Published in final edited form as:

Alcohol Clin Exp Res. 2014 December ; 38(12): 3060–3067. doi:10.1111/acer.12589.

Adjustments for drink size and ethanol content: New results from a self-report diary and trans-dermal sensor validation study

J. C. Bond, T. K. Greenfield, D. Patterson, and W.C. Kerr

Alcohol Research Group, Public Health Institute, Emeryville CA 94608

Abstract

Background—Prior studies adjusting self-reported measures of alcohol intake for drink size and ethanol content have relied on single-point assessments.

Methods—A prospective 28-day diary study investigated magnitudes of drink ethanol adjustments and factors associated with these adjustments. Transdermal alcohol sensor (TAS) readings and prediction of alcohol-related problems by number of drinks versus ethanol-adjusted intake were used to validate drink ethanol adjustments. Self-completed event diaries listed up to 4 beverage types and 4 drinking events/day. Eligible volunteers had weekly drinking and 3+ drinks per occasion with 26 reported days and pre- and post-summary measures ($n = 220$). Event reports included drink types, sizes, brands or spirits contents, venues, drinks consumed and drinking duration.

Results—Wine drinks averaged 1.19, beer, 1.09 and spirits 1.54 US standard drinks (14g ethanol). Mean adjusted alcohol intake was 22% larger using drink size and strength (brand/ethanol concentration) data. Adjusted drink levels were larger than “raw” drinks in all quantity ranges. Individual-level drink ethanol adjustment ratios (ethanol adjusted/unadjusted amounts) averaged across all days drinking ranged from 0.73–3.33 (mean 1.22). Adjustment ratio was only marginally (and not significantly) positively related to usual quantity, frequency and heavy drinking (all $ps < .1$), independent of gender, age, employment, and education, but those with lower incomes (both $p < .01$) drank stronger/bigger drinks. Controlling for raw number of drinks and other covariates, degree of adjustment independently predicted alcohol dependence symptoms ($p < .01$) and number of consequences ($p < .05$). In 30 respondents with sufficiently high quality TAS readings, higher correlations ($p = .04$) were found between the adjusted vs. the raw drinks/event and TAS areas under the curve.

Conclusions—Absent drink size and strength data, intake assessments are downward biased by at least 20%. Between-subject variation in typical drink content and pour sizes should be addressed in treatment and epidemiological research.

Keywords

alcohol; self-report; measurement; transdermal alcohol sensor (TAS); validity

INTRODUCTION

Background

Recent methodological work on alcohol consumption measurement has shown the importance of adjusting self-reported alcohol intake for drink size and ethanol content (Greenfield and Kerr, 2008). Studies confirm that among inner-city pregnant women (Kaskutas, 2000) and other women at risk for pregnancy (Witbrodt et al., 2008), assessing drink sizes and alcohol content is of great importance for accurately gauging intake and, with drinking duration, estimating blood alcohol concentration (BAC). In general population based home drink measurement studies, larger drinks were found to be normative, particularly so for wine and especially spirits drinks with considerable between-subject variation (Kerr et al., 2005). In these protocols the respondent simulates pouring typical drinks at home and measures them with previously supplied graduated beakers. Ethnic differences in men's and women's drink sizes and strength (brand/ethanol concentration) have also been observed (Kerr et al., 2009a). Bar drinks have been similarly found to involve large variation in drink sizes but with mean ethanol content larger than a standard drink, again especially so for mixed spirits drinks and wine, with bar beer drinks larger compared to home beer drinks, because more draft beer is sold in bars, often in larger glassware (Kerr et al., 2008). In fact glass size but not shape was observed in the empirical bar-drink study to affect pour sizes (Kerr et al., 2009b). International studies have also shown the importance of attending to drink ethanol, varying greatly by sizes of container and beverage %ABV, such as studies in several Indian states (Nayak et al., 2008). However, such studies have almost all relied on single-point assessments, either by using glassware or pictures of glassware (Kaskutas and Graves, 2000) measuring beakers (Kerr et al., 2005), or both (Kaskutas and Kerr, 2008). As an example, one interesting survey application in an Australian national sample involved asking detailed information on yesterday's drinking including beverage types and glassware sizes and using pictured glassware and market share means to estimate alcohol content by beverage type (Stockwell et al., 2004). A subsequent, similar study found greater coverage for such an assessment compared to standard measures (Stockwell et al., 2008). The drink ethanol adjustments from the detailed yesterday measure have been used to calibrate intake pattern measures with longer reference periods, such as the 12-month Graduated Frequencies (GF) measure included in the same survey, to empirically adjust for the ethanol in yesterday's drinks, taken this as typical of an individual's drink types and sizes (Greenfield et al., 2009). Although single time-point measures provide valuable individual-level information on drink sizes, a sufficient time sample of events is needed to obtain a reliable estimate of an individual's average drink size (liquid volume) and strength (ethanol concentration), the combined effect of both of which are accounted for here. This is because many individuals drink multiple beverage types, in multiple venues, both on- and off-premise, and in varying containers and vessels, and may vary their drinking patterns over time. Often drinking follows a weekly cycle (Room et al., 2012) potentially affecting what beverages are drunk, where and with whom, all of which are likely to affect drink strengths and sizes (Greenfield and Kerr, 2008).

