Published in final edited form as:

Addict Behav. 2015 November ; 50: 205–212. doi:10.1016/j.addbeh.2015.06.042.

Quantifying alcohol consumption: Self-report, transdermal assessment, and prediction of dependence symptoms[☆]

Jeffrey S. Simons^{a,*}, Thomas A. Wills^b, Noah N. Emery^a, and Russell M. Marks^a

^aUniversity of South Dakota, United States

^bUniversity of Hawaii Cancer Center, United States

Abstract

Research on alcohol use depends heavily on the validity of self-reported drinking. The present paper presents data from 647 days of self-monitoring with a transdermal alcohol sensor by 60 young adults. We utilized a bio chemical measure, transdermal alcohol assessment with the WrisTAS, to examine the convergent validity of three approaches to collecting daily self-report drinking data: experience sampling, daily morning reports of the previous night, and 1-week timeline follow-back (TLFB) assessments. We tested associations between three pharmacokinetic indices (peak concentration, area under the curve (AUC), and time to reach peak concentration) derived from the transdermal alcohol signal and within- and between-person variation in alcohol dependence symptoms. The WrisTAS data corroborated 85.74% of self-reported drinking days based on the experience sampling data. The TLFB assessment and combined experience sampling and morning reports agreed on 87.27% of drinking days. Drinks per drinking day did not vary as a function of wearing or not wearing the sensor; this indicates that participants provided consistent reports of their drinking regardless of biochemical verification. In respect to self-reported alcohol dependence symptoms, the AUC of the WrisTAS alcohol signal was associated with dependence symptoms at both the within- and between-person level. Furthermore, alcohol dependence symptoms at baseline predicted drinking episodes characterized in biochemical data by both higher peak alcohol concentration and faster time to reach peak concentration. The results support the validity of self-report alcohol data, provide empirical data useful for optimal design of daily process sampling, and provide an initial demonstration of the use of transdermal alcohol assessment to characterize drinking dynamics associated with risk for alcohol dependence.

Keywords

Validity of self-report; Biochemical verification; EMA; Alcohol dependence

[☆]This research was supported in part by the National Institute on Alcohol Abuse and Alcoholism Grants R01AA017433 and R01AA020519

*Corresponding author at: Department of Psychology, The University of South Dakota, Vermillion, SD 57069, United States. jsimons@usd.edu (J.S. Simons).

Contributors

All authors contributed to and have approved the final manuscript.

Conflict of interest

All authors declare that they have no conflicts of interest.

This research was conducted to examine the convergent validity of a biochemical measure of drinking, transdermal alcohol assessment, with self-reports of drinking behavior. previous research has indicated that self-reports of alcohol drinking and other substance use are reasonably valid and reliable when participants are assured of confidentiality in research settings (Del Boca & Darkes, 2003; Harrison, Martin, Enev, & Harrington, 2007; Patrick et al., 1994; Shillington & Clapp, 2000; Wills & Cleary, 1997), but Del Boca and Darkes (2003) have outlined several factors that should be explored as possibly limiting the accuracy of self-reports. For example, individuals may rely on heuristics that fail to account for deviations in normative patterns (Patrick & Lee, 2010). Measures such as the timeline follow-back (TLFB) provide prompts via use of a calendar to aid individuals in recalling their drinking experience (Sobell, Brown, Leo, & Sobell, 1996; Sobell, Sobell, Litten, & Allen, 1992). Numerous studies support the validity of the TLFB, though its strengths may lie more in obtaining an overall summary of drinking patterns, rather than the number of drinks on a specific date (Hoeppner, Stout, Jackson, & Barnett, 2010; Searles, Helzer, & Walter, 2000).

Prospective methods such as Ecological Momentary Assessment (EMA), Daily Diaries, and the Experience Sampling Method (ESM) attempt to reduce problems with retrospective recall by assessing behavior close to the time it occurs, thus reducing potential memory biases (Shiffman, 2009; Trull & Ebner-Priemer, 2009). Validity of these methods is supported by expected associations with criterion variables; in addition, the methods appear to provide information on variation in behavior across time that is difficult to accurately assess using retrospective accounts (Ebner-Priemer, Eid, Kleindienst, Stabenow, & Trull, 2009; Shiffman, 2009; Simons, Wills, & Neal, 2014). Although research supports the validity of drinking self-reports, much of the evidence comes from observing associations across multiple forms of self-reported behavior, collateral reports, or distal criteria (e.g., alcohol sales; Borsari & Muellerleile, 2009; Northcote & Livingston, 2011; Patrick & Lee, 2010; Rice, 2007). Even with an accurate drink count, differences in the alcohol content of drinks, rate of drinking, as well as individual differences in pharmacokinetics make accurately gauging the level of intoxication difficult. thus, research is needed that compares self-report approaches to biochemical measures, as it can provide valuable information regarding the relative performance of self-report measures and the type of biases that might exist (e.g., over- or under-reporting).

Moreover, drinking in community populations is a dynamic process. Risk for alcohol use disorder may be associated with a number of features of alcohol consumption that go beyond drinking frequency and broad measures of quantity (e.g., total drinks, number of drinks per week). In this regard, rate of drinking, most commonly referred to as “binge drinking,” has been highlighted as a risk factor (Collins, Kashdan, Koutsky, Morsheimer, & Vetter, 2008; Corbin et al., 2014). This metric suggests that consuming 4–5 drinks for women and men, respectively, within a couple of hours is a risk factor for alcohol use disorder as it signifies the frequency of efforts to reach a point of intoxication (e.g., BAC 0.08; NIAAA National Advisory Council, 2004). Maximum drinks in a sitting have also been identified as a risk factor (Esser, Kanny, Brewer, & Naimi, 2012), as have frequency of “extreme” binge drinking, defined as consuming 15 drinks or more per occasion (Hingson & White, 2013; Patrick et al., 2013). In line with this, individual differences in the acute

response to alcohol have been shown to confer vulnerability for alcohol use disorder (King, De Wit, Mcnamara, & Cao, 2011; Schuckit & Smith, 2000; Schuckit et al., 2009). That is, less intense alcohol responses impair detection of internal cues and warning signs to stop drinking, resulting in rapid-excessive alcohol consumption (i.e., binge drinking) which in turn confers risk for alcohol use disorder (King et al., 2011). Thus, it follows that early measures of alcohol dependence such as the alcohol dependence scale (Skinner & Horn, 1984), have items such as “do you gulp drinks,” suggesting that alcohol dependence may be tied to behaviors designed to cause a rapid increase in blood alcohol content. Finally, some authors have observed the tendency for a portion of heavy drinkers to titrate their consumption in such a manner as to maintain a prolonged state of moderate intoxication as opposed to a rapid rise and fall in BAC (Hughes et al., 2011). These three aspects suggest that rate of rise, peak concentration level, and area under the curve, common pharmacokinetic indices, may be relevant to understanding individual drinking behavior and associated risks (Irwin, Goodwin, Leveritt, Davey, & Desbrow, 2012; Whitfield et al., 2001).

