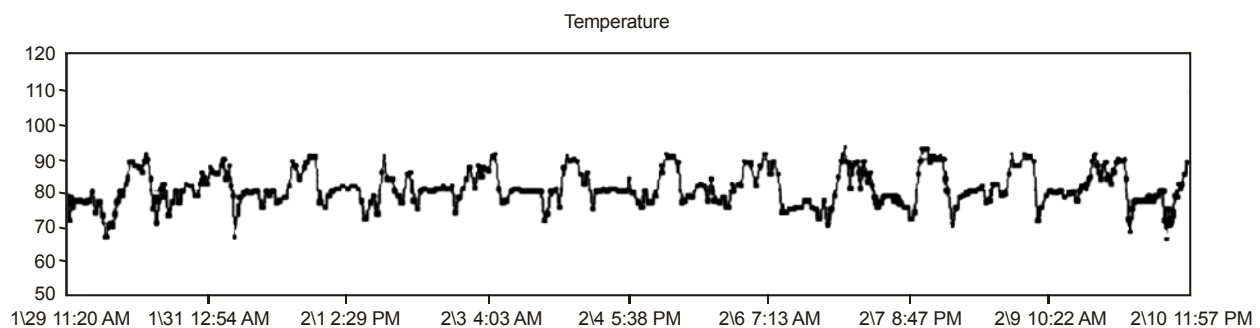


Figure 12. AMS SCRAM™ Temperature Sensor Data Over 13 Days for Subject MS10a (1/29/05-2/10/05) (Source: AMS screen capture)

To evaluate circumvention, PIRE staff wore SCRAM™ devices on both ankles for 2 days. On day one, our staff consumed alcohol and achieved a BAC near .08 g/dL. The purpose of this was to establish that both sensors responded in approximately the same way to the same alcohol signal, which they did. On day two, one of the sensors was blocked. We selected a thin but absorbent material (cotton) that was very close to the color of the skin and inserted it under the analog (fuel-cell) side of the SCRAM™ unit. The blocker was inserted just after hearing the sensor take a sample of air, thus allowing at least 30 minutes before another sample would be taken. Inserting the material was no problem; the material was intended to absorb the alcohol, and due to skin color matching conjectured that it might prevent the infrared sensor from detecting the barrier.

The result of this was that while the alcohol was absorbed, the infrared sensor issued an alert because the infrared (IR) voltage signal rose from 3 up to 5. The alcohol signal was attenuated on the blocked ankle (from .08 g/dL during the unblocked night to less than .02 g/dL the following night) but the infrared detector issued a tampering alert. The circumvention effort was detected. A variation on this study was tried with other materials (plastic, tape), but it appears that matching color of the blocking substance is not sufficient to fool the sensor. There is also the matter of reflectivity. The ideal substance will likely have to have a surface that reflects like skin but through which body heat can warm the temperature sensor.

In another study, we took unfair advantage of our knowledge of the SCRAM™ bracelet to determine if an offender with equivalent knowledge could block the signal. We knew exactly where the infrared sensor was located and this time inserted blocking material (thin plasticized wallet card) between the skin surface and the SCRAM™ bracelet, but not so far up the bracelet as to cover the infrared sensor. In this case, we did have partial success in preventing the skin alcohol signal from

reaching the fuel cell. The maximum BAC of .085 g/dL resulted in less than .02 g/dL TAC – there was no alcohol alert and no infrared alert (tamper alert) was issued.

Two cautions apply. This type of intervention is probably impractical for any extended period since it was both uncomfortable and required special knowledge of sensor locations. Possibly a redesign of the SCRAM™ unit that positioned the infrared sensor in between the vapor sampling areas would make it considerably more difficult to implement something like this, even knowing where the infrared sensor was located.

Although there are many materials or approaches that can probably prevent alcohol detection, the bigger challenge is to do so in a way that does not activate the infrared detector. A tamper is a tamper, so there is a whole series of circumvention efforts that do not need to be tested with alcohol. Our experience has shown that it is difficult to find absorbent or blocking barriers that do not lead to significant and noticeable changes in infrared readings.

By hearsay, we were informed that liquid skin or liquid bandage type products might be able to block the alcohol. This product is a substance that can be sprayed or brushed on and secures a wound and serves the same purpose as a bandage. We examined the contents (available in pharmacies) and found that it was about 7 percent alcohol. So to test it, it was necessary to get the product on the skin (after the SCRAM™ bracelet was installed) and then dry the alcohol before the device takes a sample. We determined that the substance can be applied without causing an infrared alert but that it has no ability to block alcohol. The TAC recorded was normal, a result that is not surprising as this product is advertised as “breathable.”

Environmental Challenges

The protocol to investigate the effects of exercise-induced sweating and low temperature exposure was only adequate to make a qualitative estimate of the direction of effect of each intervention condition. We imposed the low temperature manipulation for 10 trials, and the treadmill exercise on 10 trials. Our inability to more quantitatively investigate these variables was due to three factors:

- Humane considerations (subjects cannot be made cold for very long or required to exercise for very long).
- The surprisingly low TAC responses (relative to BAC), particularly for SCRAM™, that were attained in the laboratory dosing.
- The unexpectedly long delay in time to peak TAC following a BAC peak making it difficult to observe an effect on alcohol detection within the timeframe allotted for the lab-dosing studies.

Procedurally, subjects usually consumed their alcohol dose by 10:30 a.m. The planned scheduling (drink in the morning and test in the afternoon) accommodated a delay of 2 to 3 hours following the BAC peak to evaluate the effect of the low-temperature protocol and the exercise-induced sweating protocol.

Regarding the low ambient temperature protocol, we exposed alcohol-dosed subjects wearing shorts and a thin top to a cold room for 30 minutes. The room’s temperature was lowered with a special air-conditioning unit, plus a high-output fan that blew air directly on the subjects, after passing through 10 pounds of frozen blue ice. The temperature-lowering protocol was adequate to result in strong shivering and gooseflesh, and the data in Figure 13 show an exemplary plot of skin temperature

decline for leg ($\sim 6^{\circ}\text{C}$) and arm ($\sim 5^{\circ}\text{C}$) during the 30-minute period. Nearly all subjects responded with a similar degree of cold.

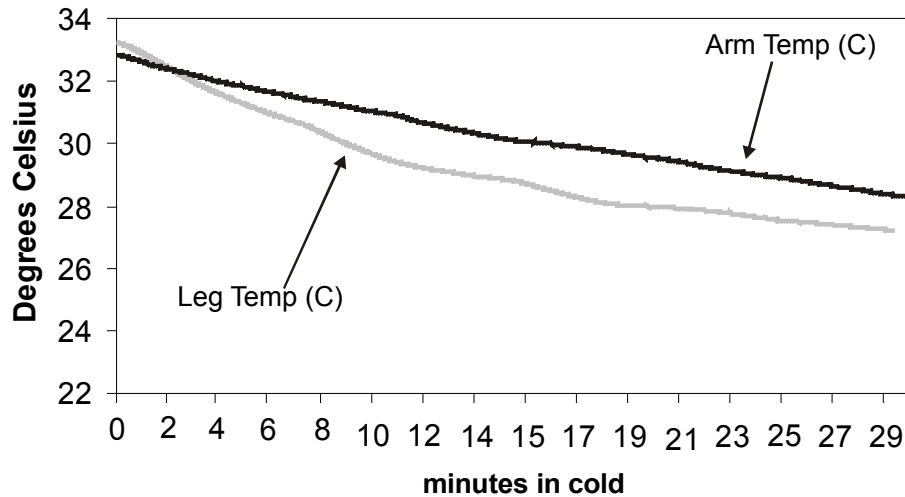


Figure 13. Skin Temperature on Arm and Leg of Subject While in the Cold Room

The qualitative effect of lowered skin temperature was small, but in about half the cases, there was a discernable but brief lowering of the alcohol signal by less than $.02\text{ g/dL}$. It would not have been apparent unless someone knew where to look for it in the data record. Other than alcohol, the effect of the environmental temperature manipulation could be seen on the SCRAMTM temperature sensor in which a drop of approximately 5° to 6°C was found. The alcohol response in the presence of this temperature drop was of three types: either there was no response because alcohol had not yet begun to be detected, there was a transient dip in the alcohol signal consistent with the temperature signal, or the temperature manipulation introduced some variability. One could infer from these results that if there is a TAC reduction due to the cold, it is temporary and small. It would be difficult to give a good test of the question of whether cold lowers SCRAMTM detection of alcohol within the constraints of human subjects' protocols.

The treadmill settings for evaluating the effect of elevated skin temperature and sweating involved 30 minutes of exercise at a setting of 3 miles per hour but at a grade of 9. This was like a walk up a very steep hill at a normal walking pace. It was sufficient to induce sensible (liquid) sweating in all subjects. The skin thermistors could not be reliably affixed in a way to document the change in skin temperature over the 30 minutes due to equipment limitations; however, at the end of the 30-minute exercise period, 3 to 5 minutes of skin temperature data were gathered to document the rise in skin temperature at that time. Most subjects showed a 2°C (arms) to 5°C (legs) degree rise in skin temperature over a 4-minute period as excess heat escaped past the skin surface and sweating continued. With starting skin temperatures of approximately 33°C , there are only a few degrees on the upside that can be attained. Normal core temperature is 37°C .

The treadmill sweating protocol affected the SCRAMTM infrared sensor reading, so this may mean that exercise could lead to tamper alerts. The general tendency was for the alcohol TAC signal to increase in the presence of sweating or due to heating of the skin surface, but as TAC was already elevated, no new alcohol alerts were possible. The limited duration of the exercise intervention and the slower-than-expected response of the sensors to alcohol made it difficult to offer more than these qualitative observations of the data.

Wearability and Acceptability

Overall, subjects tolerated the alcohol-sensing devices reasonably well. There were individual differences in the ability of subjects to tolerate either the SCRAM™ or WrisTAS™ devices, but none of the subjects found the problems to be great enough to warrant dropping out of the study. All subjects who began the study completed the study. Several commented that the bonus money for sticking with the study all the way through dispelled any idea of dropping out, even when they found the devices annoying. The largest complaints centered on the manner in which the device interfered with participation in exercise, particularly swimming, but also to a minor extent, soccer and running. Many subjects commented that they did not want to be seen as offenders and were a little embarrassed by the devices, often wearing long pants when they would have ordinarily worn shorts.

Having summarized these comments from research subjects, we found it interesting that similar comments were offered by the actual offenders who participated in focus groups to help us better understand their experiences as SCRAM™ offenders. More commentary on wear and wear-related problems are summarized in Appendices D and E.

Focus Groups of SCRAM™ Offenders

Seven SCRAM™ offenders participated in a focus group to help us understand some of the experiences reported by end users of this technology. A more complete accounting of this information can be found in Appendix E.

Users experienced both problems and benefits. Problems usually fell into the same categories mentioned by our experimental research subjects: some discomfort with wear; public embarrassment or embarrassment with close acquaintances; and some equipment-related problems, such as having to use landlines rather than mobile phones for the daily modem data transfers. In addition, the court-ordered offenders reported difficulties with the cost of the SCRAM™ device (typically \$12 to \$15 per day), and there was a strong fear of having their freedom (from going to jail) dependent upon the potential whims of provider or probation department employees who sometimes did not understand the devices. They strongly believed the devices were largely foolproof, and the general opinion reflected the view that it would be very stupid to risk any drinking.

In addition to problems, users also reported benefits. Chief among these was the enforcement of abstinence so that they experienced the benefits that derive from not drinking, such as a sharper mind, better sleep, better work performance, and better relations with family. In addition, the SCRAM™ providers also offered proof to the probation authority that the offenders' claims of abstinence were genuine. There was a resigned sense of gratitude expressed by several people that the alternative to SCRAM™, jail, was worse than having to endure the discomfort.

When asked about circumvention strategies, focus group participants might have tried, but they strongly expressed beliefs that there was no way to beat it. They felt the consequences of being caught tampering were not worth it. These seven offenders may not be typical as there was no random selection possible, so their comments should not be viewed as representative of SCRAM™ users.

Interview and Contacts with SCRAM™ Providers

Discussions with SCRAM™ providers who are installing devices on end users were conducted both in small groups at meetings and via e-mail and telephone. The general feeling among these providers (often private contractors who lease or purchase the equipment from AMS) was that the SCRAM™ system works well. In their view, about 20 percent of offenders violate the system and that, when violations occur, they are readily discernable. Among the violations discussed were cutting the strap and removing the device and various attempts to circumvent or block the signal, similar to those we tested. They claimed circumvention efforts were easily detected and, when the offenders learn they cannot successfully circumvent, most stop drinking and stop trying to beat the device. Of course, if someone devised a perfectly successful circumvention, the providers, by definition, would not know it. There was a strong feeling among the providers that the SCRAM™ devices are useful and that many of their offenders benefit from having to wear the bracelets. More detail from provider conversations can be found in Appendix H.

Discussion

This research on transdermal alcohol sensing adds information to the sparse literature in this field. In so doing, it confirms some facts that are well known and uncovers some previously undocumented findings that bear on the nature of transdermal alcohol sensing. As noted earlier, this project was not conducted for the purposes of comparing two devices in order to pick a winner. SCRAM™ and WrisTAS™ represent the universe of available technologies in the larger category of transdermal electrochemical alcohol detection. This technology, already in use in the United States, can only benefit from more active independent study of its strengths and weaknesses.

Depending on the application intended, electrochemical alcohol-sensing units mounted on the skin surface can have advantages over other forms of protracted alcohol detection such as biomarkers or frequent regular breath, blood, or saliva tests. The devices collect and store data on their own, accumulate results over time, and can be combined with tamper protocols. These features reduce human monitoring time, and the time needed for results analysis. Blood, urine, saliva, or hair biomarkers all require chain of custody sample collection, a potentially long time for measurement and /or significant costs per sample. BAC monitoring by telephone must by law exclude large portions of a 24-hour period. Transdermal alcohol-sensing technology does not have these limitations.

Of the electrochemical devices under evaluation, only the SCRAM™ unit is commercially available at present, but the technology underlying both devices is commercially viable. Both companies have cooperated fully with the project, combining the best of openness with their product information and an eagerness to acquire any new information that might help improve their products in future versions.