Study Aims

Collecting health data using diaries has a long history (Verbrugge, 1980). Here we use detailed 28-day prospective drinking diaries collected during a methodological self-report measurement study to (a) assess a given individual's mean drink ethanol content, (b) determine how this affects average intake volume, (c) examine individual demographic and personal characteristics that might influence an individual's mean drink ethanol, (d) for a subsample of person-drinking-events with simultaneous transdermal alcohol sensor data, assess whether the adjusted measures, rather than the raw number of 'drinks', better fit the physiological data, and (e) investigate whether the drink-ethanol adjustment factor applied to the raw reported number of drinks per day independently adds to the prediction of both dependence symptoms and number of alcohol-related consequences.

MATERIALS AND METHODS

Sample and Procedures

Data come from an R01 project on "Improving Self-Report Alcohol Consumption Measurement" (AA013309, T. Greenfield, PI) involving detailed 28-day daily drinking event diaries designed to assess drink sizes and strengths, validated in a subsample using a trans-dermal alcohol sensor. The sample included volunteers, either contacted by a random digit dial (RDD) local-area telephone survey ($n=3,025$ originally screened) or later, those recruited through advertisements placed on Craigslist, a web-based bulletin board. All respondents were screened for eligibility using a Computer Assisted Telephone Interview (CATI) and a 12-month GF measure (Greenfield, 2000) to have at least weekly drinking and at least one day with 3 or more (3+) drinks in the prior year. This criterion was used in an attempt to ensure study participants randomized to the TAS group (described below) reported a sufficient number of drinking days during the study so that validity analyses could compare diary with TAS readings. Those eligible by these minimal drinking criteria were told about the experimental tasks and, if agreeing to participate in the IRB-approved protocol with compensation of up to \$175, were randomized into one of three groups. Groups 1, 2 and 3 (i.e., all groups) completed by telephone pre- and post-retrospective summary measures (covering a 28-day period, see Measures). In addition, Group 2 ($n = 65$) completed drinking diaries each day, collected or mailed back weekly for 4 consecutive weeks (28 days), giving detailed information about up to 4 drinking occasions/day (see Measures which describes diary measures in detail). In addition to the diaries and pre- and post-surveys, Group 3 (and pilot participants, combined $n = 155$) also wore on their wrist a non-invasive physiological measuring device, the Wris-TAS™ (Versions 5 and 6, early in the experiment; version 7 in the final months) transdermal alcohol sensor (TAS) (Marques and McKnight, 2009; Swift, 2000), for 2 of the 4 weeks. For the analyses here we use self-report prospective diary data only from Groups 2, 3 and the Pilot study who also completed all measures and who provided valid data on at least 26 of the 28 diary days (total $n = 220$). Table 1 provides demographics and other characteristics for several groups used in the present analyses (see Results).

Measures

Graduated Frequencies (GF) summary measures were administered by CATI for comparison to the daily diary data. GF measures have been detailed elsewhere (Greenfield, 2000; Greenfield et al., 2009; Room, 1990; World Health Organization, 2000). Two versions of the GF were collected: the first assessing drinking over the past 12 months; the second assessing drinking over two 28-day periods (both prior to study entry as well as over the same 28-days assessed by the drinking diaries). The GF measure can yield volume (Rehm, 1998) in a way that accounts for heavy not just usual drinking occasions (Rehm et al., 1999).

2005 NAS 12-Month GF Screener—This consumption measure was adapted for use as an eligibility screener in this study. The initiating question (maximum) begins: “Think of all kinds of alcoholic beverages combined, that is any combination of bottles or cans of beer, glasses of wine, drinks containing liquor of any kind, or coolers, flavored malt beverages or pre-made cocktails. In this question, one drink is equal to a 12 ounce bottle of beer or cooler, a four to five ounce glass of wine, or one shot of liquor (1.5 ounces). During the last 12 months, what is the largest number of drinks you had on any single day? Was it...” Categories begin with 24 or more drinks, then 12-23, 8-11, 5-7, 4, 3, 2 drinks and finally 1 drink. (Previous maximum/day items used 3-4 (combined) and 1-2 drink categories (Greenfield et al., 2006).) Thus defined, “drinks” are the US standard, here taken as 14 g ethanol (Kerr et al., 2005; Turner, 1990). The reported maximum determines for which GF quantity levels the frequencies will be asked, beginning with (if relevant) 12+ drinks, then, in descending order, 8-11, 5-7, 3-4, 2 drinks, and 1 drink (frequency of 24 drinks is not asked). For each of the quantity ranges, a categorical frequency is asked: “Every day or nearly every day”, “3-4 times a week”, “1-2 times a week”, “1-3 times a month”, “Less than once a month”, “Once in those 12 months”, and “Never in those 12 months” (i.e., 7 categories). Algorithms to compute volume are adjusted if summations exceed 365 days (Greenfield et al., 2009). Practical concerns regarding the GF measure have been raised (Gmel et al., 2006; Graham et al., 2004), which were addressed during intensive interviewer training.