To get at these methodological questions about self-reported drinking, transdermal alcohol assessment provides a means of objective and unobtrusive assessment of alcohol consumption in the natural environment. A small amount of alcohol is secreted continuously through the skin, which can be measured via an electrochemical sensor (Tempelman, Moeller, & Schmitt, 2010). These devices are commonly used in the criminal justice system to monitor enforced abstinence (Marques & Mcknight, 2009). The wrist transdermal alcohol sensor (WrisTAS) is a device about the size of a watch that provides continuous readings of alcohol secreted through the skin (Tempelman et al., 2010). In addition, it contains temperature and skin conductance sensors that are used to determine that the device is being worn. research comparing the WrisTAS with breath alcohol concentration (BrAC) as well as self-reported drinking events supports the validity of the WrisTAS transdermal alcohol assessments (Bond, Greenfield, Patterson, & Kerr, 2014; Luczak & Rosen, 2014; Swift, 2000). There is roughly a 30-minute delay in alcohol readings via transdermal assessment and BrAC (Tempelman et al., 2010). In addition, there is greater individual variability in secretion of alcohol through the skin relative to breath analysis. Hence, the WrisTAS is not a direct quantification of current alcohol levels (Tempelman et al., 2010). These features provide limits to the use of transdermal assessment to directly quantify blood alcohol concentration without individually calibrating devices based on a laboratory alcohol challenge, though equations have been developed with some transdermal devices (e.g., scam; Dougherty et al., 2012; Hill-Kapturczak et al., 2014). Nonetheless, transdermal alcohol assessment can be used to accurately examine within-person variation in alcohol consumption and hence provide some validation of self-report, if not precise accuracy of self-reported alcohol consumption. On an exploratory level, we also examined the ability of transdermal assessment with the WrisTAS to differentiate patterns of alcohol consumption as related to between-person variability in alcohol dependence symptoms.

The present study has four broad aims. First, we seek to validate self-report alcohol data against biochemical data from transdermal alcohol assessment. Second, we provide data on the reactivity to wearing the WrisTAS device. For the purposes of this paper, reactivity is conceptualized as changes in reported drinking behavior as a function of wearing the device.

That is, do participants report more or less drinking when they wear the device relative to days they do not wear the device? Third, we seek to compare associations between transdermal alcohol assessment and three approaches to daily drinking assessment: experience sampling via (1) random prompts, (2) self-initiated retrospective assessments of drinking the previous night, and (3) a brief TLFB. These three forms of daily self-report data vary in respect to ease, cost, and participant burden. Hence the results may be useful for researchers seeking to design studies to obtain an efficient estimate of daily drinking. Fourth, we calculate several pharmacokinetic indices from the transdermal data to examine both within- and between-person associations with symptoms of alcohol dependence. The study extends previous research by testing the incremental validity of the WrisTAS assessments in predicting daily symptoms (i.e., within-person effects) of alcohol dependence over and above drinking self-report. Moreover, we examine between-person associations between pharmacokinetic indices from natural drinking episodes and symptoms of alcohol dependence.

1. Method

1.1. Participants

Participants were 60 young adults aged 18–21 ($M = 19.57$, $SD = 1.00$). Fifty-two percent were women, 97% were white, 1.5% were Native American, and 1.5% were black. Two percent identified as Hispanic. Recruitment was conducted through e-mail notices and advertisements in local media. This is a subsample randomly selected to wear the WrisTAS device as part of a larger study (Simons et al., 2014).

1.2. Procedure

Undergraduates between ages 18–25 years who drank at least moderately (i.e., 12 drinks/week for women and 16 drinks/week for men; i.e., N12 drinks/week for women and N16 drinks/week for men; Sanchez-Craig, Wilkinson, & Davila, 1995) were invited to participate in the experience sampling (ESM) study. Invited participants provided informed consent for the study, completed a set of baseline questionnaires, and were then trained in the use of the palmtop computers (PDAs). The PDAs were programmed with PMAT software (Weiss, Beal, Lucy, & Macdermid, 2004), modified by Joel Swendsen and CNRS, France. The program was configured to prompt participants to complete brief ~2 minute assessments at random times within 2-hour blocks between 10:00 a.m. and 12:00 midnight. In addition, participants completed a self-initiated assessment each morning shortly after waking, which included assessments of alcohol consumption and dependence symptoms the previous night (e.g., drinking more than intended). The study had a burst design, that is, participants carried the PDAs for 1–2 week measurement bursts over the course of 3 semesters. A random sample of participants (20%) was designated to wear the WrisTAS device, and in this subsample, 2 bursts were randomly chosen for wearing the WrisTAS. The current sample is restricted to this subset of the larger study. Finally, at the end of each measurement burst, participants completed the TLFB for the past week on the computer. Further description of study procedures can be found in Simons, Wills, and Neal (2014).