Giner WrisTAS™

When discussions began with Giner Inc. to acquire WrisTAS™ units for evaluation, they were in a transition phase between their version 5 and version 6 devices. Giner Inc. is a small research and development company that manufactures sensors for environmental gases and other products, many unrelated to biomedical applications. The alcohol detection product built around their proton exchange membrane is not available for the commercial market. Their version 5 prototype has had known problems with data integrity, and it does not have a remote telemetry system built into it as does the newer version 6. We expressed a desire to evaluate version 6, but it was not yet available. We ultimately agreed to evaluate their older version since it had already been the subject of several research studies by Swift and others. The data retrieval interface with version 5 is via a DOS-based program that uploads data from the unit through a hard-wired serial port (a serial port on a port replicator will not work). The program will run under DOS or in the DOS emulation box of some, but not all, Windows versions. Unlike the sensor itself, this aspect of version 5 is surely not cutting-edge technology.

As the research began, we learned to appreciate the distinction the Giner staff makes between WrisTAS™ devices and “pre-qualified” WrisTAS™ devices. When not pre-qualified, the devices have a higher rate of data loss. As our results demonstrate, we readily duplicated the data loss problem. We found that some WrisTAS™ devices became erratic that were once completely stable and dependable. There are different types of data loss, any of which rendered large chunks of data unavailable and formed the basis for the judged category of “missing data.” Sometimes a device

would be working well when it went out the door on the wrist of a research subject only to discover 2 weeks later that it had stopped collecting data within an hour after leaving the office.

We do not know where the WrisTAS™ data go when lost, but we do know that 38 percent of the data was missing and another 16 percent was too noisy to be used. We often did not know if the data were collected in the device and then lost, never collected, “stuck” so numbers could not be downloaded to the computer, erased by some stray electromagnetic signal, hit by a physical injury, or some other factor. Sometimes data were retrieved, but they were meaningless and patternless. We have estimated that slightly less than 25 percent of the BACs that exceeded .02 g/dL was detected as a hit even though all subjects (with a one-week exception) wore WrisTAS™ devices throughout the study program, and those devices were regularly serviced and downloaded at weekly or biweekly intervals. On the other hand, the signal detection analysis determined that when the WrisTAS™ correctly detects an episode of drinking, it can detect low, medium, or high levels of drinking with approximately equal sensitivity. The 4-week charts of a single subject (Appendix F) demonstrate how well this device can work when it works properly. Also evident in this single subject’s data is the occasional problem with shifting baseline and occasionally erratic readings. With this subject’s data, however, the WrisTAS™ device recovered nicely from noisy intervals, even though the baseline remained elevated.

Among actual true-positive hits, we found that the mean WrisTAS™ accuracy was quantitatively quite good in the laboratory, but when the WrisTAS™ registers the long duration self-dosing events, the estimated level of TAC runs high. This conclusion is based on data as represented in Figure 6a wherein the linear fit line shows divergence of TAC and BAC at the upper end of the measurement. This is not a serious problem for most applications because the error level is systematic and therefore correctable by software or scoring adjustments. In addition, we also know that the functional zero point on the WrisTAS™ record can vary widely. For the applications needed by most alcohol monitoring programs, high sensitivity even as low as .02 g/dL is highly desirable, but as currently implemented, the only method for determining a true TAC is to visually subtract a positive deviation from an effective zero baseline that can range as high as .10 g/dL (on the nominal y-axis).

In addition to the sensitivity evidence that we report here and that Swift and colleagues reported on as was discussed in the Introduction of this report, University of Indiana alcohol clamping studies (personal communication with Dr. Robert Swift) further corroborated this effect. In that research, subjects were held to a near constant BAC for an hour or more through feedback to an alcohol infusion pump. Their results showed that WrisTAS™ can precisely rise to stable levels consistent with BAC and fall in equally precise fashion once the alcohol signal decays. It is unfortunate that all the WrisTAS™ devices reported on here cannot match this degree of performance or the performance even of the generally excellent device shown in Appendix F.

The most plausible explanation for WrisTAS™ data loss is the one offered by Giner itself – that is, the chipset that was designed to handle data I/O functions in version 5 is faulty. Giner believes the new version 6 has solved many of these problems because it is based on a new chipset. This explanation, together with the sensitivity data we have seen when the device is working well, suggests the sensor itself may not be the problem.

Obviously, however, until the new WrisTAS™ version 6 or subsequent versions can be evaluated and shown (a) to have overcome the data loss problems, (b) that they are designed with adequate circumvention detection protocols, (c) that they can be secured so removal is difficult, (d) that they can be made more affordable, and (e) that they have functioning telemetry features that can reliably transmit data, it will be premature to consider WrisTAS™ devices as suited for use in offender

monitoring. On the other hand, if the promise of this sensor can be harnessed, packaged, and made available for offender monitoring at an affordable cost, it could form the basis for new applications of transdermal alcohol detection. We know of nothing available like it that detects alcohol in such a sensitive manner and in such a compact package. If the small size could be retained, and the I/O problem overcome, as anti-circumvention and communication capabilities are added, it would show great promise.

AMS SCRAM™

The SCRAM™ device is currently in use by courts and correction agencies and therefore more thorough discussion is warranted about its performance. SCRAM™ devices are reportedly in service around the United States; devices are leased or sold to direct service providers or government agencies. From their own promotional materials, AMS claims to have more than 800 courts and agencies using SCRAM.

Accuracy and Sensitivity

There were two significant problems with the SCRAM™ bracelets uncovered in this evaluation. These are a higher-than-expected false-negative rate and the possibility that sensitivity and accuracy of the devices decline over time. Because sensitivity declines over time, and because the laboratory studies were a poor emulation of real drinking it is not very meaningful to focus on an overall average performance. Nonetheless, overall, across laboratory and field evaluations, 57 percent of the BAC events were judged hits and another 43 percent were misses for various reasons. Half of the missed events were judged as evident TAC responses but less than the cutoff of .02 g/dL, and therefore, less than the level at which AMS' own automated system would issue an alert to the program monitor. Unlike the WrisTAS, there were few TAC responses to BAC \geq .02 g/dL that fell into the categories that we judged and labeled "missing" or "noisy." Among the 43 percent non-hits, after accounting for all response conditions, 15 percent were judged non-response false negatives, meaning no evident response to the elevated BAC. The 43 percent (including pre and post valve change) categorized as not true-positive hit is unexpectedly high.

The declining accuracy over time to a great extent accounts for the reduced SCRAM™ true-positive hit rate. As can be seen in Figures 9a and 9b, initially near-perfect true-positive hit rates and accuracies fall off with time in service. We have come to understand this reduction in sensitivity and accuracy may be associated with water accumulation in the sensor area. For devices that are often assigned for 3 or more months of wear, a halving of sensitivity during the first 2 weeks is troublesome.

Concerns about both accuracy and sensitivity all come with the same caveat. SCRAM™ bracelets are mostly used with people who are supposed to not consume alcohol. All of our subjects were required to consume alcohol. We cannot know if the measured decline in sensitivity and accuracy over time (as shown in Figure 9b) is related to alcohol exposure or is a consequence of wear (and consequent exposure to vapor sweat) independent of alcohol. Does time itself and exposure to vapor sweat reduce sensitivity and accuracy, or is decline in accuracy and sensitivity over time a consequence of repeated alcohol exposure? Our experience of occasionally low accuracy with light drinkers makes us somewhat skeptical that repeated alcohol exposure is the cause of the decline, but it cannot be ruled out. If it is the former, then it raises the question as to whether SCRAM™ would be capable of detecting more than half the drinking after the first couple weeks of wear. If it is the latter, then it is probably less important as SCRAM™ users, with some exceptions, should not be drinking.

When we designed this research study, we did not provide for exposure effect because it was unanticipated. It would be a straightforward study to conduct and could be addressed by waiting for 2 or more weeks of wear before testing the devices for sensitivity to alcohol. Alternatively, it could be evaluated by simultaneous BAC or alcohol biomarker assessments of offenders, or by promptly evaluating the sensitivity and accuracy differences of devices going out or coming back from offenders. AMS staffers have mentioned a “water management” issue with their devices, and we know that some research subjects were more prone than others to cause “water spikes” or transient elevations in the alcohol sensor readings that are clearly unrelated to drinking. Accuracy problems over time may reflect moisture accumulation in the sensor area that dilutes the alcohol signal and leads to extended periods of low amplitude signals long after we might expect the actual transdermal alcohol to have declined.

As we initiated a plan to evaluate SCRAM™, it seemed sensible enough to control conditions in a laboratory setting so that paid research subjects could consume known quantities of alcohol and the response of SCRAM™ monitored. At the outset, we also decided to propose a field research element using handheld fuel-cell PBT devices. Self-dosing and self-BAC measurement served to significantly increase the positive BAC episodes available for evaluation. The biggest risk with this study element was entrusting each subject with the two transdermal devices and the PBTs (about \$3,500 worth of equipment); fortunately this proved not to be a problem. As is evident now, we found that the field-collected data provided a better evaluation of SCRAM™ sensor accuracy relative to maximum BAC level. It was surprising how poorly the laboratory testing served as a proxy for normal drinking and detection of that drinking. In reality, people do not drink that way, so the laboratory model was a poor emulation of “normal drinking,” and proved to be a suboptimal method for evaluating the SCRAM™ devices.

The laboratory non-response false-negative rate for SCRAM™ was no worse than self-dosed (both around 15%), but the most striking difference was the accuracy as measured by the percentage of deviation of SCRAM™ TAC from the maximal BAC within an episode of drinking. SCRAM™ TAC was 45 percent less than maximal BAC in the laboratory testing but only 12 percent lower than the maximal BAC in field testing. This relative discrepancy was less a problem in the laboratory for the WrisTAS™ TAC (21% less than lab BAC maximum versus 86% more than field BAC maximum). It may be that there is something about the laboratory protocol of rapid consumption that does not allow sufficient time for movement of the alcohol across the various dermal membranes before the enzymes convert it. If, as the Widmark equation suggests, the average rate of metabolism is .017 g/dL/hour, then even a 2-hour delay in peak TAC means that .034 g/dL of detectable signal would have been diminished (approximately a 45% reduction). It may be that a transdermal device’s accuracy has to be tuned either for rapid consumption or protracted consumption. Most of the WrisTAS™ trials have been laboratory-based studies, whereas the SCRAM™ devices are calibrated against stable wet bath alcohol concentrations before they are released for detection of field self-initiated drinking. This might explain why each device has its own domain of accuracy with different types of alcohol administration.

Communications

A unique feature of the SCRAM™ product is the Scramnetwork server and the fairly seamless way that it communicates and uplinks data from the ankle bracelet. The server authorization process allows for varying levels of security and access that is granted based on the security clearance of the person who wants to see the data. From the company’s own top-level access, multiple call center supervisors might monitor several agencies, agents nest within agencies, and clients within agents.

For each new client introduced into the Scramnetwork server, the equipment is initialized and a specific bracelet is linked to a specific modem.

During initialization, when a new client begins a period of wear, the modem and bracelet communicate first with each other, then the modem calls into the server and initializes the process. Equipment is “checked out” so that dates of service initiation are clearly logged, and when service ceases, the equipment is checked in. This is done remotely without the need for anyone at AMS to intercede. Because the modem is typically set to call into the network every 24 hours during the sleep hours, the local manager/supervisor can examine new data daily after uploading. Between successive clients, the local installation center uses a disinfecting kit for the surfaces that come in contact with skin.

The problems we encountered with the communication system were minimal. The biggest problem that we found is the current requirement that it be used only with landline (wired) telephones. This has become a problem as many people now use mobile phones exclusively for communication. Occasionally other problems, even with landline telephones, have occurred that may be associated with DSL lines or other competing sources of signal that may interfere with transmission. However, if call-in is prevented for 24 hours, the server flags this and issues an alert to inform the agent (in this case, our research staff), who then communicate by voice with the individual and resolve the problem. Overall, the data aggregation and reporting system is well designed and works well.

Other Considerations

Finding a doubling of SCRAM™ detection rates (from 36 to 63%) with males after the check valve problem was resolved may partially explain problems with the earlier equipment. We do not understand why the improved valve did not also result in improved results with females or why the overall percentage of accuracy improvement relative to BAC was only marginally better than it was before the valve change. The female difference might be related to sample size (although there were an overall 86 test trials with females, 20 in the lab and 66 in the field). The mere marginal improvement after check valve removal may be a consequence of allowing some of the water out of the SCRAM™ device, but not enough of it. With the check valve removed, there is an escape path for water vapor, but it does not mean that all the water is in a vapor form, and the sensor area could still be partially saturated.

But as we puzzle over why female TAC accuracy was low, it should be noted that the problem was not restricted to SCRAM™. Both SCRAM™ and WrisTAS™ performed more poorly with female subjects. Relative to males, females had a greater percentage deviation of TAC from BAC for both types of devices in both lab-dosing and self-dosing situations. This may be worthy of focused study, and it would be interesting to know whether skin pH or the higher subcutaneous adiposity of females accounts for flattening of the TAC peaks. We are not aware of other transdermal studies commenting on sex differences in detection.

As to wear and comfort, more females reported discomfort with the wearing of the SCRAM™ device than did men. Some suggested the size of SCRAM™ was more suited to men than women. Women in the study, and in focus groups, were more likely to complain about the tightness of SCRAM™ bracelets and of SCRAM™ limiting clothing options. Women also found it annoying to have to work around the bracelet when shaving their legs, and in one case, a clear false-positive was found during use of “Skintimate” shave cream, a product that contains triethanolamine. These spikes following brief external exposure to alcohols are easily distinguished from elevated BAC (and are not flagged by the automated alert system due to the rapid rise and fall), but there is the larger problem of

detecting exposure to various alcohols or industrial chemicals on a regular basis if such substances are in someone's work environment. AMS is aware of these problems and handles them on a case-by-case basis when they occur.

Signal Detection Analysis

Writing an automated algorithm for classifying transdermal events in a 2x2 table of true or false positives and negatives poses several challenges. There are transient spikes to exclude and fluctuating baselines to accommodate, but one big problem was not having BAC test logs recorded during times when subjects were not drinking. That would have provided a larger database from which to find more true negatives rather than relying solely on the first BAC samples on laboratory test days. To calculate the signal detection analysis (SDT) analysis with the ROC curves, all four cells need data. Our only solidly zero (0 g/dL) BAC data were those obtained before those morning laboratory-dosing events. Accordingly, the false-positive rates are unnaturally high for both devices because we took a wide 3-hour interval around those zero BACs (to get 3 TAC data points on SCRAM™) and any positive TACs that were not sufficiently filtered out by the algorithm would have registered as a false positive. Nonetheless, the ROC curves have allowed us to determine that SCRAM™ and WristAS™ differ on low-end sensitivity. WristAS™ had similar sensitivity at all test BAC levels and, as is shown in Figure 11, with overlapping ROC curves (at fixed false positive rates). In Figure 10, SCRAM™ sensitivity was much improved as the test BAC level increased. The pattern of sensitivity detected in the automated SDT analysis is very similar to the pattern from AMS alcohol alerts found in Figure 4, which shows the plotted results of the AMS algorithm for detection. In both of these very different types of detection exercises, the true-positive rates increased from the low to mid 60s for all BACs of .02 g/dL or greater and ranged up through the mid to high 80s for BACs of .08 g/dL and higher.