Pre- and Post-Diaries 28-day Summary GF: A new 28-day GF measure used the original quantity-range categories (12+, 8-11, 5-7, 3-4, 1-2 drinks). Following the maximum question, the frequency of drinking each amount was asked *in days* (of the prior 28 days). The CATI program required reconciliation with the interviewee if the total initially given exceeded 28 days, so no summations of days could exceed the 4-week (28-day) reference period. The pre period covered the 28-day period prior to study entry and was used, along with the 12-month GF, for in study screening criteria. The post 28-day GF measured covered the same 28-day period as assessed by the drinking diaries, as described next.

Self-Report Measurement Study 28-Day Diaries: The drinking diary consisted of one page per day (Day, Date, and ID Number were pre-entered), divided down the page into sections for up to 4 “Drinking Sessions” with data on each entered in three rows with 8 columns. Column headings from left to right were: Type of alcohol (W, B, S); # of Drinks (always record a whole number); Location (H=Home, R=Restaurant, B=Bar, P=Party, O=Other); Brand-Type (Examples: Bud Lite, Mondavi Chardonnay, Meyers Dark Rum);

Drink Size in Ounces (½, 1, 1 ½ 2 4 6 8 12 16 24 32 40); Finished Drink? (# of drinks not finished, % of drink left); Meal within 1 hr? (S=Snack, M=Meal, N=No). For each session, respondents recorded Time began and Time ended. A 30-minute training session introduced the diary and explained the meaning of the columns (e.g., how to figure out drink sizes, meaning of “session”, use of shorthand, etc.). The letters/numbers given above were pre-printed to allow circling. Data were mailed weekly in pre-paid envelopes, or collected. A significant monetary incentive was associated with completion of diaries and surveys with a bonus for full completion (up to a total \$175 for all tasks).

Drink Ethanol Adjustments and Adjustment Ratios: Analyses examining drink ethanol adjustments are based on diary data (Groups 2 and 3). For each drinking occasion, the unadjusted number of diary drinks was estimated as the sum of the raw number of reported ‘drinks’/day (assumed to each be 14 g ethanol); the adjusted number of diary drinks (expressed in numbers and decimals of standard drinks) corrected ethanol content consumed for each drink. Adjustment for drink ethanol was made by using self-reported brand, type and drink size data in the diaries and information from prior Center methodological studies which had assembled a database of detailed alcohol type and brand percent alcohol by volume (%ABV); for more details see (Kerr et al., 2005; Kerr et al., 2008). For each respondent, an adjustment ratio was estimated as the sum of the total number of adjusted diary drinks divided by the total number of unadjusted diary drinks.

12-month Alcohol Dependence Symptoms and Alcohol-related Consequences: These 12-month scales, used in a second validation analysis, were included in the screener. Developed earlier (Clark and Hilton, 1991) these “Hilton” scales (Hilton, 1988) assess current number of alcohol dependence symptoms (9 dichotomous items) or alcohol-related consequences (11 dichotomous items). Dependence symptoms include 3 loss of control items (e.g., “I deliberately tried to cut down but I was unable to do so”; 2 withdrawal items, e.g., I have taken a strong drink in the morning to get over the effects of last night's drinking”; and other items such as “I had awakened the next day not being able to remember some of the things I had done while drinking”, “I have skipped a number of regular meals while drinking”, and “I was afraid I might be an alcoholic.” Hilton (1988) reported acceptable internal scale reliabilities for dependence symptoms (Cronbach's alphas .74 to .76). The Consequences scale included social and health problems such as 3 people who'd like you to drink less or act differently when you drank (spouse, parents); 3 job-related problems, e.g., “I have lost a job or nearly lost one because of my drinking”; 3 legal problems, e.g., “A policeman questioned or warned me because of my drinking” and I have been arrested for driving after drinking”; and 2 health-related problems, e.g., I felt that my drinking was becoming a serious threat to my physical health. Hilton (1988) reported a range of reliabilities for the Consequences scale (alphas .56 to .69). For a complete list of both scale's items, see the Hilton's Appendix (Hilton, 1988, p 1427) and see further (Midanik and Greenfield, 2000).

Transdermal Alcohol Sensor—We used the Wris-TAS™ developed by Giner, Inc., of Newton, MA, a noninvasive wrist-worn device, to indirectly estimate blood alcohol level and by inference alcohol intake from ethanol eliminated from the human body via the skin in

perspiration (Leffingwell et al., 2013). Giner, Inc. has developed a patented electrochemical transdermal alcohol sensor for monitoring patterns of drinking objectively, conveniently and passively (Swift et al., 1992). Although the TAS units themselves are effectively prototypes and thus still in the test phase, results to date suggest a strong lagged (Swift and Swette, 1992) relationship with Blood Alcohol Level (BAL) and self-report (Swift et al., 1999; Swift et al., 1992). Therefore, although not considered here an infallible gold standard, the TAS is at a stage of development where it becomes important to include it as a further validity check on the diaries.