1.3. Measures

1.3.1. Alcohol consumption—On each random in situ assessment, participants reported the number of standard drinks they consumed over the past 30 min on an 8-point scale (0–7 or more drinks). The total number of drinks reported each night is the experience sampling measure of drinking. Definitions of standard drinks were provided during the palmtop training. In addition, each morning the participants initiated an assessment and provided the total number of drinks consumed the previous evening on a 0–25 point scale. This is the daily retrospective assessment of drinking. Finally, at the end of the week, participants completed a 7-day TLFB assessment on a computer. The program provided standard instructions and definitions of standard drinks. The program displayed a calendar in which the participant worked backward from the previous night and entered the number of drinks for each night. The TLFB measure differs from the original in four ways: (1) it is presented on the computer, (2) it is self-administered, (3) it does not provide personal memory cues (e.g., birthdays), and (4) it is a brief 1-week assessment rather than the extended 6 month to 1 year versions. Because of the brief nature of the TLFB, we do not think that the lack of personally relevant memory cues is a substantial limitation. Similar brief and self-administered TLFB surveys have been used in previous research (Collins et al., 2008; Hoeppepner et al., 2010; Maisto, Conigliaro, Gordon, McGinnis, & Justice, 2008), and demonstrate excellent test–retest reliability. However, the pronounced difference in timeframe makes the results reported here of limited relevance to the validity of the TLFB procedure over longer time periods.

1.3.2. Alcohol dependence symptoms—Random in situ assessments included a 7-item checklist with the following choices: (a) felt sick or vomited; (b) drank when you promised yourself not to; (c) had withdrawal symptoms; (d) tried unsuccessfully to limit your drinking, cut back, or stop; (e) drank more or for a longer time than you intended; (f) drank more than usual to get drunk; and (g) felt alcohol effects less than usual for amount used. The morning assessment assessed whether the participant (a) passed out, (b) blacked out, (c) vomited, (d) needed to drink more than usual to get the desired effect, (e) felt less effects than usual for the amount drank the previous night, (f) a.m. withdrawal symptoms, and (g) hangover. Definitions of blackouts, withdrawal symptoms, and hangovers were provided during the palmtop training. The morning assessment was designed, in part, to pick up symptoms that might have been missed in the random sampling. Items in the morning assessment that corresponded to items in the random assessments (e.g., felt effects less for amount used) were recoded to zero if the item had been endorsed in the evening. This prevents inadvertently counting an event twice. After recoding, the sumtotal of items endorsed for the repeated nighttime assessments and the morning assessments was the measure of acute dependence symptoms. Previous research with this and another sample supports the validity of the approach (Simons, Dvorak, Batien, & Wray, 2010; Simons et al., 2014). In addition, the alcohol dependence scale (Skinner & Horn, 1984) was administered at baseline. This was used as the between-subjects assessment of alcohol dependence symptoms.

1.3.3. Transdermal BAC monitoring—Transdermal monitoring was measured with the wrist-worn Giner WrisTAS™ 7 alcohol sensor (Tempelman et al., 2010). Approximately 1–

2% of alcohol consumed by humans is excreted through the skin; when a sample of the perspiration is trapped in front of an electrochemical ethanol sensor this measurement of transdermal alcohol can be correlated to blood alcohol concentration (BAC). The WrisTAS has three auxiliary sensor elements. Two monitor skin temperature and conductance to confirm/record that the sensor is being worn against the skin. A third confirms that the sensor is hydrated to guard against false negative alcohol readings. These tamper-resistant features provide extra data integrity/security. The WrisTAS sensors have been successfully tested in laboratory and field settings (Bond et al., 2014; Swift, 2000). The WrisTAS sensor detects readings from 10 to over 250 mg/dl BAC. Since the sensor is sealed against the skin with medical-grade foam, the sensor does not detect environmental alcohol or other external environmental species. Transdermal alcohol levels do not rise as quickly as blood and breath alcohol levels after the consumption of alcohol. Hence, transdermal sensing has a lag time due to the pharmacokinetics of alcohol in the tissue and skin compartment, so while the alcohol sensor has an intrinsic response of several minutes, transdermal detection of the onset of drinking lags by about 30 min. Recent research indicates a sensitivity of 85.6% and specificity of 67.5% for the WrisTAS in detecting self-reported drinking events (Bond et al., 2014).

2. Results

2.1. Descriptive statistics

Participants responded to 82% of the random experience sampling prompts and 95% of the self-initiated morning reports. Participants in this moderate to heavy drinking sample reported drinking for 34.62% of the days on the PDA and for 36.17% of days on the TLFB. For comparison, the WrisTAS indicated drinking on 31.48% of days. TLFB assessment captured 87.27% of drinking days (based on ESM data). Similarly, TLFB results corresponded with 89.76% of non-drinking days (based on ESM data). On the morning reports, participants reported drinking an average of 7.61 ($SD = 5.06$) drinks per drinking day. On the TLFB, participants reported drinking an average of 8.04 ($SD = 5.47$) drinks per drinking day. Participants removed the WrisTAS device on 9.89% of days and the sensor failed (e.g., faulty electronic connectivity) on 17.93% of days. Descriptive statistics are in Table 1.

2.2. Reactivity to WrisTAS

Reactivity was tested by two multilevel negative binomial hurdle models in Mplus 7.1 (Muthén & Muthén, 2012). For the multilevel analyses, days are nested within person. the negative binomial hurdle analysis is a 2-part model that first tests a logistic model to predict zero versus non-zero values (i.e., abstaining vs. any drinking) and then estimates a truncated negative binomial count model (i.e., number of drinks), conditional upon the likelihood of a nonzero value, indicating that drinking occurs (Atkins, Baldwin, Zheng, Gallop, & Neighbors, 2013). Thus, in the current analyses, we estimate whether the individual reports *no drinking* on a given day in the logistic part of the model, and then estimates the number of reported drinks (the negative binomial count) conditional upon the likelihood of drinking. Self-reported drinks from the morning reports were the criterion variable. This is a daily assessment of the number of drinks consumed the previous night. for the first model, this