Circumvention Detection

It is possible for a motivated and knowledgeable offender to engage in at least some degree of drinking without detection, but this is not likely to be a serious problem, especially among offenders who will typically be required to wear SCRAM™ transdermal devices for months. Theoretical circumvention and practical circumvention are not the same. Other technologies that detect alcohol, such as alcohol ignition interlock devices, can also be circumvented with adequate motivation, but in both cases, the proportion of offenders who could regularly circumvent with success is likely to be very small.

Due to the poorer low-end sensitivity, a concern with SCRAM™ could be detection of low-level consumption because two drinks spread out over a few hours at typical metabolism rates would keep BAC and TAC below the detection level. That degree of drinking is not the type of drinking that led the offender to the attention of authorities in the first place. However, if there were an abstinence order in effect, such drinking would nonetheless represent an undetected violation.

Also in our environmental conditions, we showed that temperature and sweat probably do affect the detection of skin alcohol for reasons that make physical sense (i.e., alcohol volatility and skin blood flow are both dependent on local temperature). However, the amount that we could detect in a laboratory evaluation of temperature factors was small. It is conceivable that someone drinking outside in the cold with an ankle bracelet fully exposed to the low air temperatures could reduce the alcohol signal.

Perceptions: Offenders and Vendors

A focus group with seven post-SCRAM™ offenders opined that it would be foolish to try to circumvent the device, and none offered methods or admitted to attempts to do so as their freedom from jail was a strong motivating factor in compliance. These participants were selected only on the basis of their availability and willingness and therefore cannot be thought of as representing average users. Despite discomfort complaints by most, the bracelet was generally regarded as a net good by those who thought there was no way to “beat it” and who subsequently reported no longer drinking. Their commentary suggested a clear parallel to the bogus pipeline paradigm well-known among behavioral researchers as a method for improving the accuracy of self-disclosure. People will be more truthful and more compliant when they fear the consequences if they are caught lying. With that as a context, SCRAM™ can likely still serve its intended purpose even though the version we studied is imperfectly accurate and relatively insensitive to low-end alcohol after a couple weeks of service.

Service providers we spoke with were all supportive of SCRAM™ and its ability to detect noncompliant offenders. Unlike the offenders in our focus group, service providers deal with the full range of offender types. Of course, as providers, they may have some personal stake in the product, so their views are not objective. Although they will have become aware of unsuccessful circumvention efforts, if there are offenders who found a way around SCRAM™, they would probably not know about it.

In informal conversations, AMS has said that about 20 percent of all who are forced onto SCRAM™ have to be removed from the program and confined to jail because they cannot find the self-discipline to not to drink. This is not a surprisingly high proportion considering that only about a third of those who enter alcohol treatment successfully sustain abstinence for 6 to 12 months after initiation.

In our interviews, we learned that judges differ in their reaction to drinking while on SCRAM. Some send offenders to jail right away, others do little. The biggest fear among the offenders we interviewed seemed to be a result of encounters with capricious or uninformed monitoring staff who did not understand the difference between external interfering substances spikes and real drinking, and who had the authority to send them back to jail.

Final Note and Recommendations

In general, the sensitivity and accuracy of these devices were poorer than we expected. But if they are not yet perfected, further product improvement is likely to get them closer. It may be impossible to ever expect the alcohol signal at the skin surface to be a precise estimate of BAC if Anderson and Hlastala (2006) are correct. Their model suggests that the stratum corneum, the outer most layer of skin, and other systemic factors importantly affect the measurable ethanol gas concentration near the skin. Individual differences or state differences within individuals in hydration, temperature, and other factors theoretically affect the transdermal alcohol signal strength greatly.

On the other hand, the monitoring of alcohol consumption does not depend on precise measurement of BAC; it depends on the ability of a technology to detect abstinence violations as measured by a signal in excess of some minimal amount, such as .02 g/dL. As a monitoring device for offenders, the transdermal concept is valid and the benefit appears evident despite the limitations of the actual equipment with false-negatives rates that are too high. Nonetheless, these devices warrant further development and further study.

AMS has reportedly redesigned the interior of its SCRAM™ devices to solve the water problem by reducing the potential spaces where water can accumulate inside the housing. It is likely that the accuracy problem, like the sensitivity problem, is due to the water accumulation. The problem of declining accuracy over time in this regard is a very significant concern. AMS needs to demonstrate that equipment in service beyond 2 weeks has adequate sensitivity.

It is likely that female complaints about size and fit of the bracelet are valid and worthy of attention by AMS staff. Miniaturization and improved accuracy will only happen if the industry achieves the kind of success that warrants further investment in research and engineering.

On that point, the communication of the SCRAM™ bracelet with its remote server, along with data retrieval and reporting technology, is exceptionally innovative. The sensor technology that is at the heart of the Giner device also appears to be exceptional.

This report, or any like it, that concludes transdermal alcohol sensing has many more benefits than problems is unlikely to spark a rush to transdermal sensing as a monitoring remedy for alcohol abusive offenders; average daily cost is still a barrier, particularly with DWI offenders. The typical daily cost of leasing SCRAM™ is around \$12. This compares very well with regular house arrest, which runs about \$12/day, but quite poorly with alcohol ignition interlock devices that usually cost about \$2.50/day. Accordingly, the use of transdermal sensing is probably going to be less used with routine DWI offenders than for multi-problem alcohol abusers; people who need to be entirely prevented from drinking, not just drinking-and-driving, are the natural market for SCRAM™. Transdermal technology could make a very important contribution to the societal problem of low penetration of alcohol ignition interlocks among DWI offenders if it were an alternative to the interlock rather than simple license suspension, as is now most typical.

References

- Anderson, J. C., & Hlastala, M.P. (2006). The kinetics of transdermal ethanol exchange. *Journal of Applied Physiology*, 100(2), 649-655.
- Bjerre, L. (2003). An evaluation of the Swedish alcohol ignition interlock program. *Traffic Injury Prevention* 4(2), 98-104.
- Bland, J.M., & Altman, D.G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, February 8, 1986, 307-310.
- Bock, S. (2003). Michigan DOC runs BETA test of new remote transdermal alcohol monitoring system. *The Journal of Offender Monitoring*, 16(1), 9-10, and 13, 19-20.
- Brown, C.D., & Ridder, T.D. (2005). Framework for multivariate selectivity analysis, Part I: Theoretical and practical merits. *Applied Spectroscopy*, 59(6), 787-813.
- Gilg, T., Buchholtz, U., & Huth, O. (2000). CDT (carbohydrate deficient transferrin) and other alcohol markers in the MPA (medical and psychological assessment) of alcohol offenders regaining driving licenses – Results of an empirical study in Germany. In H. Laurell & F. Schlyter (Eds.). *Proceedings of the 15th International Conference on Alcohol, Drugs and Traffic Safety* (Vol. 4, pp. 1183-1184). Stockholm, Sweden: ICADTS.
- Gjerde, H., & Morland, J. (1987). Concentrations of carbohydrate-deficient transferrin in dialysed plasma from drunken drivers. *Alcohol and Alcoholism*, 22(3), 271-276.
- Gjerde, H., Sakshaug, J., & Morland, J. (1986). Heavy drinking among Norwegian male drunken drivers: A study of gamma- glutamyltransferase. *Alcoholism: Clinical and Experimental Research*, 10(2), 209-212.
- Greenfield, T.K., Tujague, J., Bond, J., & Kerr, W. (2005). *A pilot study of self-report alcohol consumption: Diary versus transdermal alcohol sensor accuracy*. Paper presented at the Research Society of Alcoholism Meeting, Research Society of Alcoholism, Santa Barbara, CA.
- Jacobi, U., Gautier, J., Sterry, W., & Lademann, J. (2005). Gender-related differences in the physiology of the stratum corneum. *Dermatology*, 211(4), 312-317.
- McCartt, A.T., Geary, L.L., & Nissen, W.J. (2002, September 2002). *Observational study of the extent of driving while suspended for alcohol-impaired driving* (DOT HS 809 491). Washington, DC: U.S. Department of Transportation, National Highway Safety Administration.
- Marques, P., Bjerre, B., Dussault, C., Voas, R.B., Beirness, D.J., Marples, I.R., & Rauch, W.R. (2001). *Alcohol ignition interlock devices. I: Position paper*. Oosterhout, Netherlands (Online: <http://www.icadts.org/reports/AlcoholInterlockReport.pdf>. Accessed Nov. 16, 2004): International Council on Alcohol, Drugs and Traffic Safety (ICADTS).
- Phillips, M., Greenberg, J., & Andrzejewski, J. (1995). Evaluation of the Alcopatch, a transdermal dosimeter for monitoring alcohol consumption. *Alcoholism: Clinical and Experimental Research*, 19(6), 1547-1549.
- Ridder, T.D., Brown, C.D., & Ver Steeg, B.J. (2005). Framework for multivariate selectivity analysis, Part II: Experimental applications. *Applied Spectroscopy*, 59(6), 804-815.

- Ridder, T.D., Hendee, S.P., & Brown, C.D. (2005). Noninvasive alcohol testing using diffuse reflectance near-infrared spectroscopy. *Applied Spectroscopy*, 59(2), 181-189.
- Roizen, J., Lichtor, L., & Lane, B. (1990). A "band-aid" to detect alcohol levels in blood. *Anesthesiology*, 73, A510.
- Sakai, J.T., Mikulich-Gilbertson, S.K., Long, R.J., & Crowley, T.J. (2006). Validity of transdermal alcohol monitoring: Fixed and self-regulated dosing. *Alcoholism: Clinical and Experimental Research*, 30(1), 26-33.
- Swift, R. (2000). Transdermal alcohol measurement for estimation of blood alcohol concentration. *Alcoholism, Clinical and Experimental Research*, 24(4), 422-423.
- Swift, R.M., Martin, C.S., Swette, L., LaConti, A., & Kackley, N. (1992). Studies on a wearable, electronic, transdermal alcohol sensor. *Alcoholism: Clinical and Experimental Research*, 16(4), 721-725.
- Swift, R.M. (2003). Direct measurement of alcohol and its metabolites. *Addiction*, 98(Suppl 2), 73-80.
- Widmark, E.M.P. (1932). Principles and applications of mediolegal alcohol determination. Berlin: Urban & Schwarzenberg.
- Zettl, J.R. (2002). *The determination of blood alcohol concentration by transdermal measurement: A white paper*. Alcohol Monitoring Systems, Inc. Retrieved August 16, 2004 online: <http://www.alcoholmonitoring.com/pdf/TransdermalWhitePaper.pdf>

APPENDICES

Appendix A – Literature Search Sources

Scientific Literature

At a minimum, the following Dialog databases will be examined:

- *Law enforcement databases:*
 - Criminal Justice Periodicals Index [Dialog file #171]
 - Gale Group Business A.R.T.S. SM [Dialog file #88]
 - Gale Group Legal Resource Index™ [Dialog file #150]
 - NCJRS [Dialog file #21]
 - PAIS International [Dialog file #49]
 - Periodical Abstracts PlusText™ [Dialog file #484]
- *Social science, psych, and medicine databases:*
 - American Medical Association Journals [Dialog file #442]
 - Dissertation Abstracts Online [Dialog file #35]
 - ERIC [Dialog file #1]
 - Gale Group Health & Wellness DatabaseSM [Dialog file #149]
 - General Science Abstracts/Fulltext [Dialog file #98]
 - MEDLINE® (1966-present) [Dialog file #155]
 - MEDLINE® (1990-present) [Dialog file #154]
 - NCJRS [Dialog file #21]
 - NTIS - National Technical Information Service [Dialog file #6]
 - New England Journal of Medicine [Dialog file #444]
 - Newsweek™ [Dialog file #482]
 - PAIS International [Dialog file #49]
 - PsycINFO® [Dialog file #11]
 - Social SciSearch® [Dialog file #7]
 - Wilson Humanities Abstracts Full Text [Dialog file #436]
- *Transportation and safety databases*
 - Occupational Safety and Health (NIOSHTIC®) [Dialog file #161]
 - Wilson Social Sciences Abstracts [Dialog file #142]
 - Wilson Applied Science & Technology Abstracts [Dialog file #99]

- *Government agencies*

National Institute on Alcohol Abuse and Alcoholism

National Highway Traffic Safety Administration

National Institute of Justice

Appendix B – AMS Letter Regarding Check Valve

March 24, 2005

Mr. Paul R. Marques, Ph.D.
Senior Research Scientist
Pacific Institute for Research and Evaluation
Calverton Office Park
11710 Beltsville Drive, Suite 300
Calverton, MD 20705



9135 S. Ridgeline Blvd.
Suite #190
Highlands Ranch, CO 80129
303.989.8900
Fax 303.791.4262
www.alcoholmonitoring.com

Dear Paul,

As you know, we have seen results during the PIRE research project not previously experienced in the field or in any of our controlled testing. In the last few weeks, we have conducted numerous rounds of laboratory testing in a moist alcohol environment, and we have finally been able to duplicate the results seen at PIRE.

In one particular unit, after several days of normal operation, sensor voltage abruptly dropped, and we no longer detected an alcohol standard. Upon further investigation, we determined that the check-valve at the unit's sample system exit had become stuck, was blocking the flow path, and dramatically reducing the volume of air flowing across the sensor's surface. After removing this valve, bracelet operation immediately returned to normal.

Examination of the check-valve revealed that the internal o-ring had swelled, preventing the valve from working properly. This particular o-ring material was part of a new pressure check-valve that was incorporated into the bracelets in Version 3.0.8—the same version as the first set of bracelets sent to PIRE. The manufacturer changed the material used in the o-rings when the change to the lower pressure check-valve was made.

Research into this failure mode and the o-ring material has led to the following conclusions:

- All 3 compounds (water, ethanol, and sulfuric acid) that may come in contact with the o-ring produce material swelling. This swelling can range from 1% to 10%, depending on the concentrations of each chemical and duration of exposure.
- Only 1% to 2% swelling is required to cause the check-valve to stick. A stuck valve will dramatically reduce the air flow and prevent any significant voltage from being read on the sensor.
- Even if the valve does not stick, every 1% swell in the o-ring causes more than a 10% reduction in flow. By the time 9% swell is achieved, the flow path is completely blocked.
- Stuck valves can be readily duplicated on the bench-top at the component level.