The basis of the sensor measurement is continuous electrochemical oxidation of alcohol (ethanol) to acetic acid at a fixed potential and measurement of the resultant current, which is proportional to the transdermal alcohol concentration (TAC). A foam seal captures a small space above the skin where alcohol vapor is measured. The device also monitors skin temperature and conductivity to confirm subject compliance with wearing the device. Data is downloaded via USB port with Excel files output for further analysis. For the validity analyses reported here, a simple estimate of amount consumed/day was used—the Area Under the Curve (AUC) in mg/dL from the recorded TAS output (TAC) associated with a particular drinking event and aggregated over 24 hours (Bond et al., 2009b). The WrisTAS™ sensor was worn for 2 of 4 weeks of diary keeping by Group 3 and Pilot study participants and used here to validate the ethanol adjustments under the assumption that the physiologically derived indicators of alcohol intake must be free of self-report biases (Bond et al., 2009b). A preliminary analysis indicated that when physiological measures agreed on a diary drinking event, “peak height measurements” appeared to correspond with small or large [diary quantities]” (Greenfield et al., 2005). Among those completing diaries and GF measures, Group 3 and the Pilot study group also wore the WrisTAS™ for 2 weeks, used in validity analyses. Figure 1 provides an example TAS recording (TAC).

Selection of Cases for use in TAS-Diary Validity Analyses: A subsample ($n = 30$ participants) with high quality TAS recordings (among WrisTAS wearers) was chosen for validating the drink ethanol adjustment. To assure a meaningful analysis, the sample was limited to cases having ‘higher quality’, face-valid recordings implying adequate TAS functioning given not-uncommon problems with equipment reliability (Greenfield et al., 2005). High quality TAC curves were indicated by an alcohol signal that rose and fell in a way plausibly indicative of likely (albeit delayed) BAC (see Figure 1). A large variety of TAS alcohol signal behaviors were observed that could not plausibly be associated with an underlying BAC or that limited the ability to estimate peaks or AUCs. Such issues included erratic signals with many spikes or dips suggesting device or seal problems, removal of the device (as indicated by drops in thermistor temperature), overall high variability in the TAS signal precluding valid TAC analysis, and other problematic signal characteristics clearly indicating sensor failure. For each drinking event, we measured area under the TAC curve (AUCs) for each day the TAS was worn.

These 30 participants yielded 149 person-days of diary data for which drinking was reported in the diary and where the TAC curve also indicated that drinking occurred. Before recording area under the TAC curve for each of these days, some necessary adjustments were made to TAC curves to remove effects of alcohol signal problems as well as issues

with compliance (i.e., the wearer taking off the TAS unit). Such adjustments included (a) accounting for upward drift in the baseline signal in about 30% of the cases (e.g., in Figure 1, after 5/21/2004, the alcohol signal returns to a baseline somewhat above zero), (b) smoothing of spikes in the alcohol signal near the peak (typically associated with the wearer taking the TAS off for a short period of time, about 20% of the cases), and (c) smoothing areas of high alcohol signal variability (35% of the cases). Area under the TAC curve was estimated using the simple rectangle rule (i.e., sum of the curve heights) and area associated with drift (either upward or downward) in the baseline curve was subtracted.

Analyses: Within each of the GF quantity bands (1-2, 3-4, 5-7, 8-11, 12+ drinks), each diary drinking day used the unadjusted number of drinks to define the drinking level for that day. The average number of unadjusted and, separately, adjusted diary drinks were then estimated across all individuals and diary days combined, with the adjustment ratio within each GF level defined as the ratio of these two quantities. Linear regression was used to predict, at the individual level, the adjustment ratio (total number of adjusted divided by number of unadjusted diary drinks) as a function of demographics. Validity of drink ethanol adjustments was examined by assessing its incremental contribution to the prediction of current (12-month) alcohol dependence and alcohol-related consequences beyond raw diary reported numbers of drinks. In evaluating the effect of the drink ethanol adjustment on the alcohol-problem outcomes, it was important to include this volume measure both because of its well-established relationship to these outcomes (Caetano et al., 1997) but also since the adjustment is referenced to this variable. The criterion (dependent) variable in these models was (a) number of dependence symptoms and, separately, (b) number of consequences reported. Outcomes were predicted from the 'raw' reported number of drinks, and entering a difference measure ($= \text{Ethanol-adjusted drinks} - \text{Unadjusted or 'raw' drinks}$), as well as a range of covariates known to affect alcohol intake and alcohol-related problems. Finally, correlations were then estimated between the TAC area under the curve and a) unadjusted number of drinks and, separately, b) unadjusted number of drinks using data from all drinking events combined across usable TAS records (described below).