was predicted by an indicator variable signifying whether the WrisTAS was worn (days it was assigned and worn (1) vs. days that it was not scheduled (0)) at level 1 (i.e., within-person level) and the subject mean of this variable at level 2 (i.e., between-person level). The subject mean of the indicator is the percent of days that individual wore the device. This tests whether individuals who were more compliant with wearing the WrisTAS differed in reported drinking at the between-subjects level. The level 1 predictor was centered at the person mean and the level 2 predictor was centered at the grand mean to disaggregate the within- and between-person effects (Enders & Tofighi, 2007). The first model had 60 participants and 2188 person-days. In the following analyses, we report the incident rate ratios (IRR) for the count part of the model and odds ratios for the logistic part of the model. The IRR indicates the percent change in the count for every unit change in the predictor. The OR indicates the percent change in the odds for every unit increase in the predictor (Hilbe, 2008). The number of self-reported drinks on drinking days did not vary as a function of wearing the WrisTAS ($IRR = 1.02, p = .687$) nor did the likelihood of reporting abstaining ($OR = 0.83, p = .060$). Hence, reported drinking on days when the sensor was worn did not differ from days when the individual was not scheduled to wear the sensor. There were no significant contextual effects at level 2. A similar model compared days it was worn (0) vs. days the participant took it off (1). This model contained 60 participants and 434 person-days (i.e., only includes days it was scheduled). Consistent with the above, there was no relation between removing the sensor and self-reported drinks ($IRR = 0.44, p = .075$). However, taking the WrisTAS off was associated with an increased like lihood of reporting abstaining ($OR = 6.55, p = .004$) at level 1. There were no significant effects at level 2. Individuals may remove the device because they are engaging in activities that are incompatible with both the device and drinking (e.g., sports participation). Overall, evidence for reactivity seems minimal.

2.3. Convergent validity analyses

Sensitivity, specificity, and area under the curve (from the derived receiver operating curve) were calculated from a multilevel logit model in Stata 13 (Statacorp, 2013) predicting self-reported drinking days in the experience sampling data from peak transdermal readings from the WrisTAS. The WrisTAS correctly classified 85.74% of self-reported drinking events at a cutscore of .4, sensitivity was 72.35%, and specificity was 92.94%. The receiver operating curve (roc) area was .89. The WrisTAS thus performed well in detecting self-reported drinking via experience sampling.

Next, we estimated three multilevel negative binomial hurdle models in Mplus 7.1 (Muthén & Muthén, 2012) to compare correspondence between daily peak WrisTAS assessments and three self-report measures (see Table 2). The self-reports included (1) nightly totals of random in situ assessments of drinks in the past 30 min, (2) morning assessments on the PDA of total number of drinks the previous night, and (3) 1-week computerized TLFB assessments. Each model included gender and the person-mean scores on the WrisTAS as covariates. Daily peak WrisTAS score was a level 1 (within-person) predictor and centered at the person-mean. Gender and the subject mean peak WristAS score were level 2 (between-person) variables and were centered at the grand mean. The in situ assessment model included number of nightly assessments as an exposure variable. As expected, daily

peak readings on the WrisTAS predicted self-reported drinks. The IRRs for drinks in the count portion were quite consistent across models (peak $\rightarrow$ in situ assessment, IRR = 1.022, $p < .001$; peak $\rightarrow$ morning reports, IRR = 1.021, $p < .001$; peak $\rightarrow$ TLFB, IRR = 1.026, $p < .001$). Thus, for each 0.01 g/dl increase in the WrisTAS peak, the number of reported drinks increased by 2.1–2.6%. For the logit portion, the ORs predicting not drinking were 0.93, $p < .001$ (in situ), 0.80, $p < .001$ (morning reports), and 0.82, $p < .001$ (TLFB). These results suggest that each self-report approach is comparably associated with the transdermal readings, though the random in situ reports may be more likely to miss a drinking episode. Fig. 1 displays associations between the WrisTAS and self-reported drinks from the morning reports for drinking days. (See Fig. 2.)

2.4. Incremental validity analysis

In this analysis, we test whether the WrisTAS transdermal alcohol assessment predicts daily reported dependence symptoms over and above the effects of more commonly used alcohol self-report assessment. We tested a multilevel model in Mplus 7.1 (Muthén & Muthén, 2012) predicting daily reported dependence symptoms from the experience sampling data from nightly totals of random in situ assessments of drinks in the past 30 min and the area under the curve of the transdermal alcohol signal. The level 1 variables (i.e., nightly drinking self-report and AUC of transdermal signal) were centered at the person mean and the person means of the drinks and AUC variables were included at level 2 (centered at the grand mean). Gender was included as a covariate at level 2. Dependence symptoms were specified as count variable with a negative binomial distribution. We first tested a reduced model that included only self-reported drinking and gender. Self-reported drinking was a significant predictor at L1 (IRR = 1.31, $p < .001$) and l2 (IRR = 2.12, $p < .001$). The effect of gender was not significant (IRR = 0.70, $p = .281$). We then added the AUC of the transdermal signal within- (l1) and between- (l2) person predictors to the model. Both self-reported nighttime drinks and the transdermal AUC were significant predictors at both within- and between-person levels (see Table 3). Thus, both self-reported drinks from the experience sampling as well as transdermal alcohol assessment provide unique information in the prediction of alcohol dependence symptoms. To determine the incremental validity of the transdermal assessments relative to self-report, we calculated a likelihood ratio test comparing the full and reduced models with a scaling correction for the MLR estimator (Muthén & Muthén, 2012; $\chi^2(2) = 54.67$, $p < .001$). Thus, the model including the transdermal data was a significant improvement over the self-report model.

2.5. Drinking style and dependence symptoms

The final analysis tested between-person associations between alcohol dependence symptoms (measured at baseline) and subject means of two pharmacokinetic values derived from drinking episodes: peak concentration and time to reach peak concentration. This analysis was conducted on 50 individuals with valid WrisTAS readings of drinking episodes. One individual was a multivariate outlier and removed from the analytic sample, resulting in a sample of 49. Scores on the alcohol dependence scale were regressed on gender, subject mean peak concentration, time to reach peak concentration, and the peak $\times$ time to reach peak interaction. Continuous variables were centered at their mean before forming the interaction. The negative binomial regression model was significant $\chi^2(4, n =$

49) = 10.26, $p = .0363$, see Table 4. As expected, average peak concentration from the WrisTAS was positively associated with scores on the alcohol dependence scale obtained from the baseline assessment. Furthermore, time to reach peak concentration was inversely associated with scores on the alcohol dependence scale, indicating that individuals with greater alcohol dependence tend to drink in a manner to reach peak concentration more rapidly. This effect was qualified by the significant peak concentration $\times$ time to reach peak concentration interaction. For individuals reporting high average concentrations (i.e., $m + 1$ SD) the association between time to reach peak concentration and alcohol dependence symptoms was not significant (IRR = 0.86, $p = .619$). However, at the mean (IRR = 0.34, $p = .006$), and 1 SD below the mean (IRR = 0.13, $p = .007$) of peak concentration, there was an inverse association between time to reach peak concentration and alcohol dependence symptoms. This is consistent with the premise that alcohol dependence is associated with rapid ingestion of alcohol during drinking episodes.