This failure mechanism is consistent with PIRE test result in the following ways:

- Swelling of the o-ring material causes a slow degradation of bracelet performance over time. As swelling increases, flow is reduced and sensor voltage is reduced.
- When a valve becomes stuck or as flow is significantly reduced, moisture accumulation is accelerated.
- The swelling of the o-ring and the resulting problem with the check-valve would result in low alcohol readings or missed alcohol detection.
- Upon return to AMS, the vast majority of the bracelets have returned to normal operation. This is because the valves have become unstuck, the sensors have dried out, or both.



The only purpose of the check-valve is to prevent gross damage to the alcohol sensor if the bracelet is submerged in water—something all clients are contractually restricted from doing, and something that, in our history, has caused very few bracelet failures.

We are taking the following actions to rectify the check-valve issue:

- Since the purpose of the check-valve is only to deal with an extreme water condition that is relatively uncommon, we are currently preparing bracelets with the valve removed.
- In addition, we have stopped shipping units with the current check-valve and o-ring, and have accelerated the procurement of new materials that do not have the check-valve.
- We have implemented a program with our customers that allows them to conduct a simple test on Version 3.0.8 bracelets that are in the field. If a bracelet does not pass this test, the unit will be shipped back to AMS and upgraded with the new components. Bracelets that continue to pass this brief field test will be modified during the unit's normal service cycle.

We currently plan to ship replacement units to PIRE the week of March 28th. Moving forward, we'd like to schedule a conference call at your earliest convenience in order to discuss the next steps for getting your testing cycle back on track.

Sincerely,



Mark Wojcik
Vice President, Products



9135 S. Ridgeline Blvd.
Suite #190
Highlands Ranch, CO 80129
303.989.8900
Fax 303.791.4262
www.alcoholmonitoring.com



Appendix C – Participant Agreements and Instructions

1. Informed Consent Form
2. Participant Agreement Form
3. Subject Do's and Don'ts Checklist

**INFORMED CONSENT FORM
FOR EVALUATION OF TRANSDERMAL ALCOHOL DEVICES**

The Pacific Institute for Research and Evaluation (PIRE) is conducting an evaluation of the effectiveness of 2 devices for measuring alcohol levels from the skin's surface. This research study is funded by the National Highway Traffic Safety Administration. The purpose of the study is to determine the accuracy levels of the two different devices. If you meet all the qualifications for participation, you are invited to take part in this study.

You are invited to participate only if you: are at least 21 years of age but under 35, drink alcohol on a regular basis, have never had an alcohol or drug related health or criminal problem, are not taking medication of any kind, are not pregnant, and are in good health. You will be asked to complete two short questionnaires regarding alcohol use before qualifying for the study. You will also be asked to provide a urine sample to screen for drug use. Females will be asked to take an early pregnancy test.

Participation in the research project is voluntary. If you are eligible to be in the study, you may receive up to \$800 for your time. Note that the IRS requires us to send you a 1099 tax form for any payment over \$600, so you will receive a 1099 form for tax reporting purposes if you complete the entire study and receive \$800.

You will be asked to wear two transdermal devices (one on your ankle and the other on your wrist) 24 hours a day for 28 consecutive days. During that four-week period, you will be asked to keep a daily diary or log of your drinking, and home breath tests. You will also need to come to the PIRE offices and spend up to 20 hours total across 5 occasions. Some of that time will be used to install and remove the devices as well as discuss the study tasks with you.

On 2 occasions when you come to the PIRE offices, you will participate in alcohol dosing lab studies that will involve your drinking vodka and juice in a brief period. During this time, your blood alcohol concentration (BAC) will be regularly monitored with breath tests so that your BAC does not exceed .08. To test the effect of different temperatures and conditions on the transdermal devices, you will also be exposed to a cold room (60 degrees F) and a hot room (90 degrees F) and asked to walk on a treadmill for approximately 10-15 minutes. Transportation from and to your home will be provided. You will not be allowed to leave the PIRE offices until your BAC level has decreased to at least .02, after which time you will be given a ride directly to your home. A certified paramedic will be monitoring your status throughout the study.

To protect your confidentiality, your data will be coded with a unique ID number. The key linking your name and ID number (along with the consent form) will be kept secure in a locked file cabinet separate from the data collection forms. Only senior staff will have the key. Your name will not be used when reporting the results of the study. No information traceable to you will be released to anyone.

There are some risks to you by participating in this research. You could be embarrassed by your behavior while under the influence of alcohol, although you

may stop participating at any point in the study. You may become sick as a result of drinking alcohol, although a paramedic will be on hand to assist and provide medical attention if needed. You will benefit by receiving \$100 for each of 4 weeks in the study and an additional \$400 upon final completion of the study for your participation. You will be contributing to understanding the results of using alcohol sensing transdermal devices.

If you have any questions about the research study, you can contact Dr. Paul Marques who is in charge of the study at the Pacific Institute for Research and Evaluation, 11710 Beltsville Drive, Suite 300, Calverton, MD 20705, Phone (301) 755-2700. If you have questions about your rights as a research subject, you can contact Elysia Oudemans, Manager of Research Integrity Compliance, at the same address and phone number.

I have read the consent form and agree to participate in the evaluation of transdermal devices.

Participant

Date

Witness

Date

PIRE TRANSDERMAL ALCOHOL STUDY PARTICIPANT AGREEMENT FORM

Participant Name _____ **Date** _____

I, _____, am participating in a study of transdermal technology for the Pacific Institute for Research and Evaluation (PIRE). I understand that for participating in this research study, I will be paid at a rate of \$100 per week plus and additional \$300 upon satisfactory completion of the full 28 days of the study, for a maximum payment of \$700. I agree to comply with all study requirements set forth in this Agreement and to strictly follow the instructions of PIRE research staff. I understand that any failure by me to comply with this Agreement or the instructions of PIRE research staff will be considered a violation of this agreement and may constitute grounds to be dropped from the study prior to becoming eligible for the full \$700 payment.

SCRAM™ EQUIPMENT

As a condition of my participation in the study, I agree to properly use the Secure Continuous Remote Alcohol Monitoring™ ("SCRAM™") equipment provided to me by PIRE research staff. In that regard, I will wear the SCRAM™ Bracelet on my ankle for the duration of the Program and will allow the SCRAM™ Modem to be connected to my home or office telephone or as agreed with research staff. I understand that the SCRAM™ Bracelet will, at pre-programmed intervals, test me for the presence of a positive blood alcohol concentration by the measurement of ethanol, which is being emitted as vapors through my skin. When the SCRAM™ Bracelet detects the presence of ethanol, it will record a positive reading and will later transmit a reading to the SCRAM™ Modem. The SCRAM™ Bracelet also contains systems designed to detect interference or tampering and will also transmit a tampering alert to the SCRAM™ Modem in the event of a tampering attempt.

I understand that I have a responsibility to do everything reasonable, within my power to protect the SCRAM™ equipment from damage. I understand that if PIRE does not return the equipment in good working condition, they will be charged for the repair or the replacement of the equipment as follows:

- Full replacement of the SCRAM™ Bracelet - \$1,400.00
- Full replacement of the SCRAM™ Modem - \$ 700.00
- Front strap replacement - \$ 75.00
- Back strap replacement - \$ 175.00
- Back comfort strap replacement - \$ 75.00
- Clip and battery replacement - \$25.00

I agree to allow authorized personnel to inspect and maintain the SCRAM™ Bracelet and SCRAM™ Modem.

While participating in the study, I agree to wear a non-removable SCRAM™ Bracelet that will be attached by PIRE research staff. I agree not to remove the device during the course of the study. I agree not to tamper with the SCRAM™ bracelet, or place any obstruction material between the SCRAM™ bracelet and my leg, except where directed to do so as part of PIRE's investigation of circumvention of the SCRAM™ device. Under no circumstances will an attempt to circumvent the device involve acts that might damage the SCRAM™ equipment Only in a life-threatening

emergency or with the prior permission of PIRE research staff will I remove the SCRAM™ Bracelet by cutting the front strap where indicated by the words 'Cut Here'.

I agree to maintain an analog telephone line and electrical service in my residence at my own expense, unless some other arrangement has been agreed to by PIRE staff.

I understand that PIRE staff and I will agree on a time everyday when I will be in range of the SCRAM™ modem. I agree to be physically in range of my SCRAM™ Modem for 15 minutes prior to the designated reporting time. I will not leave SCRAM™ Modem range while the green light is blinking. SCRAM™ Modem range is within the same room as the SCRAM™ Modem or within 30 feet of the SCRAM™ Modem.

If I experience problems with the SCRAM™ Bracelet or SCRAM™ Modem, or if I lose electrical power at my residence, I agree to call PIRE research staff immediately.

If I am unable to reach PIRE staff immediately, I agree to call and leave a message on their answering machine including my name, the date, the time, and the nature of my problem. If there has been a power problem, I agree that I will call and leave another message when the power is restored. I also agree to notify PIRE research staff of any problems with my telephone service as soon as I am able to do so.

I understand that as a participant in the study, I am to refrain from the following practices which are potentially harmful to the SCRAM™ equipment:

Tampering – I understand that attempts to tamper with SCRAM™ devices or alter their readings will be considered a violation of this Agreement.

Swimming & Bathing – I understand that I am not to submerge the SCRAM™ Bracelet in water. Showers are the only permitted bathing method.

Personal Hygiene - I agree, that when bathing, I will thoroughly rinse with clean water and dry underneath the SCRAM™ Bracelet. I understand that failure to rinse away all soap may result in a mild skin rash.

I agree to return all SCRAM™-related equipment in my possession, including the SCRAM™ bracelet and SCRAM™ modem, to PIRE promptly upon completion of my role in the study.

WARNING: The SCRAM™ Bracelet contains caustic liquid and a lithium battery. Should the plastic casing become cracked or otherwise visibly damaged or there is a burning sensation on my skin or other apparent health risk, I will remove the SCRAM™ Bracelet immediately by cutting the front bracelet strap where it says "Cut Here." If this occurs, I will immediately notify PIRE staff of this problem.

I acknowledge receipt of SCRAM™ Bracelet number _____ and SCRAM™ Modem number _____.

WrisTAS™ EQUIPMENT

The WrisTAS™ (wrist-worn transdermal alcohol sensor) has different care requirements than the SCRAM™ bracelet. Unlike SCRAM™, the WrisTAS™ should not get wet.

I understand that I am to wear the WrisTAS™ sensor for the duration of the study, except that I will remove it for the time it takes to take a shower. I will put it back on when I finish showering. I understand that the WrisTAS™ device is sensitive to moisture and physical shocks. While I am wearing the WrisTAS™ I will completely refrain from activities that that would cause the device to be submerged (e.g., swimming and bathing) and avoid activities that would subject the device to dampness. I will avoid activities that subject the device to shocks (e.g., contact sports, heavy manufacturing and construction work). If for some reason I am forced to remove the device for a period of more than 30 minutes I will put it into an airtight zip-lock bag. I will not allow anyone else to wear the WrisTAS™. I understand that the WrisTAS™ device is extremely expensive (\$2,500) and will endeavor to protect it from damage and theft while it is in my possession. I will return the WrisTAS™ device to PIRE when my participation in the study is over.

I acknowledge receipt of WrisTAS™ Bracelet number _____.

INTOXILYZER BREATH TESTER

I understand that I am to keep the Intoxilyzer 400 with me when I will be drinking and use it when I have two or more drinks to take breath measurements. I will start taking breath measurements approximately a half hour after I have my first drink and continue taking measurements at least once an hour until my BAC returns to .000, or until I go to bed for the night (whichever comes first). I will record BAC measurements and enter them into my drinking log. It is important to obtain accurate measurements that do not include the effects of alcohol in the mouth. In order to do this I will rinse my mouth with water and/or wait 15 minutes after drinking alcohol before taking a breath measurement.

I will bring the Intoxilyzer with me when I come in for dosing studies so that the data can be downloaded from the Intoxilyzer. These data will be used as confirmation of the breath measurements in your log. To prevent creating confusing breath data files, I will refrain from using the breath testers to test anyone but myself.

If I believe that the Intoxilyzer 400 is not functioning properly I will contact PIRE to discuss whether it is malfunctioning and how to address the problem. I understand that the Intoxilyzer 400 is quite expensive (\$700) and will endeavor to protect it from damage and theft while it is in my possession. [NOTE: PIRE experience with use of breath testers in the field has shown that breath testers are often the targets of theft attempts.] I will return the Intoxilyzer 400 to PIRE when my participation in the study is over.

I acknowledge receipt of Intoxilyzer 400 number _____.

DRINKING LOG

During my participation in the study I will keep a log of my eating and drinking activity and BAC readings for days when I drink alcohol. I will send a report to PIRE everyday by e-mail. If I am reporting on a drinking day the report will contain my drinking log for that day. If I am reporting on a non-drinking day the report will simply state that I did not drink that day. I will include the day and date to which the drinking log applies. I will send each drinking log as a separate e-mail message.

Entries for eating will contain:

- The time at which the food was consumed
- A brief description of the type of food consumed (e.g., “hamburger and fries,” or “bowl of cereal”)
- Whether you consider it a “snack,” a “small meal,” a “medium meal” or “large meal”

Entries for drinking alcoholic beverages will contain:

- The time during which drinking took place (e.g., 4 p.m. – 7 p.m.)
- The number of drinks consumed, where a “drink” is one 12 oz. beer, one 4 oz. glass of wine or a liquor drink containing one shot of liquor. Larger amounts of beer should be counted based on their size (e.g., 16 oz. pints of beer would be counted as 1.3 drinks, 22 oz. “schooners” of beer would count as 2 beers). Liquor served as “doubles” would count as 2 drinks.

Entries for breath tester readings would contain:

- The time at which the reading was taken; and
- The breath test reading.

A sample drinking log is attached as an Appendix to this agreement form.

You will need to keep a note pad to record your breath test readings – it is highly unlikely that you will be able to remember your readings until you fill out the daily log. As long as you are recording your breath test readings you will probably want to record eating and alcoholic drinks on the same note pad. As soon as possible, at the end of the day or the next morning, transfer your drinking log information to an e-mail and mail it to _____@pire.org.

NATURE OF DRINKING

When study participants are at PIRE for the dosing portion of this study, PIRE will control the amount of alcohol being consumed by participants. As mentioned in the consent for signed by participants, subjects may discontinue drinking during the dosing study if they feel that drinking will make them sick or feel uncomfortable.

When study participants are not participating in the dosing study, but are drinking as they normally would in the course of their normal lives, participants will drink only as much as they decide they want to drink. Participation in this study should not be considered encouragement by PIRE to drink any more than a participant would normally choose to drink. PIRE will not be held liable for adverse consequences resulting from a participant’s decision to drink.