RESULTS

Demographic Comparisons Across Samples

Table 1 shows demographic distributions across several samples used in the present analyses. The first column provides rates for the n=30 sub-sample from the Pilot and Group 3 with good-quality TAS recordings, complete diary reports (≥ 26 of 28 days) and at least 2 drinking diary day reports; the second column for the n=220 with complete diary reports; the third of all those enrolled into groups 2, 3, and the pilot (n=250). For the first column, comparisons of rates were between the n=30 group indicated and the remainder sample of n=155 in Groups 3 and the Pilot (i.e., respondents who wore the TAS). For the second column, comparisons of rates were between the n=220 group shown and the remainder of those in the combined Groups 2, 3 and the Pilot (i.e., $n=250-220=30$). None of the distributions were significantly different at the .05 level.

Results Related to Drink-Ethanol Adjustment and Summary GF Measure

Ethanol content in drinks varied by beverage: wine drinks averaged 1.19, beer, 1.09 and spirits 1.54 times the ethanol content in a US standard drink (14 g). Overall, the sample's adjusted alcohol intake was 22% larger when accounting for the daily reports of drink sizes and strengths. One purpose of the study was to examine the drink ethanol adjustment by quantity levels of the GF measure (i.e., 1-2 drinks, 3-4, 5-7, 8-11 and 12+ drinks). For each level, all days on which quantities were reported in the given range (e.g. 1 or 2 drinks, etc.) were examined. The 'raw' (unadjusted) daily number of drinks provided an empirical distribution from which a mean of the GF quantity range is calculated. Following adjustments for drink ethanol calculated as described earlier, all daily quantity GF levels on average required similar upward adjustment from the 'raw' number of drinks when taking account of drink ethanol (excepting the 12+ level). Table 2 gives, for each GF quantity band (or level): (a) the arithmetic mean (e.g., the arithmetic mean of 5-7 drinks is 6 drinks); (b) the 'raw' diary mean (typically below the arithmetic mean because of the shape of the underlying drinking distribution); and (c) the drink ethanol adjusted mean (in all cases found to be above both other means). For example, a mean of 5-7 reported drinks yielded 7.36 ethanol-adjusted drinks instead of 5.79 (unadjusted), 8-11 drinks adjusted mean was 10.08 vs 9.02 standard drink equivalents, while the adjusted mean of 12+ drinks was 15.25 drinks.

At each of the lowest GF drink ranges, i.e., 1-2, 3-4 and 5-7, the difference between the 'raw' number of drinks and the drink-ethanol-adjusted mean number (each expressed in standard drinks) was highly significant using a paired sample t test (all $p < .001$). Table 2 also provides the adjustment ratios (adjusted mean number of drinks / unadjusted 'raw' drinks) for each of the GF quantity bands. Adjustment factors were between 1.30 and 1.26 for the levels from 1-2 to 5-7 drinks. At 8-11 drinks the difference fell just short of significance ($p = .07$) with an adjustment ratio of 1.12 (Table 2). In standard practice, computing volume based on the GF measure has relied on algorithms summing the GF's series of $Q \times F$ s (a) using the arithmetic mean for each Q level and (b) a conservative value of 13 drinks for the unbounded top threshold of 12+ drinks. Regarding (a), the diary distributions show that the 'raw' empirical mean is below the arithmetic mean (as plausible given the approximately log normal drinking distribution) but more importantly, the drink-ethanol adjusted mean, using drink size and strength information, is considerably above the arithmetic mean (see Table 2). Regarding (b), the values of the 'raw' and drink ethanol adjusted mean of the empirical distribution at 12 drinks and above is closely similar at 15.26 and 15.25 drinks, respectively. Both are substantially greater than the 13 drinks previously used. The above results are based on considering person days as the unit of analysis. We turn now to considering results with the individual as the unit of analysis.

Individual's Drink Ethanol Adjustments

Individual-level drink ethanol adjustment ratios (ethanol adjusted / unadjusted standard drink amount) averaged across all a person's drinks (over the 28-day period) ranged from 0.73 to 3.33 with a group mean of 1.22. This means that some individuals typically consumed 'drinks' that are more than 3 times a standard drink (14 g ethanol), while others consumed smaller drinks (as low as a mean of $\frac{3}{4}$ of a standard drink). To understand factors influencing mean drink strength, we estimated regression models to predict the within-

individual adjustment ratio. Independent variables included 12-month drinking pattern variables (usual quantity, frequency, and 5+ drinking days from the GF) and demographic factors (gender, age, Hispanic group, marital status, employment, family income (under \$30k/year vs. over), and education (< 4 year degree vs. 4 year degree or more), summarized in Table 3.

In the full model, controlling for all other factors, an individual's mean drink adjustment (i.e., averaged over 28 days) was marginally and positively (but not significantly) related to individual's mean quantity (drinks per drinking day), frequency of drinking and days of heavy drinking (all $ps < .10$). Size/strength adjustments were independent of many demographics such as gender, age, and Hispanic ethnicity but those who had incomes of \$30,000 or less ($p < .01$) drank stronger/bigger drinks. A parsimonious model eliminating several less significant factors confirmed that lower income earning individuals tended to have larger drinks with other characteristics such as typical quantity, heavy drinking (positively) and being married or partners (negatively) only marginally (but not significantly) involved ($p < .15$).