3. Discussion

The current study adds to literature on the validity of self-report assessment of alcohol use and a growing body of literature on the use of transdermal alcohol assessment (Barnett, Meade, & Glynn, 2014; Bond et al., 2014; Leffingwell et al., 2013; Luczak & Rosen, 2014). Given proper assurances and appropriate assessment methods, research participants can provide reliable and valid assessment of their alcohol consumption. The consistency between reported alcohol use on days the device was assigned and those it was not, suggests that the reliability of self-report does not vary as a function of the prospect of biochemical verification. Though there was a significant association between taking the WrisTAS device off and self-reported abstaining, we suspect this is more likely due the device being removed to participate in activities in consistent with both alcohol use and wearing the sensor (e.g., playing sports). Overall, there was high convergence between the WrisTAS and self-report in detecting drinking days, supporting the validity of each approach to assessment.

The results add to the literature in three ways. First, we provide data on the correspondence between transdermal assessment and three common approaches to assessing daily drinking data. Second, we demonstrate that the WrisTAS and experience sampling drinking data each showed an independent association with alcohol dependence symptoms, providing evidence of incremental validity as well as criterion validity. Finally, we provide preliminary evidence utilizing unique aspects of the transdermal data to characterize dynamic aspects of drinking styles and their implications for risk of alcohol dependence. These issues are discussed in turn below.

3.1. Validity of TLFB, experience sampling, and daily reports of drinking

Peak concentration assessed by the WrisTAS was significantly associated with TLFB, experience sampling, and daily drinking reports at both the within- and between-person levels. Incident rate ratios (IRR) were highly consistent across the approaches, indicating that the number of reported drinks on drinking days increased by 2.1–2.6% for each 0.01 g/dl rise in the transdermal peak concentration. There was more variability in the prediction of abstaining. The odds ratio (OR) for the morning assessments and TLFB were 0.80 and 0.82, respectively. This indicates that the odds of abstaining decreased by 18–20% with each

0.01 g/dl rise in peak concentration. In contrast, the OR for the experience sampling analysis was 0.93, indicating that the odds of abstaining decreased by only 7% for each .01 g/dl rise in peak concentration. Hence, the experience sampling, which in this case, obtained random 30-minute samples of behavior every couple of hours, may miss low frequency drinking events. In summary, though the performance was quite consistent across methods, the less burdensome daily reports and TLFB methods did very well and were superior in detecting low intensity drinking. Nonetheless, researchers are often interested in not only assessing number of drinks in these studies but also assessing predictors such as affect, which may be more influenced by memory biases or establishing temporal associations within days, in which case experience sampling approaches may be necessary. A combination of experience sampling and daily morning reports (Gaher et al., 2014; Simons et al., 2010, 2014) or use of EMA (Piasecki et al., 2011) may be necessary to accurately assess low intensity drinking events.

3.2. Criterion and incremental validity of transdermal assessment and ESM

Area under the alcohol signal curve (AUC) and ESM drinking data each accounted for unique variance in alcohol dependence symptoms at both the within- and between-person level. This provides evidence of the criterion validity of the transdermal assessments as well as ESM assessments. Furthermore, the associations between the experience sampling of dependence symptoms and transdermal AUC data provide support for the criterion validity of the alcohol dependence assessment. Interestingly, both the AUC and ESM indices of drinking had significant independent relations to alcohol dependence symptoms. Given that both measures provide some estimate of drinking level, it is unclear what unique information each provides. One possibility is that the random behavioral samples in the experience sampling capture instances of rapid succession of drinks (e.g., having had several shots prior to a random prompt). Such behavior may not be adequately quantified in the AUC. Alternatively, this could reflect shared method variance between the experience sampling data of drinking and dependence symptoms. However, secondary analyses (not presented) showed similar associations with data from the TLFB as well as morning assessments, which would not be accounted for by either of the above explanations. This issue warrants further consideration.

3.3. Pharmacokinetic characteristics of drinking episodes and risk for alcohol dependence

The present study was able to provide information about dynamic characteristics of drinking episodes. Peak alcohol concentration during drinking episodes was positively associated with alcohol dependence symptoms at baseline as expected, a finding that is in line with previous studies examining variability in pharmacokinetics and risk for dependence (Whitfield et al., 2001). In addition, time to reach peak concentration was inversely associated with alcohol dependence. Furthermore, peak concentration and time to reach peak concentration interacted, such that time to reach peak concentration was predictive of alcohol dependence primarily for those who drank at mean or lower intensities. Hence, alcohol dependence was associated with a pattern of drinking characterized by both high peak concentration of alcohol as well as rapid rise of alcohol to peak concentration. This suggests that for some, differences in the manner in which they reach peak concentration may be indicative of risk for alcohol dependence despite reaching lower levels of

intoxication on average. This finding is qualified by the relatively small sample of drinking episodes for analysis and warrants replication with a larger sample. However, it is an interesting demonstration of the potential to advance understanding of alcohol dependence by evaluating the pharmacokinetic indices calculated from natural drinking episodes.

3.4. Limitations

The primary limitation of the study is the relatively small sample size, both in respect to number of participants and number of assessment days. Additional work with larger samples is needed to replicate and extend the results. In addition, the incorporation of an initial alcohol challenge to provide data to convert person-specific transdermal alcohol assessment to estimates of BAC would provide further information to aid in interpreting the transdermal signal. Previous research utilizing the SCRAM device reported high correspondence between a random effects model prediction and observed BAC in a laboratory experiment (Dougherty et al., 2012; Hill-Kapturczak et al., 2014). However, in the current study we focus on the prediction of self-reported drinks and dependence symptoms rather the ability of the device to model BAC levels.