PIRE will provide a letter to participants that can be used to help explain the nature of the devices to officials needing information about them. An example would be security personnel in airports. This letter can not be construed by participants as permission to drink, or have alcohol in their system, under circumstances where drinking or having alcohol in their system is unlawful. For example, a study participant who is stopped under suspicion of driving under the influence can not use the letter to avoid possible legal consequences.

Participants must understand that using the breath test equipment to check whether they are under the “legal limit” (i.e., the illegal-per-se level of .08 g/dL) for driving under the influence is not a

guarantee that they can drive safely or be free from possible charges of driving under the influence. First, these devices are not guaranteed to be accurate for measuring blood alcohol levels for determining who is subject to prosecution for drinking driving. Secondly, drivers with BACs below the State's "legal limit" can nevertheless be convicted of driving under the influence if it can be shown that the alcohol in their system has impaired their driving. Lastly, research shows that driving ability is negatively affected at blood alcohol levels below .04 g/dL BAC.

I acknowledge that I have received a copy of this Agreement and that it has been explained to me before signing.

(PARTICIPANT)

(DATE)

(PIRE AGENT)

(DATE)

Sample Drinking Log

7:00AM - Bowl of cereal – small meal
12:30 PM - 3 pieces of pizza – medium meal
12:30 PM - one 12 oz beer with lunch
4:00PM - 7:00PM – four 12 oz beers
4:32 PM - BAC = .030
5:29 PM - BAC = .061
6:31 PM - BAC = .045
7:30 PM - BAC = .032
7:30 PM – Steak, salad, baked potato – large meal
8:30 PM - BAC = .000
11:00 PM - two shots of tequila
11:30 PM - BAC = .036
12:30 AM - BAC = .012
12:45 AM - bed

Do's and Don'ts Checklist

Do:

- Take proper care to protect all equipment from damage and theft
- Record a drinking log and email it to alcoholstudy@pire.org everyday
- Record breath tests every hour when drinking
- Rinse mouth and/or wait 15 minutes before taking a breath test
- Take off the WrisTAS™ before showering and keep WrisTAS™ in ziplock bag when not wearing it
- Be near the SCRAM™ modem at the designated download time

Don't:

- Get WrisTAS™ wet
- Tamper with equipment
- Drink more than you normally would
- Make drinking-driving decisions based on breath test results
- Use the breath tester to test other people
- Put WrisTAS™ on other people.
- Eat heavy meals before dosing
- Drive to dosing appointments

In case of problems or questions call:

Scott McKnight – 301-755-2735

Paul Marques – 301-755-2723

Eileen Taylor – 301-755-2719

Appendix D – Wearability of the SCRAM™ and WrisTAS™ Devices

One of the evaluation tasks in this transdermal alcohol device study was an informal debriefing of subjects on the question of wearability. That is, with four continuous weeks of wearing the devices how did they interfere with daily activities? The questions posed addressed issues of comfort, limitations, as well as expected and unexpected difficulties in complying with the conditions of the research agreement. Wear issues were very different for the two devices and answers to these questions are summarized separately. Because the debriefing was informal with no particular structure, so too this summary will be mostly unstructured.

SCRAM™ Ankle Bracelet

As part of the orientation we informed subjects that according to the manufacturer most people find wearing the bracelet to be most difficult during the first 2 to 3 days. And whether through this suggestion or their own experience, this is very much what the subjects reported back to us at debriefing. The following is from a sequence of debriefing reports of some male subjects:

- It seemed too tight for 2 days then seemed to loosen
- Agreed with our comment that the first 3 days were the hardest but then he accepted it and found it to be relatively non-problematic
- Commented that the first week was tough with SCRAM™ but then more acceptable
- For 48 hours he found it terribly annoying and then almost overnight it seemed to become virtually unnoticeable
- First three days were annoying while trying to sleep.
- First four days were uncomfortable but after that it was no big deal.

Overall it seems that more females reported wear problems than men. Many thought that the physical size of it was more suited to a male than female and there may have been some additional interaction between the device and changes in body water retention. The following is a selection of commentary from female subjects:

- She found the first two weeks to be very uncomfortable particularly because she is physically very active and she discovered the tendon on the back of her ankle throbbed quite a lot when she exercised with the device on.
- Biggest issues were showering and shaving legs and was unable to play soccer since when kicking the ball there was discomfort at device contact with leg.
- Running was a pain but not a stopper. Found that depending on where she was on the puffy/not-puffy continuum the ankle bracelet would slide around somewhat and the more it slid around, the more uncomfortable it was.
- Subject reported bruising on her ankle and device left a clear impression of the SCRAM™ faceplate that lasted for months after the study ended.

WrisTAS™ Wrist Device

The WrisTAS™ device is a prototype and not available for the criminal justice market so accordingly the version tested (version 5) is not designed with anti-circumvention features. Accordingly, subjects would have to remove the wrist device prior to showering and then reattach it after the shower. Since the removal and reattachment of the device is completely under control of the research subject, tightness was not a problem. However several subjects did report rash and itch related problems in long term wear of the WrisTAS™. The following is sampling of characteristic comments...

- WristTAS led to chafing and found comfort in switching it from wrist to wrist on occasion. He also had the need to change pads regularly due to sweating.
- WrisTAS™ pad became sticky and stuck to his skin.
- The wrist device took a long time to become accustomed to since he did not regularly wear wrist watches.
- He reported that the wrist device got stinky after a while despite a swap out of wrist devices at one point.
- WrisTAS™ gave him skin rashes. He would move it from arm to arm and to different places on his arm. This resulted in a number of pink dots the size of the pad on his arms.
- Comfort was no problem.
- No skin problems. He left it on his left arm the whole time.

The general experience of subjects seemed to depend somewhat on individual differences in skin sensitivity. We found that it was a good idea both hygienically and pragmatically to change the pads regularly so that they do not sweat out or distort in a way that blocks the vapor flow from skin to sensor. Probably it is best to change them at least weekly or as needed.

In summary, there are individual differences in the ability of subjects to tolerate either the SCRAM™ or WrisTAS™ devices but no subjects found the problems to be great enough to warrant dropping out of the study. All subjects who began the study completed the study. Several commented that the bonus money for sticking with the study all the way through dispelled any idea of dropping out even though when they found the devices annoying. The largest complaints centered on the manner in which it interfered with participation in exercise, particularly swimming, but also to a minor extent soccer and running. Many subjects commented that they did not want to be seen as offenders and were a little embarrassed by the devices, often wearing long pants when they would have ordinarily worn shorts.

Appendix E – SCRAM™ Focus Group and Users Feedback

This appendix is a report on a Focus Group (FG) conducted with people who had worn the SCRAM™ device as ordered by courts in response to various alcohol offenses. The following caveats apply to this and other focus groups. FGs are intended to give insight into beliefs and experiences of participants, they constitute a small sample of the overall population of interest, and the results do not provide information that can be generalized to the larger population. Participants are encouraged to talk honestly and openly, but there is no guarantee that comments are not filtered to please the facilitator. A primary purpose of this FG was to learn about problems the offender participants had with the SCRAM™ system. Participants may have taken this opportunity to complain to someone who they believe will have the ear of someone in authority. Results of such a discussion will necessarily be somewhat negative.

Focus Group

A Focus Group (FG) meeting was held in Dallas, Texas, with seven people who had completed sentences to wear the SCRAM™ device. Following Federal human subjects protections guidelines, FG subjects were recruited by PIRE with contacts provided by the Dallas-area SCRAM™ provider. All subjects volunteered freely and were given a \$100 money order each to participate. Subjects were both male and female. The ages ranged from 21 to 45. Most of the subjects were at the younger end of that age range. Most, but not all, were sentenced to SCRAM™ as a result of an alcohol-related driving offense. Subjects had worn their bracelets for periods ranging from 2.5 months to almost a year. The meeting was held in a hotel meeting room and lasted about 2 hours.

Subjects spoke about problems that they had with the SCRAM™ system, benefits of SCRAM™, and other issues.

Problems

Physical Discomfort

Discomfort caused by the bracelet was a common complaint. All subjects reported problems sleeping with the bracelet on during the first few days of wear, and about half said that they had problems sleeping during the entire time they wore it. Subjects also complained of skin irritation and a few reported severe bruising. One offender complained of being cut by the bracelet so that he was bleeding, and that the bracelet prevented the cut from healing. The problem was eventually solved by having the bracelet switched to the other leg. In some cases the bracelet would loosen and bounce on the ankle bone. This was solved by wearing socks or a head band around the leg, though this was done at the risk of potential tamper alerts should the bracelet work its way down over the cloth. Another common complaint was related to being overheated due to wearing long pants in hot conditions, rather than wearing shorts that showed the bracelet.

Embarrassment

Most subjects preferred that others not know that they were wearing the bracelet. Situations in which the bracelet was visible or audible caused them embarrassment. Offenders reported being embarrassed when others heard sounds coming from the bracelet at work, in school, in church and

during a wedding ceremony. Offenders reported a range of sounds from the bracelet that may have been interpreted by others as a mobile phone vibrating, someone moaning or passing gas. Offenders discussed embarrassment as part of what was called the “shame factor.” In general, offenders accepted the notion that “shame” was an intended part of the punishment and that this was not particularly unfair. Offenders accepted that, while they were not particularly happy at being sentenced to SCRAM™, it was something that happened due to their bad behavior.

Equipment-Related Problems

Occasionally there would be equipment failures that affected offenders. This resulted in frequent phone calls to the offenders in an attempt to determine what was going on, and frequent orders to come into the provider’s office to have the bracelet checked. Offenders objected to being called into the office to address equipment issues as often as once every week or two at inconvenient times. In some cases, SCRAM™ provider staff members traveled to the offender to perform maintenance. The most commonly reported equipment problem was related to batteries that failed prematurely. However, most batteries last for a month or more. One subject reported having two bracelets fail completely, however this seemed to be due primarily to abuse while participating in sports (e.g., sliding feet-first into base while wearing the bracelet). One subject reported having long pants torn at the bottom by the bracelet.

There were also complaints of modems malfunctioning. This was most often a function of the phone system, not the modem. Sometimes modems would pick up during the day (not at the scheduled communication time) and interrupt phone calls and Internet connections. Internet problems mostly involved dial-up connection but included offenders with DSL connections using micro-filters as required. One subject dealt with this problem by leaving the modem unplugged and forcing a download once a day by powering up the modem. The fact that the modem requires a land line caused inconvenience to some offenders who had only mobile phones. One made arrangements to download daily at a neighbor’s house. Others went to the provider’s office to download data.

Costs

Offenders reported having problems paying to use SCRAM™. One had opted to use SCRAM™ rather than an interlock before understanding the difference in price. By the time he realized the cost of SCRAM™ it was too late to go back to the interlock option. One person used funds from a student loan to pay for SCRAM™. Others felt fortunate that they could afford SCRAM™ with the help of their families. In addition to the problems caused by the total cost, there were problems related to inflexibility in payment – when payments were due the client needed to be able to pay or drop out of the program and go to jail. Offenders pointed out that, while they managed to pay for SCRAM™, many people could not, therefore the benefits of SCRAM™ to society (e.g., getting non-violent offenders out of jail) might be increased if SCRAM™ could be made more affordable.

Possible Misinterpretation of Data

A few subjects said that they had problems when they were forced to defend themselves against people who weren’t able to correctly interpret SCRAM™ data. In one case an offender reported being called into the office by a provider staff member and accused of drinking based on a misinterpretation of an interferrant spike. A supervisor was eventually contacted who correctly read the chart, but the offender said that it was extremely distressing, in the meantime, feeling that he may be sent back to jail when innocent of drinking. FG participants reported that the provider was sending charts of SCRAM™ data to parole officers who were not trained to read them. Parole officers

were said to have been questioning offenders in an accusing manner about TAC data that did not represent confirmed drinking events. For this reason the provider eventually stopped sending charts with reports.

The SCRAM™ system provides alerts when data indicate that offenders may have attempted to tamper with the bracelet. Not all of these are actual tamper attempts. These are generally investigated by the provider and judged to be confirmed or unconfirmed tampers. Offenders reported that for a while all potential tampers, including unconfirmed tampers, were being reported to parole officers as “tampers.” This caused problems when offenders were forced to explain that they had not actually attempted to tamper with the device.

Offenders were very bothered at the prospect of having their fate decided by people who might misinterpret the SCRAM™ data, and by the fact that they had no recourse if they were pronounced guilty of drinking or tampering. No participants were aware of anyone who had been taken off the program due to misinterpreted data.