Predictive Validity of Diary Drink Ethanol Adjustments for Dependence and Consequences

To determine whether the drink-ethanol adjustment was important in predicting alcohol-related problems over and above average number of raw drinks/day, we separately estimated models predicting dependence symptoms and alcohol-related consequences, controlling for the same independent variables as in the full model in Table 3. Table 4 summarizes the regression models predicting numbers of alcohol dependence symptoms and alcohol-related consequences. Degree of adjustment independently predicted alcohol dependence symptoms ($p < .01$) with age ($p < .001$) and income ($p < .05$) also influential. For predicting number of consequences the degree of adjustment for drink ethanol was also influential ($p = .01$) after controlling for covariates with gender ($p < .05$) and, though not significantly age and income (both $p < .10$) also marginally influential. Based on R^2 , fit appears to be somewhat better when predicting dependence than consequences.

Predictive Validity of Diary Drink Ethanol Adjustments for the TAS

In the sub-sample of 30 participants, for each of the 274 total days for which there was either a diary record or a TAS curve available, only 214 (77%) of these days had a sufficiently discernable TAS-derived TAC curve permitting a drinking vs. no drinking assessment and for which the diary and TAS measurements corresponded to the same day (e.g., study complexity resulted in several cases that had 1 day at the beginning or ending of a TAS week with no diary report or vice versa). The sensitivity of the TAS to the diaries was 85.6%; that is, 85.6% of the 174 days for which the diary indicated drinking took place, the TAS also indicated a drinking event. Similarly, the specificity of the TAS to the diaries was 67.5%; i.e., on 67.5% of the 40 days during which the diaries indicated no drinking took place the TAS also indicated no drinking occurred. Combined overall across all person-days, the correlation between the AUC and the unadjusted number of drinks was .62 whereas the correlation between AUC and the adjusted number of drinks was .73 (difference between correlation coefficient estimates $p = .04$). This difference corresponds approximately to a medium effect size difference between correlations (Cohen, 1988).

DISCUSSION

As indicated by these findings, alcohol intake measures such as the Graduated Frequencies scale (Greenfield, 2000), if they do not attend to drink size and strength, are biased downward by at least 20% in a population of slightly heavier drinking individuals—those meeting the inclusion criteria of at least weekly drinking and ever consuming 3+ in a day during the prior 12 months. This result is similar in degree to findings from an Australian national population survey also using a version of the GF measure (Greenfield et al., 2009). Moreover, we found wide between-subject variation in typical drink content both over time and on average between individuals. These results extended the usual-pour findings of a number of home- and bar-drink size/strength methodological studies conducted recently by our group (Kerr et al., 2005; Kerr et al., 2008). It is gratifying that the results on wine beer and spirits alcohol content were closely similar to those found in beaker-measured studies, which in the bar study also assayed mixed drink alcohol content using an Analox® analyzer (Kerr et al., 2008). The extensive dataset on %ABV of brands and types of drinks developed from those studies was used in converting respondents' daily reports of specific drinks and their sizes in this 28-day study. What is new here is the 4-week time sample by individual of up to 4 drinking events per day (separate events defined mainly by venues such as at home, at a party, or out at a bar, for example). As in the earlier studies, wine drinks were about 20% larger and spirits drinks one-and-a-half times bigger than a standard drink. One limitation of this study is that although the data were from daily drinking diaries, less subject to long-term memory distortions (Greenfield and Kerr, 2008), the numbers of drinks, their sizes, and brand/types relied on self-reports. Here we established that the intake adjusted for drink ethanol was more correlated with (a) AUCs from the physiological TAS measure, taken as the “gold” (or at least unbiased) standard, and with (b) two alcohol problem measures, numbers of dependence symptom and consequences (controlling for age and gender). These statistically significant findings support the validity of the strategy of adjusting for drink ethanol based on drink size and specific brand/type data, in this case recorded on an event-specific basis over 4 weeks. The amount of ethanol intake (rather than the mere number of drinks) *should* conceptually be a better predictor of both an objective indicator associated with BAC and negative outcomes.

We were also able to develop empirical adjustments to the Graduated Frequencies (GF) measure's quantity bands that account on average for drink ethanol content (Greenfield et al., 2009). The result was a 22% increment to the calculated mean volume of the full sample of those completing at least 26 diary days. One limitation is that these results, by design, apply to the eligible sample made up of individuals who on average were somewhat heavier drinkers than US drinkers in a general population (see Table 1). These results reinforce that pour sizes and specific types of drinks should be accounted for to improve treatment and epidemiological measures. Much of the epidemiological literature is based on alcohol consumption measures assuming standard drinks, and one should be aware that their intake assessments are likely to be downwardly biased on average (Greenfield et al., 2009). By broadening the time sample, present results complement findings from prior methodological analyses of home (Kerr et al., 2005; Kerr et al., 2009a) and bar drinks (Kerr et al., 2008), based on one-time assessments of typical behavior. Both types of findings will allow the GF

measures used in our National Alcohol Surveys (and potentially other beverage-specific alcohol intake assessments) to be adjusted for drink types and pour sizes resulting in less downwardly biased alcohol measurement sizes (Greenfield and Kerr, 2008). Further work to increase measurement efficiency (such as employing smart phones for ecological momentary assessment) will be beneficial and is much needed. The challenge is to efficiently but sufficiently accurately attend to drink sizes and strengths in event-based assessments. We are well on the way to ushering in new self-report alcohol intake measurement strategies that can substantially improve epidemiological and clinical studies by providing real-time monitoring of drinking behavior.