Our results indicate participants can provide reasonably valid self-report data. However, generalizability of the results to other samples should be done with caution. The validity of self-report is inherently related to both characteristics of the participant and characteristics of the assessment setting. For example, results are likely to differ in situations where individuals fear negative consequences stemming from their self-report data. In addition, though the sample is a relatively heavy drinking sample, they are well functioning young adults. Further research is needed to examine these issues in clinical samples with AUD.

Though there was some evidence of increased reports of abstaining when taking the sensor off, as we discuss above, we suspect this is not due to an under-reporting bias in the absence of verification. The current data suggests that the prospect of verification did not enhance the validity of self-report at the within-person level (i.e., self-reported drinking didn't differ on days it was assigned versus not). However, unfortunately, the study cannot test the broader question of whether the prospect of verification can enhance the validity of self-report, as all participants knew they could be randomized to receive the WristAS in the future.

3.5. Summary

Some will always view self-reported alcohol and drug use data with skepticism. Substance users have been viewed as immoral and deceitful for ages, and this has unfortunately carried over into both clinical and research realms. Certainly, the proposition of scorn, loss of benefits or freedom, and potential criminal prosecution are good reasons for individuals to minimize and hide their alcohol and other substance use. However, the results of this study demonstrate that in a research context, participants can provide valid reports of their drinking behavior. There was remarkable consistency across different self-report methods and the self-reported drinking exhibited robust associations with transdermal alcohol data. Though these data cannot determine the precise accuracy of self-report, they provided no evidence of an underreporting bias. Indeed, participants self-reported drinking on slightly

more days than detected by transdermal assessment. In respect to criterion validity, the AUC of the alcohol signal predicted alcohol dependence symptoms assessed via experience sampling at both the within- and between-person level. In addition, there were significant associations between self-reported alcohol dependence symptoms at baseline and transdermal alcohol signal during drinking episodes. This latter finding was less robust, but this may be due to the relatively small number of drinking episodes. The results provide empirical evidence of the relative performance of three approaches to daily assessment that will be useful for research designing daily process studies. Finally, the analyses provide an initial demonstration of the use of transdermal alcohol assessment to identify dynamic characteristics of drinking to advance understanding of alcohol use disorder.

Acknowledgments

Role of funding sources: This research was supported in part by the National Institute on Alcohol Abuse and Alcoholism Grants R01AA017433 and R01AA020519 to Jeffrey S. Simons.

References

- Atkins DC, Baldwin SA, Zheng C, Gallop RJ, Neighbors C. A tutorial on count regression and zero-altered count models for longitudinal substance use data. *Psychology of Addictive Behaviors*. 2013; 27:166. [PubMed: 22905895]
- Barnett NP, Meade EB, Glynn TR. Predictors of detection of alcohol use episodes using a transdermal alcohol sensor. *Experimental and Clinical Psychopharmacology*. 2014; 22:86–96. <http://dx.doi.org/10.1037/a0034821>. [PubMed: 24490713]
- Bond JC, Greenfield TK, Patterson D, Kerr WC. Adjustments for drink size and ethanol content: New results from a self-report diary and transdermal sensor validation study. *Alcoholism: Clinical and Experimental Research*. 2014; 38:3060–3067.
- Borsari B, Muellerleile P. Collateral reports in the college setting: A meta-analytic integration. *Alcoholism: Clinical and Experimental Research*. 2009; 33:826–838.
- Collins RL, Kashdan TB, Koutsky JR, Morsheimer ET, Vetter CJ. A self-administered timeline followback to measure variations in underage drinkers' alcohol intake and binge drinking. *Addictive Behaviors*. 2008; 33:196–200. [PubMed: 17720324]
- Corbin WR, Zalewski S, Leeman RF, Toll BA, Fucito LM, O'Malley SS. In with the old and out with the new? A comparison of the old and new binge drinking standards. *Alcoholism, Clinical and Experimental Research*. 2014; 38:2657–2663. <http://dx.doi.org/10.1111/acer.12514>.
- Del Boca F, Darkes J. The validity of self-reports of alcohol consumption: State of the science and challenges for research. *Addiction*. 2003; 98:1–12. [PubMed: 14984237]
- Dougherty DM, Charles NE, Acheson A, John S, Furr RM, Hill-Kapturczak N. Comparing the detection of transdermal and breath alcohol concentrations during periods of alcohol consumption ranging from moderate drinking to binge drinking. *Experimental and Clinical Psychopharmacology*. 2012; 20:373–381. [PubMed: 22708608]
- Ebner-Priemer UW, Eid M, Kleindienst N, Stabenow S, Trull TJ. Analytic strategies for understanding affective (in)stability and other dynamic processes in psychopathology. *Journal of Abnormal Psychology*. 2009; 118:195–202. <http://dx.doi.org/10.1037/a0014868>. [PubMed: 19222325]
- Enders CK, Tofighi D. Centering predictor variables in cross-sectional multi level models: A new look at an old issue. *Psychological Methods*. 2007; 12:121. [PubMed: 17563168]
- Esser MB, Kanny D, Brewer RD, Naimi TS. Binge drinking intensity: A comparison of two measures. *American Journal of Preventive Medicine*. 2012; 42:625–629. <http://dx.doi.org/10.1016/j.amepre.2012.03.001>. [PubMed: 22608381]
- Gaher, RM.; Simons, JS.; Hahn, AM.; Hofman, NL.; Hansen, J.; Buchkoski, J. An experience sampling study of PTSD and alcohol-related problems. *Psychology of Addictive Behaviors*. 2014. <http://dx.doi.org/10.1037/a0037257>.