Security Checkpoints

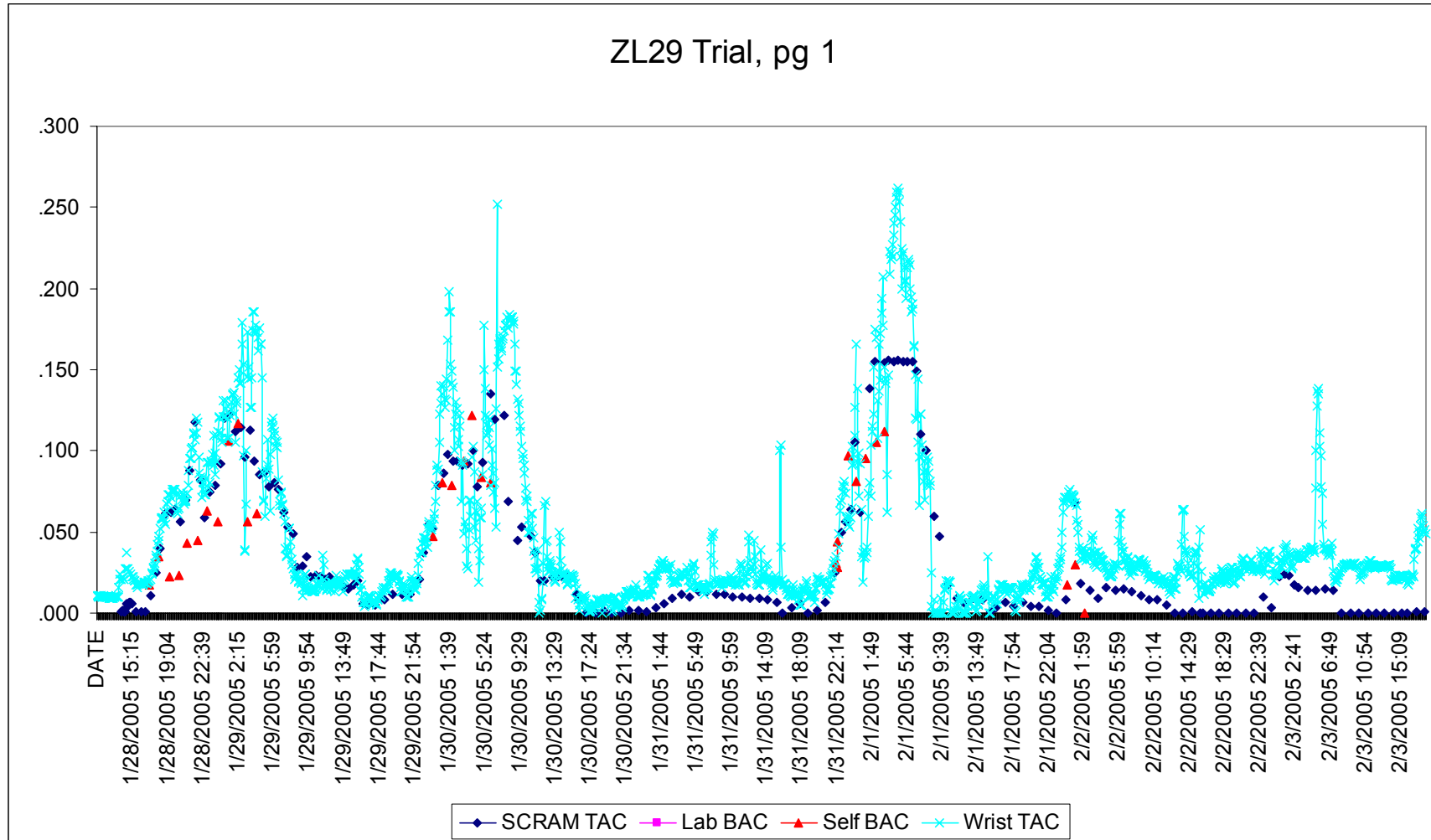
There was some question as to whether it would be difficult for SCRAM™ clients to get through security checkpoints wearing the metal bracelets on their leg. Most offenders did have experience with checkpoints, primarily when entering court buildings. Offenders reported little inconvenience from going through checkpoints, including those who went through security in airports. Aside from determining that the bracelet contained no explosives, security personnel had little interest in the bracelet aside from some curiosity as to what it was.

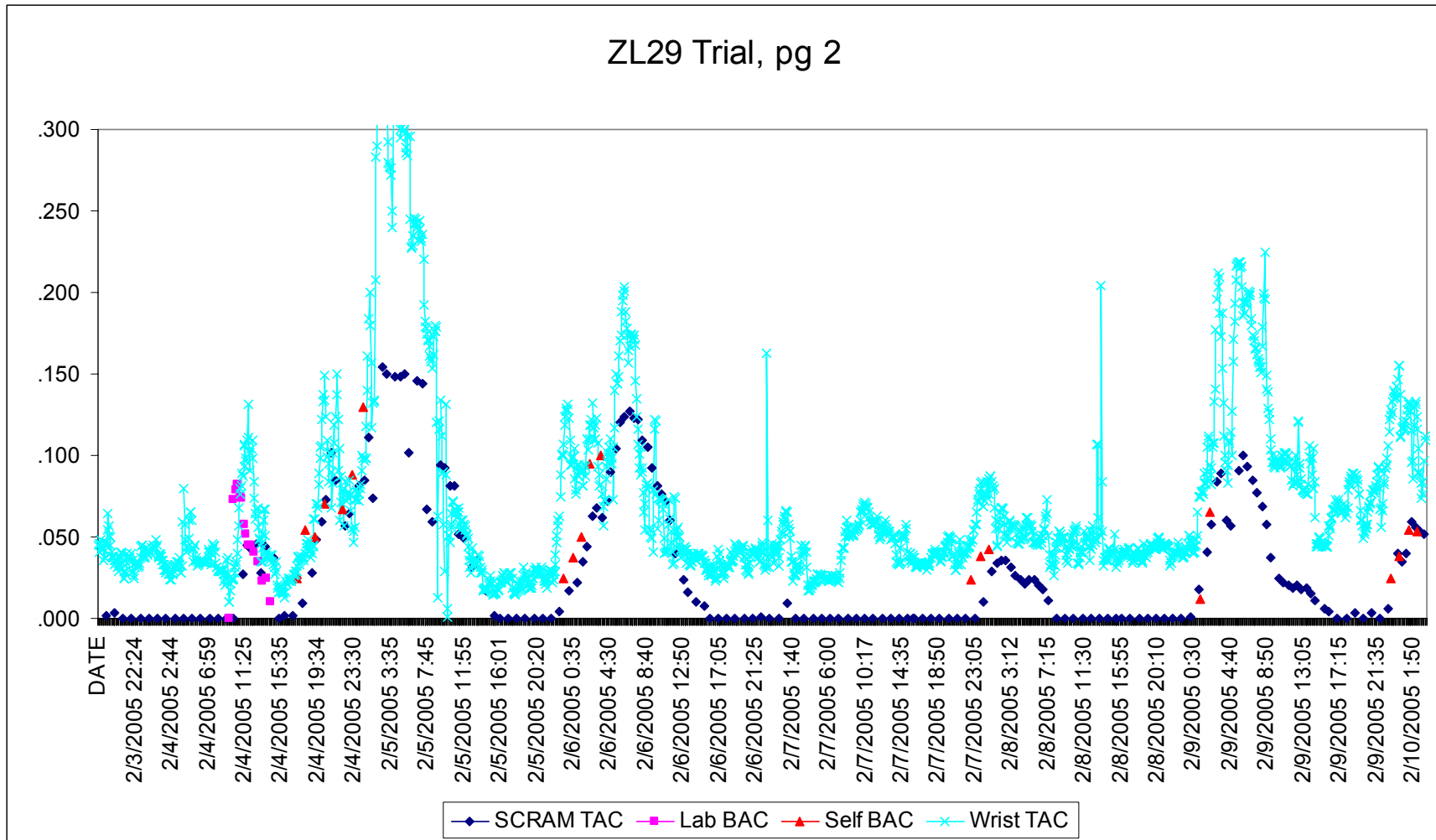
Benefits

Enforcement of Abstinence

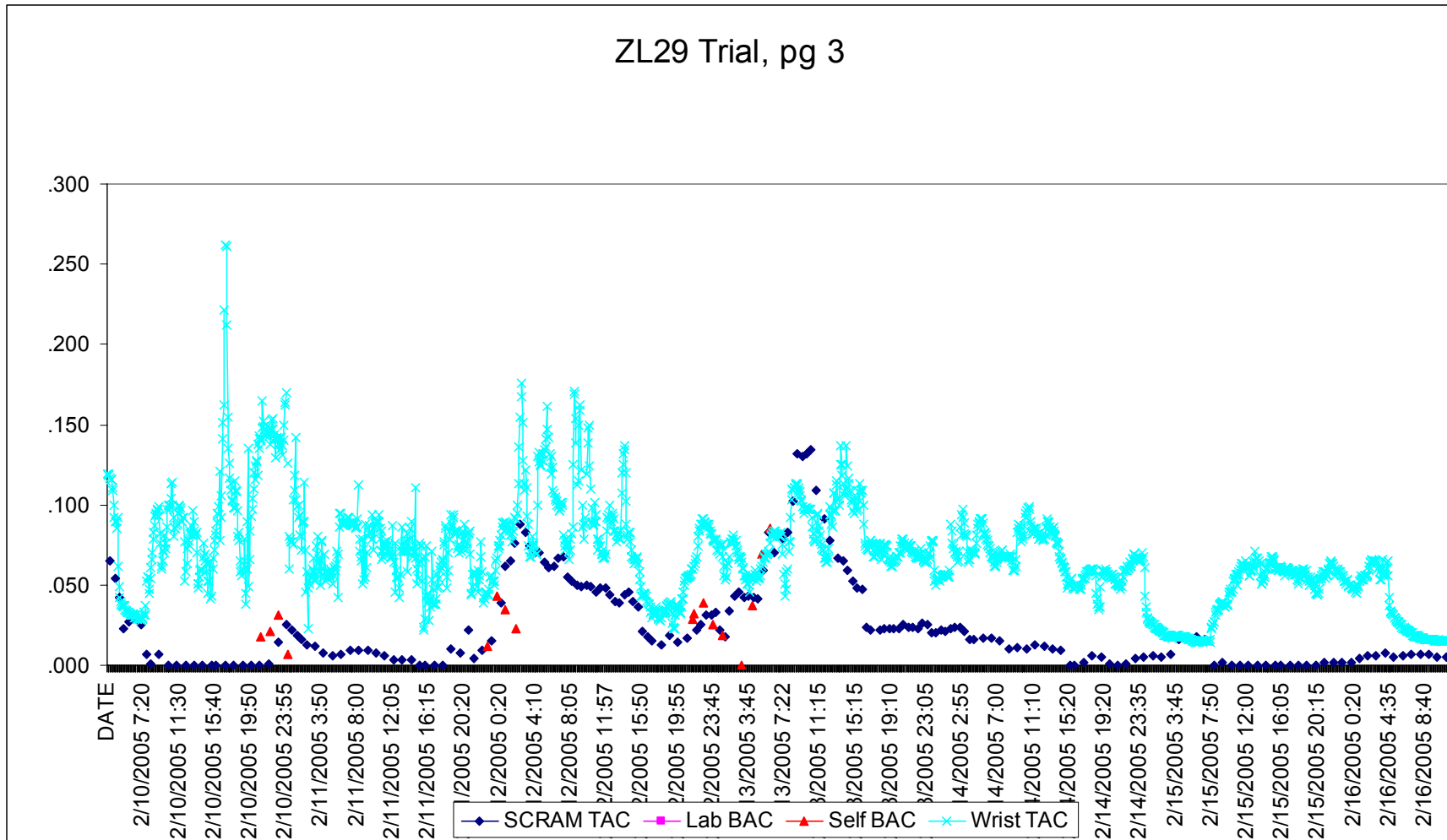
All focus group members had been through alcohol treatment, claimed to be abstinent and expressed a desire to remain so. Several had stories of how much they had been drinking and using drugs, and saw this as a problem that needed to end. Offenders also recognized that abstinence prevented impaired driving, which helped keep them safe and out of the legal system. Simply being sentenced to SCRAM™ was beneficial because it meant quitting, rather than managing drinking. Several expressed the feeling that SCRAM™ helped them to quit drinking by “keeping them honest” – any attempts to cheat would be detected. Another benefit was that complying with SCRAM™ required a definite end to drinking on a certain date rather than allowing them to put off quitting, or trying to cut down slowly, which may not have worked.

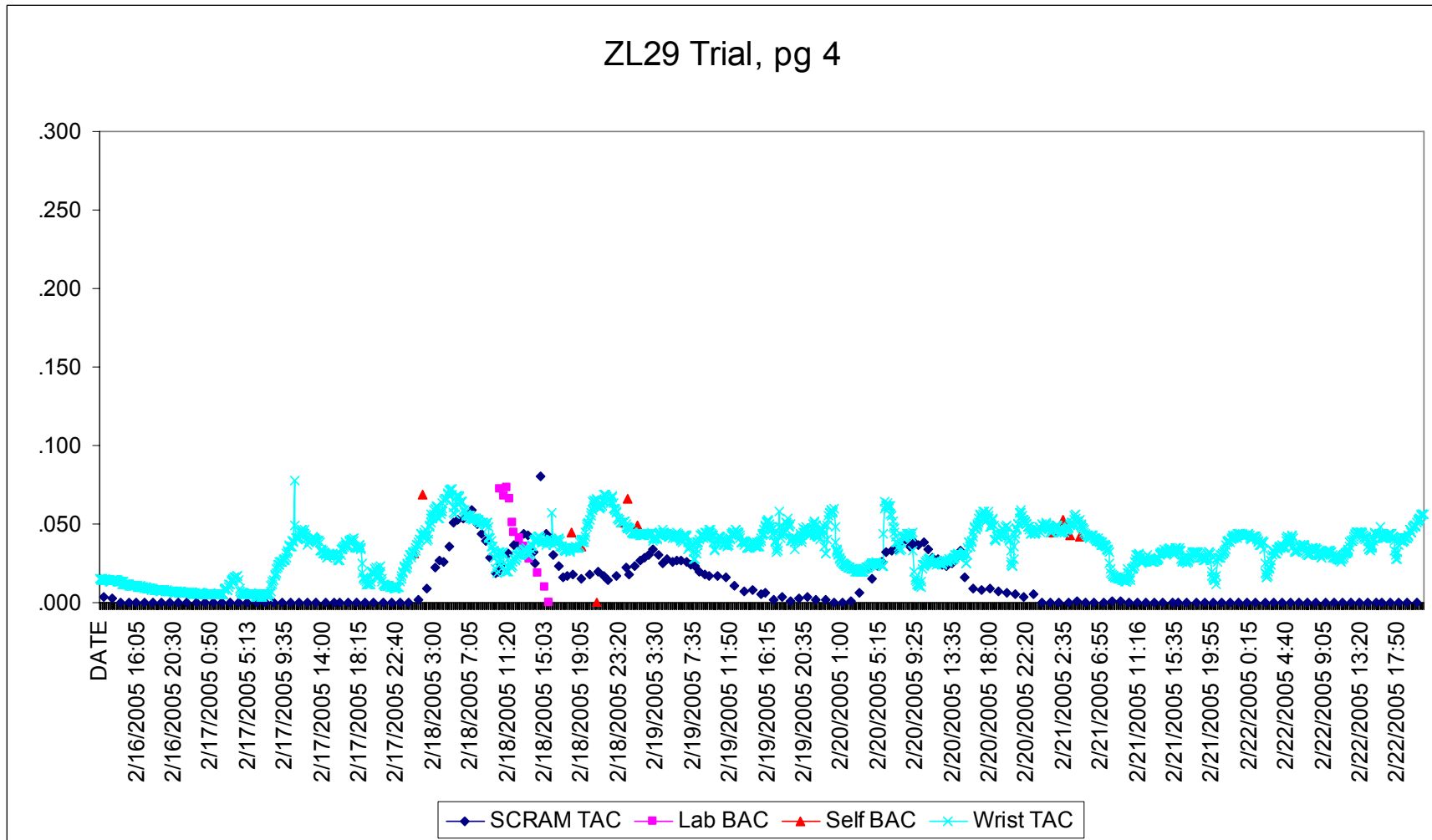
Appendix F – One Subject's Four-Week Chart