Acknowledgement

Supported by a grant to the National Institute on Alcohol Abuse and Alcoholism to the Public Health Institute (R01 AA013309) and by its Center Grant (P50 AA005595) for both of which the principal investigator was the second author (Greenfield). The views represented are those of the authors and not necessarily reflect those of the sponsoring institutions. A previous version of this paper was presented at the 37th Annual Alcohol Epidemiology Symposium of the Kettil Bruun Society for Social and Epidemiological Research on Alcohol, Melbourne, Australia: April 11-15; 2011.

References

- Bond, J.; Greenfield, TK.; Patterson, D.; Kerr, WC. Alcoholism: Clinical and Experimental Research. Vol. 33. San Diego, CA: Jun 20-24. 2009a Temporal variation in reporting of ethanol adjusted compared to unadjusted drinking diaries over a 28-day period in Research Society on Alcoholism. [Abstract] Suppl. 63A
- Bond, J.; Greenfield, TK.; Patterson, D.; Kerr, WC.; Lapp, WM. Alcoholism: Clinical and Experimental Research. Vol. 33. San Diego, CA: Jun 20-24. 2009b Expected intoxication levels after adjusting for drink size, alcohol content and characteristics of drinkers: evidence from Transdermal Alcohol Sensors, in Research Society on Alcoholism. [Abstract] Suppl. 63A
- Caetano R, Tam TW, Greenfield TK, Cherpitel CJ, Midanik LT. DSM-IV alcohol dependence and drinking in the U.S. population: a risk analysis. *Annals of Epidemiology*. 1997; 7(8):542–549. [PubMed: 9408550]
- Clark, WB.; Hilton, ME., editors. Alcohol in America: Drinking practices and problems. State University of New York Press; Albany, NY: 1991.
- Cohen, J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed.. Lawrence Erlbaum Associates; Hillsdale, NJ: 1988.
- Gmel G, Graham K, Kuendig H, Kuntsche S. Measuring alcohol consumption – should the ‘graduated frequency’ approach become the norm in survey research? *Addiction*. 2006; 101(1):16–30. [PubMed: 16393189]
- Graham K, Demers A, Rehm J, Gmel G. Problems with the graduated frequency approach to measuring alcohol consumption: results from a pilot study in Toronto, Canada. *Alcohol Alcohol*. 2004; 39(5):455–462. [PubMed: 15289203]
- Greenfield TK. Ways of measuring drinking patterns and the difference they make: experience with graduated frequencies. *J Subst Abuse*. 2000; 12(1):33–49. [PubMed: 11288473]
- Greenfield TK, Kerr WC. Alcohol measurement methodology in epidemiology: recent advances and opportunities. *Addiction*. 2008; 103(7):1082–1099. [PubMed: 18422826]
- Greenfield TK, Kerr WC, Bond J, Ye Y, Stockwell T. Improving graduated frequencies alcohol measures for monitoring consumption patterns: results from an Australian national survey and a US diary validity study. *Contemp Drug Prob*. 2009; 36(3/4):705–733.
- Greenfield TK, Nayak MB, Bond J, Ye Y, Midanik LT. Maximum quantity consumed and alcohol-related problems: assessing the most alcohol drunk with two measures. *Alcohol Clin Exp Res*. 2006; 30(9):1576–1582. [PubMed: 16930220]