- Harrison, LD.; Martin, SS.; Enev, T.; Harrington, D. Comparing drug testing and self-report of drug use among youths and young adults in the general population. Rockville, MD: 2007. (DHHS Publication No SMA 07-4249, Methodology Series M-7, RTI/0209009.186.002)
- Hilbe JM. Brief overview on interpreting count model risk ratios: An addendum to negative binomial regression. 2008
- Hill-Kapturczak N, Lake SL, Roache JD, Cates SE, Liang Y, Dougherty DM. Do variable rates of alcohol drinking alter the ability to use transdermal alcohol monitors to estimate peak breath alcohol and total number of drinks? *Alcoholism, Clinical and Experimental Research*. 2014; 38:2517–2522. <http://dx.doi.org/10.1111/acer.12528>.
- Hingson RW, White A. Trends in extreme binge drinking among US high school seniors. *JAMA Pediatrics*. 2013; 167:996–998. <http://dx.doi.org/10.1001/jamapediatrics.2013.3083>. [PubMed: 24042186]
- Hoepfner BB, Stout RL, Jackson KM, Barnett NP. How good is fine-grained timeline follow-back data? Comparing 30-day tlf and repeated 7-day tlf alcohol consumption reports on the person and daily level. *Addictive Behaviors*. 2010
- Hughes K, Quigg Z, Bellis MA, Van Hasselt N, Calafat A, Kosir M, et al. Drinking behaviours and blood alcohol concentration in four European drinking environments: A cross-sectional study. *BMC Public Health*. 2011; 11:918. [PubMed: 22151744]
- Irwin C, Goodwin A, Leveritt M, Davey AK, Desbrow B. Alcohol pharmacokinetics and risk-taking behaviour following exercise-induced dehydration. *Pharmacology Biochemistry and Behavior*. 2012; 101:609–616. <http://dx.doi.org/10.1016/j.pbb.2012.02.016>.
- King AC, De Wit H, Mcnamara PJ, Cao D. Rewarding, stimulant, and sedative alcohol responses and relationship to future binge drinking. *Archives of General Psychiatry*. 2011; 68:389–399. [PubMed: 21464363]
- Leffingwell TR, Cooney NJ, Murphy JG, Luczak S, Rosen G, Dougherty DM, et al. Continuous objective monitoring of alcohol use: Twenty-first century measurement using transdermal sensors. *Alcoholism, Clinical and Experimental Research*. 2013; 37:16–22. <http://dx.doi.org/10.1111/j.1530-0277.2012.01869.x>.
- Luczak SE, Rosen IG. Estimating BrAC from transdermal alcohol concentration data using the BrAC estimator software program. *Alcoholism: Clinical and Experimental Research*. 2014; 38:2243–2252.
- Maisto SA, Conigliaro JC, Gordon AJ, McGinnis KA, Justice AC. An experimental study of the agreement of self-administration and telephone administration of the timeline followback interview. *Journal of Studies on Alcohol and Drugs*. 2008; 69:468–471. [PubMed: 18432391]
- Marques PR, McKnight AS. Field and laboratory alcohol detection with 2 types of transdermal devices. *Alcoholism, Clinical and Experimental Research*. 2009; 33:703–711. <http://dx.doi.org/10.1111/j.1530-0277.2008.00887.x>.
- Muthén, LK.; Muthén, BO. Mplus statistical modeling software: Release 7.1. Los Angeles, CA: Muthén & Muthén; 2012.
- NIAAA National Advisory Council. Definition of binge drinking NIAAA newsletter. Bethesda, MD: National Institutes of Health; 2004.
- Northcote J, Livingston M. Accuracy of self-reported drinking: observational verification of 'last occasion' drink estimates of young adults. *Alcohol and Alcoholism*. 2011; 46:709–713. <http://dx.doi.org/10.1093/alcalc/agr138>. [PubMed: 21949190]
- Patrick DL, Cheadle A, Thompson DC, Diehr P, Koepsell T, Kinne S. The validity of self-reported smoking: A review and meta-analysis. *American Journal of Public Health*. 1994; 84:1086–1093. [PubMed: 8017530]
- Patrick ME, Lee CM. Comparing numbers of drinks: college students' reports from retrospective summary, followback, and prospective daily diary measures. *Journal of Studies on Alcohol and Drugs*. 2010; 71:554–561. [PubMed: 20553664]
- Patrick ME, Schulenberg JE, Martz ME, Maggs JL, O'Malley PM, Johnston LD. Extreme binge drinking among 12th-grade students in the United States: Prevalence and predictors. *JAMA Pediatrics*. 2013; 167:1019–1025. <http://dx.doi.org/10.1001/jamapediatrics.2013.2392>. [PubMed: 24042318]