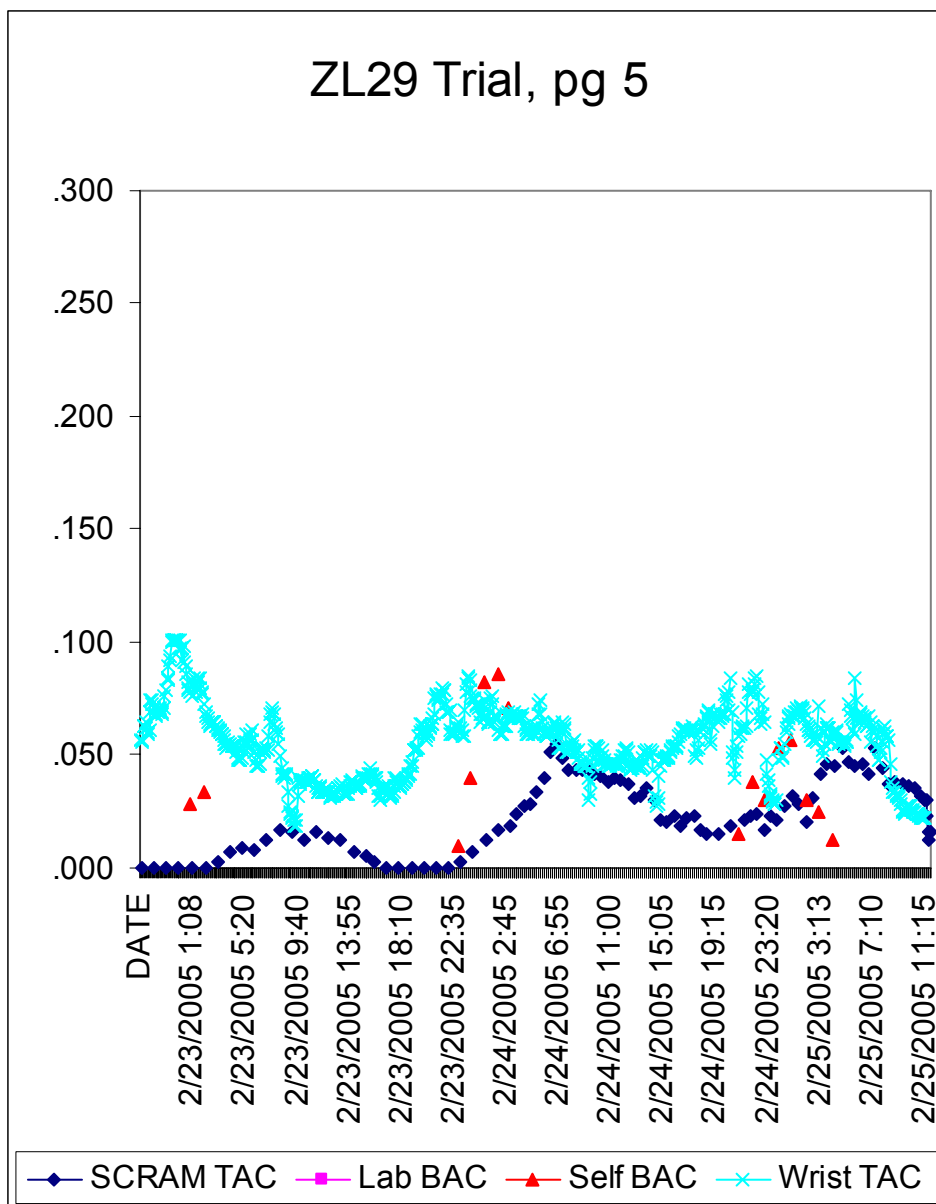




ZL29 Trial, pg 3

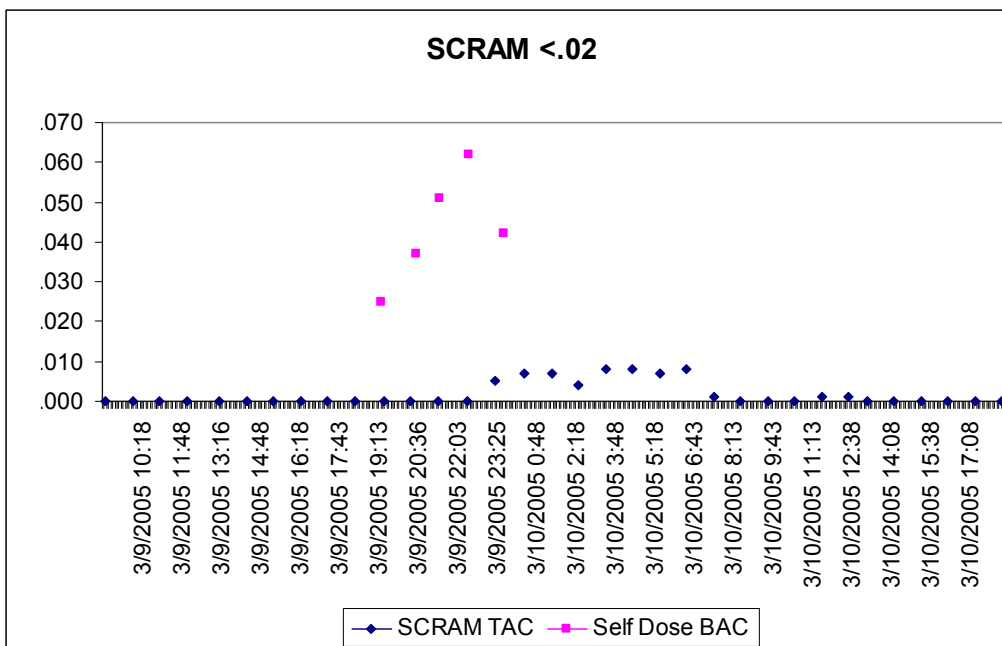
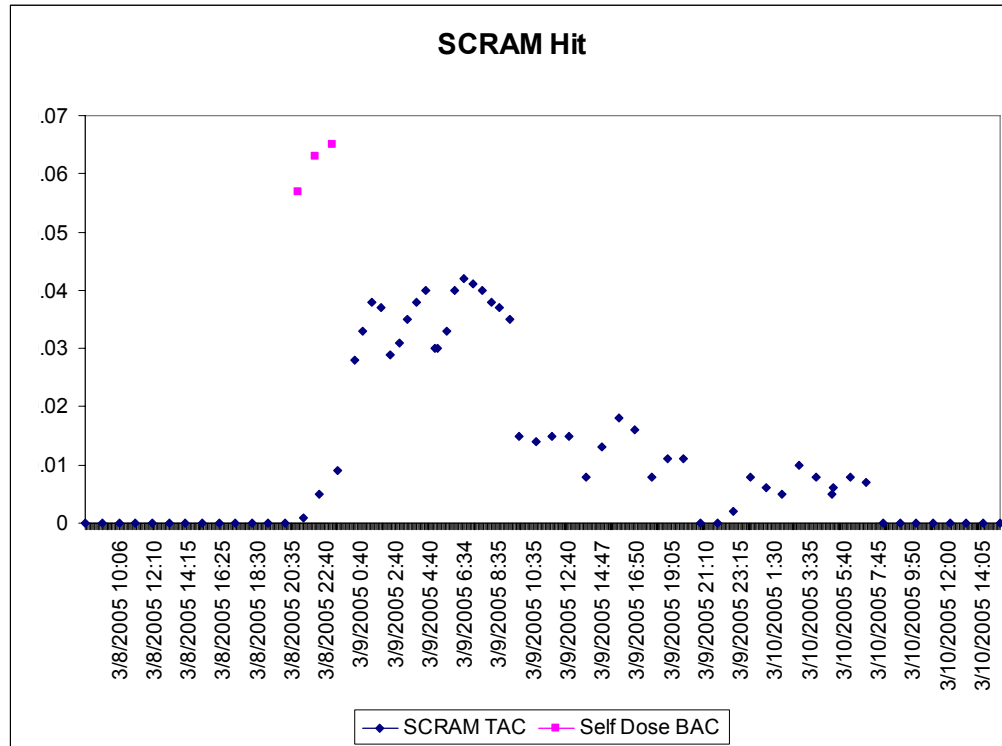


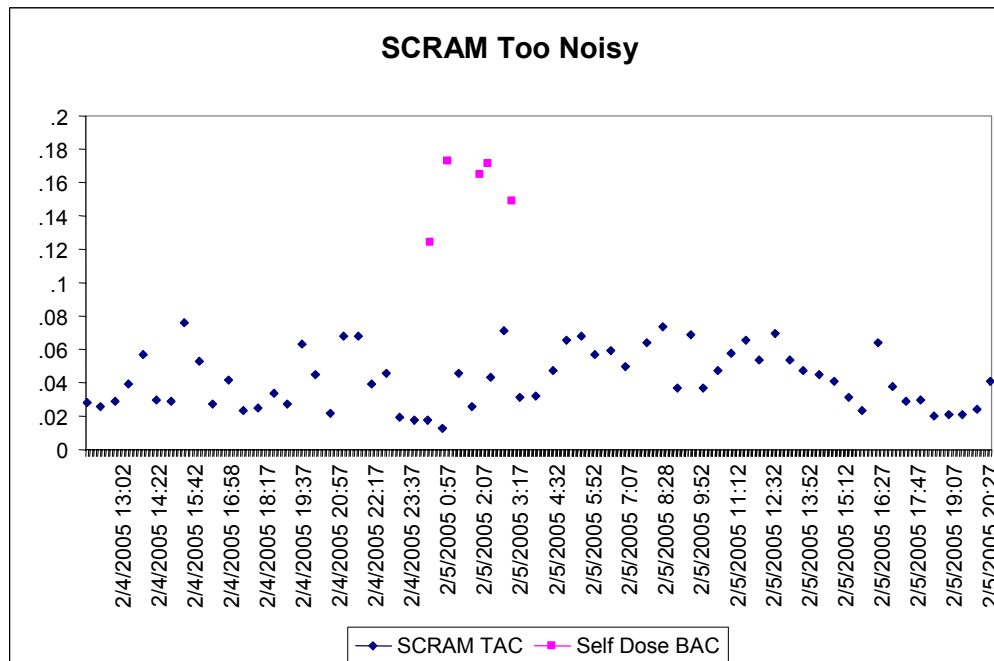
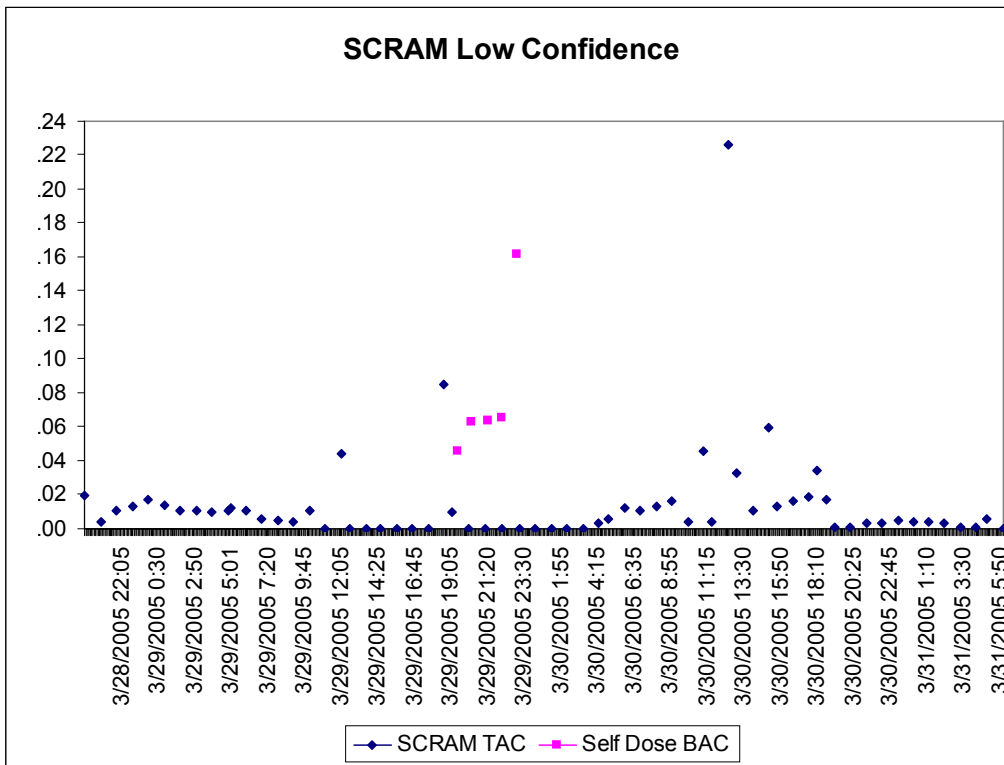


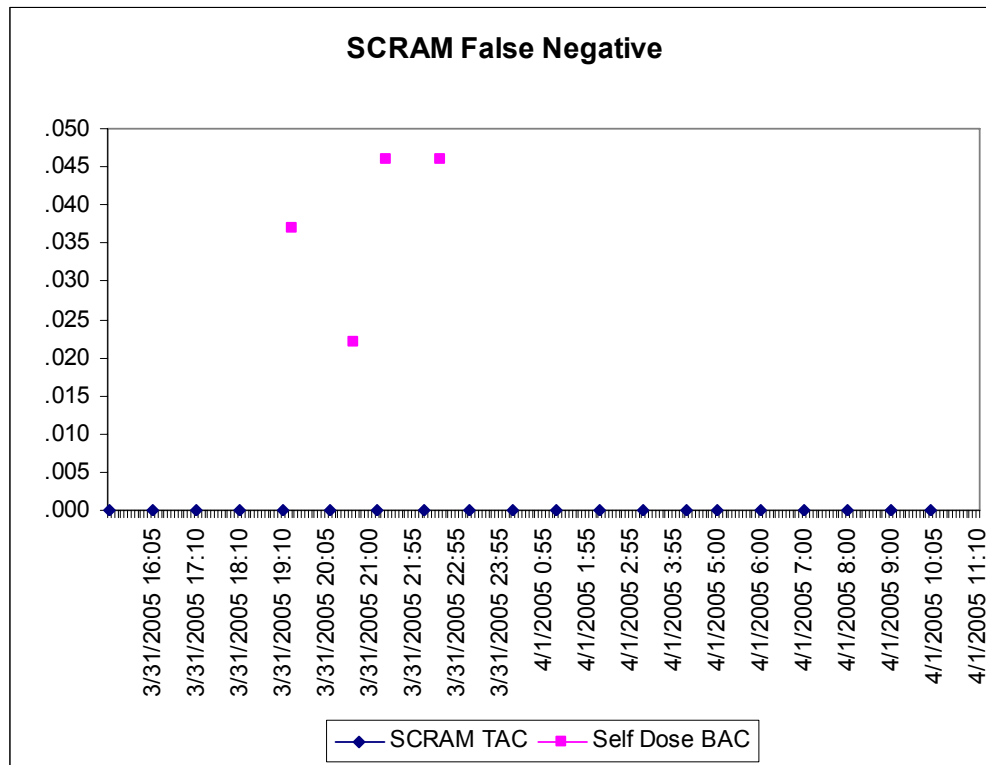


Appendix G – Coded Judgment Examples

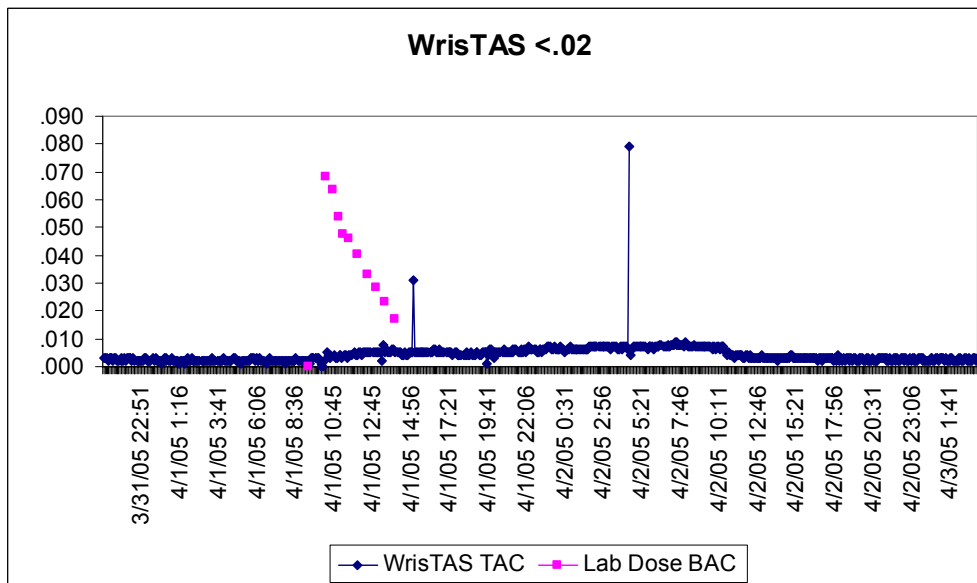
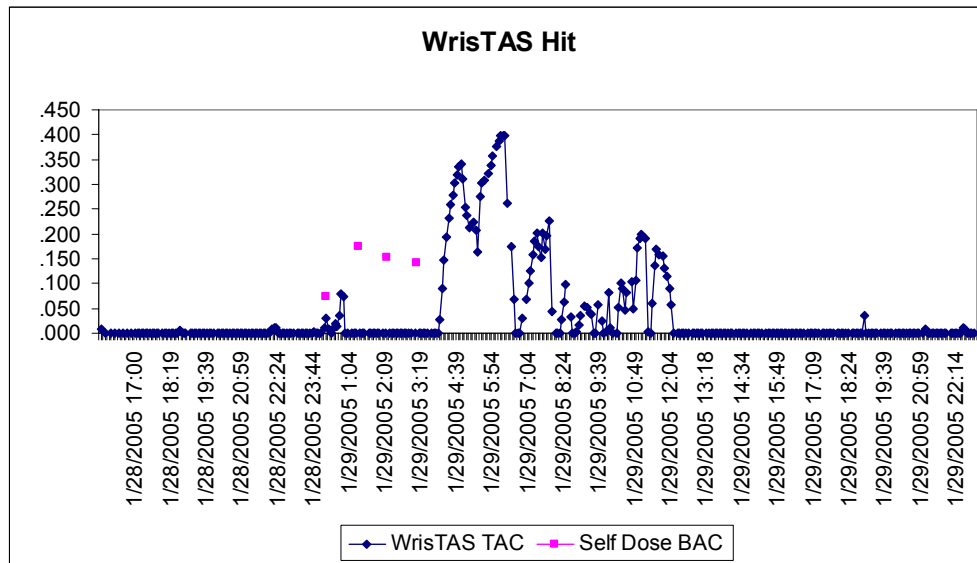
SCRAM™ Examples

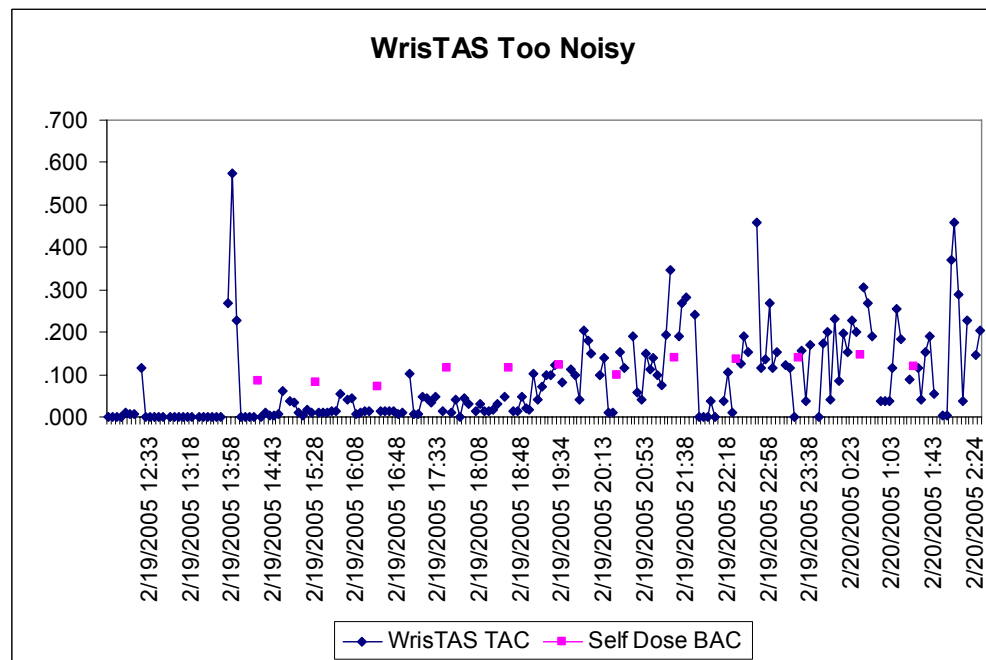
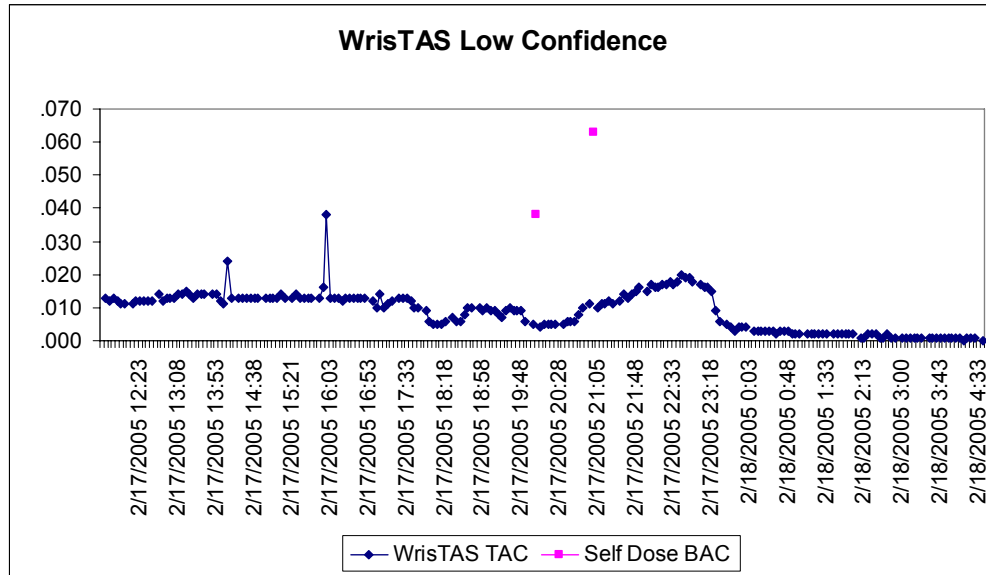


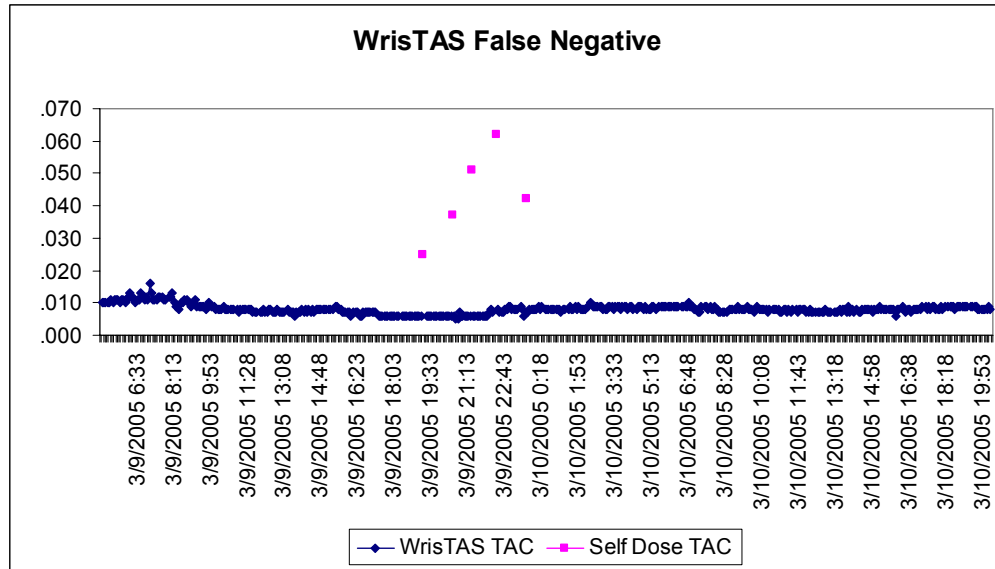




WristAS™ Examples







Appendix H – SCRAM™ Vendor Feedback

Several SCRAM™ Providers were interviewed regarding their experiences with the SCRAM™ system and offenders sentenced to SCRAM™. Interviewees included private providers, officials from parole offices and judges.

None of the people interviewed had exact numbers on the percentage of offenders who had violated the terms of SCRAM™ by drinking. Estimates from both vendors and AMS were somewhere around 20 percent. This includes people who would be identified as drinking small amounts of alcohol and/or people who violated early but stopped later. Reasons for improvement may have been: (1) realizing that the device *would* in fact detect their drinking, and (2) getting their drinking more under control over time. Offenders who drink tend to start out denying that they did so, until confronted with the data and an explanation of how the devices works, at which point they will admit to drinking. Most of the time when offenders attempt to get away with drinking, they will try to drink small amounts and obstruct the device. Providers say they can identify these situations relatively easily. They will call the offender as soon as possible and tell them they can see the drinking and/or the obstruction and tell them to stop drinking. Once offenders understand that they will get caught they improve. There are some offenders who simply decide that they are going to drink as much as ever. They tend to cut the device off. Often these offenders will also leave home and hide to avoid being arrested for violating parole. Sometimes these offenders are identified not by receiving data indicating the bracelet has been removed, but because the modem is disconnected and never calls in.

Providers discussed ways in which offenders have tried to circumvent the system. Most involve obstructing the bracelet input. Common substances used include plastic wrap and socks. Also mentioned were cardboard, Vaseline, aluminum foil, and large band-aids. Multiple providers mentioned offenders who thought they could avoid detection by using antiperspirant to prevent alcohol from exiting the body with sweat. Some of the more amusing obstruction stories involved slices of bologna and a plastic bag full of raw meat. A few offenders had managed to remove the bracelet for a while, though it was obvious from the data that it had been off if it was not shifted to someone else. There was a report of one offender who came into the office with the device obviously damaged from removal, and reinstalled upside down and on the wrong leg. There was a report of someone managing to get the bracelet off and putting it on someone else. In this case, a change in infrared distance suggested tampering and the offender eventually admitted what had happened. Providers allowed that offenders might get away with drinking if they drank only a little, spaced drinking out over time, and/or drank after eating. Providers suggested, though, that this wasn't the kind of behavior they were most concerned about, that the offenders with the most problems would not be able to limit drinking in such a way, and that drinking offenders would be identified eventually, if not immediately.

Providers felt that they were able to understand the graphs provided on the SCRAMNet site. They stressed that it was important to know the offenders and their lifestyles so that they could interpret the data properly. In particular, offenders' work situations had much to do with the likelihood of seeing interferent spikes and distance anomalies at certain times of the day.

Providers did not believe that false positives were a problem. Most potential false positives are in the form of interferent spikes. These are generally easy to distinguish from actual drinking. They are usually caused by exposure to substances that offenders have been warned to avoid. When offenders experienced these spikes they were further warned to avoid the substances. Those offenders who experienced negative consequences as a result of spikes seemed to do better at avoiding them afterward. Where spikes were not predictable (e.g., offender experienced a new substance and was not aware that it contained an interferent) or where interferents were unavoidable (e.g., occasional exposure at work), providers could interpret them appropriately by talking to the offenders and knowing the offenders' personal stories.

Providers were positive about their experiences working with AMS support staff, whom they found to be knowledgeable and helpful. There have been problems with some of the equipment which required them to bring offenders in for maintenance and equipment swapping. This was more the case in the early days of SCRAM™ than recently. One frequently mentioned problem was a bad bunch of batteries that had been sent out into the field. This problem was identified and addressed relatively swiftly.

Providers sometimes had problems working with offenders. A primary problem involved trying to get paid by offenders on a timely basis. Providers were displeased with the large and generally unpredicted amount of time they spent being "bill collectors." Other problems included getting data from offenders when they were supposed to. Some referred to a phenomenon which one referred to as the "80/20 rule" in which 20 percent of offenders accounted for 80 percent of their problems. It was suggested that the problem offenders tended to behave as they did out of a number of inter-correlated factors which included alcoholism, unstable work and home lives, lack of education, difficulties understanding what is being required of them and lack of honesty. It was reported that most problem clients either dropped out of the program or improved as they became more familiar with SCRAM™ and as they "got their act together" following cessation of drinking.

Providers reported that SCRAM™ was a beneficial experience for many of the offenders. As with the offenders who participated in the focus groups, SCRAM™ helped to quit drinking by providing a specific date after which they were not allowed to drink, and a system for "keeping them honest." Some offenders reportedly credited SCRAM™ with saving their life by facilitating drinking cessation.

DOT HS 810 875
November 2007



U.S. Department
of Transportation
**National Highway
Traffic Safety
Administration**