- Piasecki TM, Jahng S, Wood PK, Robertson BM, Epler AJ, Cronk NJ, et al. The subjective effects of alcohol–tobacco co-use: An ecological momentary assessment investigation. *Journal of Abnormal Psychology*. 2011; 120:557–571. <http://dx.doi.org/10.1037/a0023033>. [PubMed: 21443289]
- Rice C. Retest reliability of self-reported daily drinking: Form 90. *Journal of Studies on Alcohol and Drugs*. 2007; 68:615. [PubMed: 17568968]
- Sanchez-Craig M, Wilkinson DA, Davila R. Empirically based guidelines for moderate drinking: 1-Year results from three studies with problem drinkers. *American Journal of Public Health*. 1995; 85:823–828. <http://dx.doi.org/10.2105/AJPH.85.6.823>. [PubMed: 7762717]
- Schuckit MA, Smith TL. The relationships of a family history of alcohol dependence, a low level of response to alcohol and six domains of life functioning to the development of alcohol use disorders. *Journal of Studies on Alcohol and Drugs*. 2000; 61:827.
- Schuckit MA, Smith TL, Danko GP, Trim R, Bucholz KK, Edenberg HJ, et al. An evaluation of the full level of response to alcohol model of heavy drinking and problems in COGA offspring. *Journal of Studies on Alcohol and Drugs*. 2009; 70:436. [PubMed: 19371495]
- Searles JS, Helzer JE, Walter DE. Comparison of drinking patterns measured by daily reports and timeline follow back. *Psychology of Addictive Behaviors*. 2000; 14:277–286. [PubMed: 10998953]
- Shiffman S. Ecological momentary assessment (EMA) in studies of substance use. *Psychological Assessment*. 2009; 21:486–497. <http://dx.doi.org/10.1037/a0017074>. [PubMed: 19947783]
- Shillington AM, Clapp JD. Self-report stability of adolescent substance use: Are there differences for gender, ethnicity and age? *Drug and Alcohol Dependence*. 2000; 60:19–27. [PubMed: 10821986]
- Simons JS, Dvorak RD, Batien BD, Wray TB. Event-level associations between affect, alcohol intoxication, and acute dependence symptoms: Effects of urgency, self-control, and drinking experience. *Addictive Behaviors*. 2010; 35:1045–1053. <http://dx.doi.org/10.1016/j.addbeh.2010.07.001>. [PubMed: 20685044]
- Simons JS, Wills TA, Neal DJ. The many faces of affect: A multilevel model of drinking frequency/quantity and alcohol dependence symptoms among young adults. *Journal of Abnormal Psychology*. 2014; 123:676–694. <http://dx.doi.org/10.1037/a0036926>. [PubMed: 24933278]
- Skinner, HA.; Horn, JL. Alcohol dependence scale: User's guide. Toronto: Addiction Research Foundation; 1984.
- Sobell LC, Brown J, Leo GI, Sobell MB. The reliability of the alcohol timeline followback when administered by telephone and by computer. *Drug and Alcohol Dependence*. 1996; 42:49–54. [PubMed: 8889403]
- Sobell, LC.; Sobell, MB.; Litten, RZ.; Allen, JP. Timeline follow-back: A technique for assessing self-reported alcohol consumption Measuring alcohol consumption: Psychosocial and biochemical methods. Totowa, NJ US: Humana Press; 1992. p. 41–72.
- Statacorp. Stata statistical software: Release 13. College Station, TX: Stata Corporation; 2013.
- Swift R. Transdermal alcohol measurement for estimation of blood alcohol concentration. *Alcoholism: Clinical and Experimental Research*. 2000; 24:422–423.
- Tempelman LA, Moeller MP, Schmitt EW. Wireless, low maintenance transdermal alcohol sensor — Phase II final report. 2010(R44 AA014118-04)
- Trull TJ, Ebner-Priemer UW. Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: Introduction to the special section. *Psychological Assessment*. 2009; 21:457–462. (2009-22537-001 [pii]). [PubMed: 19947780]
- Weiss, HM.; Beal, DJ.; Lucy, SL.; Macdermid, SM. Constructing EMA studies with PMAT: The Purdue momentary assessment tool user's manual. 2004. from <http://www.mfri.purdue.edu/pages/PMAT/pmatusermanual.pdf>
- Whitfield JB, Zhu G, Duffy DL, Birley AJ, Madden PAF, Heath AC, et al. Variation in alcohol pharmacokinetics as a risk factor for alcohol dependence. *Alcoholism: Clinical and Experimental Research*. 2001; 25:1257–1263.
- Wills TA, Cleary SD. The validity of self-reports of smoking: Analyses by race/ethnicity in a school sample of urban adolescents. *American Journal of Public Health*. 1997; 87:56–61. [PubMed: 9065227]

Fig. 1. Prediction of drinks on drinking days from peak transdermal alcohol signal. Graph only includes the fixed effect, and hence random intercept will account for some of the observed deviation from the fitted line.

Fig. 2. Prediction of alcohol dependence symptoms as a function of peak concentration and time to reach peak concentration during drinking episodes.

Table 1

Descriptive data.

Source				
	Baseline	ESM	AM	TLFB
	52% women			WristAS
Gender				
Age	19.57 (<i>SD</i> = 1.00)			
Percent drinking days		34.62 ^a	34.62 ^a	36.17
Drinks per drinking day		N/A ^b	7.61 (<i>SD</i> = 5.06)	8.04 (<i>SD</i> = 5.47)
ADS	11.52 (<i>SD</i> = 5.39)			N/A

Note. *N* = 60, 647 days of monitoring. ADS = Alcohol Dependence Scale, ESM = the drinks from the random experience sampling, AM = number of drinks from morning reports, TLFB = timeline follow-back assessment.

^aPercent drinking days reflect whether an event was detected on either experience sampling or AM report.

^bNot reported as is not comparable due to being a random sample of time.

Table 2

Comparison of associations between transdermal assessment and three methods of self-reported drinking.

	ESM		AM		TLFB	
	OR/IRR	p-Value	OR/IRR	p-Value	OR/IRR	p-Value
<i>Abstain</i>						
Peak	0.93	<.001	0.80	<.001	0.82	<.001
SM peak	0.54	<.001	0.77	<.001	0.81	<.001
Gender	0.92	.031	1.00	.990	0.84	.534
<i>Drinks</i>						
Peak	1.02	<.001	1.02	<.001	1.03	<.001
SM peak	1.03	.008	1.03	<.001	1.03	<.001
Gender	1.17	.314	1.29	.016	1.38	.009

Note. ESM = the drinks from the random experience sampling. AM = number of drinks from morning reports, TLFB = timeline follow-back assessment. Models are multilevel negative binomial hurdle models. OR = odds ratio (for the prediction of zeros). IRR = incident rate ratio (for the prediction of counts). ESM ($n = 59$, $N = 463$), AM ($n = 59$, $N = 383$), TLFB ($n = 59$, $N = 455$). Ns are less than total due to missing data. Bold face rows are between subject effects.

Table 3

Associations between experience sampling assessment of drinks, transdermal AUC, and alcohol dependence symptoms.

	IRR	SE	<i>p</i>-Value
Drinks	1.17	0.06	0.01
AUC	1.01	0.00	<.001
SM drinks	1.71	0.17	0.002
SM AUC	1.02	0.00	<.001
Gender	0.74	0.28	0.284

Note. $N = 455$, $n = 59$. Daily dependence symptoms are the outcome. Bold face signifies between-person effects.

Table 4

Associations between peak concentration, time to reach peak concentration, and alcohol dependence symptoms.

	IRR	SE	p-Value
Peak concentration	1.003	0.001	.001
Time to reach peak	0.339	0.135	.006
Peak $\times$ time	1.008	0.004	.025
Gender	1.151	0.137	.236

Note. $N = 49$. $\chi^2(4, N = 49) = 10.26, p = .0363$