- Greenfield, TK.; Tujague, J.; Bond, J.; Kerr, WC. Alcoholism: Clinical and Experimental Research. Vol. 29. Santa Barbara, CA: Jun 25–29. 2005 A pilot study of self-report alcohol consumption: diary versus transdermal alcohol sensor accuracy, in Research Society on Alcoholism. [Abstract] 130A
- Hilton ME. Trends in drinking problems and attitudes in the U.S.: 1979-1984. *Br J Addict.* 1988; 83(12):1421–1427. [PubMed: 3266087]
- Kaskutas LA. Understanding drinking during pregnancy among urban American Indians and African Americans: health messages, risk beliefs, and how we measure consumption. *Alcohol Clin Exp Res.* 2000; 24(8):1241–1250. [PubMed: 10968664]
- Kaskutas LA, Graves K. An alternative to standard drinks as a measure of alcohol consumption. *J Subst Abuse.* 2000; 12(1-2):67–78. [PubMed: 11288475]
- Kaskutas LA, Kerr WC. Accuracy of photographs to capture respondent-defined drink size. *J Stud Alcohol Drugs.* 2008; 69(4):605–610. [PubMed: 18612577]
- Kerr WC, Greenfield TK, Tujague J, Brown S. A drink is a drink? Variation in the alcohol content of beer, wine, and spirits drinks in a U.S. methodological sample. *Alcohol Clin Exp Res.* 2005; 29(11):2015–2021. [PubMed: 16340459]
- Kerr WC, Patterson D, Greenfield TK. Differences in the measured alcohol content of drinks between black, white and Hispanic men and women in a US national sample. *Addiction.* 2009a; 104(9): 1503–1511. [PubMed: 19438419]
- Kerr WC, Patterson D, Koenen MA, Greenfield TK. Alcohol content variation of bar and restaurant drinks in Northern California. *Alcohol Clin Exp Res.* 2008; 32(9):1623–1629. [PubMed: 18616674]
- Kerr WC, Patterson D, Koenen MA, Greenfield TK. Large drinks are no mistake: glass size, not shape, affects alcoholic beverage drink pours. *Drug Alcohol Rev.* 2009b; 28(4):360–365. [PubMed: 19594789]
- Leffingwell TR, Cooney NJ, Murphy JG, Luczak S, Rosen G, Dougherty DM, Barnett NP. Continuous objective monitoring of alcohol use: twenty-first century measurement using transdermal sensors. *Alcohol Clin Exp Res.* 2013; 37(1):16–22. [PubMed: 22823467]
- Marques PR, McKnight AS. Field and laboratory alcohol detection with 2 types of transdermal devices. *Alcohol Clin Exp Res.* 2009; 33(4):703–711. [PubMed: 19170663]
- Midanik L, Greenfield TK. Trends in social consequences and dependence symptoms in the United States: the National Alcohol Surveys, 1984-1995. *Am J Public Health.* 2000; 90(1):53–56. [PubMed: 10630137]
- Nayak MB, Kerr WC, Greenfield TK, Pillai A. Not all drinks are created equal: implications for alcohol assessment in India. *Alcohol Alcohol.* 2008; 43(6):713–718. [PubMed: 18832137]
- Rehm J. Measuring quantity, frequency, and volume of drinking. *Alcohol Clin Exp Res.* 1998; 22(Suppl. 2):4S–14S. [PubMed: 9603301]
- Rehm J, Greenfield TK, Walsh G, Xie X, Robson L, Single E. Assessment methods for alcohol consumption, prevalence of high risk drinking and harm: a sensitivity analysis. *Int J Epidemiol.* 1999; 28(2):219–224. [PubMed: 10342682]
- Room, R. Measuring alcohol consumption in the United States: methods and rationales. In: Kozlowski, LT.; Annis, HM.; Cappell, HD.; Glaser, FB.; Goodstadt, MS.; Israel, Y.; Kalant, H.; Sellers, EM.; Vingilis, ER., editors. *Research Advances in Alcohol and Drug Problems.* Vol. 10. Plenum Press; New York: 1990. p. 39-80.
- Room R, Mäkelä P, Benegal V, Greenfield TK, Hettige S, Tumwesigye NM, Wilsnack R. Times to drink: cross-cultural variations in drinking in the rhythm of the week. *Int J Public Health.* 2012; 57(1):107–117. [PubMed: 21553132]
- Stockwell T, Donath S, Cooper-Stanbury M, Chikritzhs T, Catalano P, Mateo C. Under-reporting of alcohol consumption in household surveys: a comparison of quantity-frequency, graduated-frequency and recent recall. *Addiction.* 2004; 99(8):1024–1033. [PubMed: 15265099]
- Stockwell T, Zhao J, Chikritzhs T, Greenfield TK. What did you drink yesterday? Public health relevance of a recent recall method used in the 2004 Australian National Drug Strategy Household Survey. *Addiction.* 2008; 103(6):919–928. [PubMed: 18482414]

- Swift R, Bird C, Santa Ines M, Almeida A, Davidson D. A comparison of transdermal alcohol detection and self report for determining alcohol consumption. *Alcohol Clin Exp Res*. 1999; 23(5): 129A. [Abstract].
- Swift RM. Transdermal alcohol measurement for estimation of blood alcohol concentration. *Alcohol Clin Exp Res*. 2000; 24(4):422–423. [PubMed: 10798575]
- Swift RM, Martin CS, Swette L, LaConti A, Kackley N. Studies on a wearable, electronic, transdermal alcohol sensor. *Alcohol Clin Exp Res*. 1992; 16(4):721–725. [PubMed: 1530135]
- Swift, RM.; Swette, L. Assessment of ethanol consumption with a wearable, electronic, transdermal alcohol sensor/recorder. In: Allen, J.; Litten, R., editors. *Methods to Assess Alcohol Consumption*. Humana Press; Totowa, NJ: 1992. p. 189-202.
- Turner C. How much alcohol is in a 'standard drink'? An analysis of 125 studies. *Br J Addict*. 1990; 85:1171–1175. [PubMed: 2224197]
- Verbrugge LM. Health diaries. *Med Care*. 1980; 18(1):73–95. [PubMed: 6986517]
- Witbrodt J, Kaskutas LA, Korcha R, Armstrong MA. Under-estimation of alcohol consumption among women at-risk for drinking during pregnancy. *Contemp Drug Prob*. 2008; 35(1):37–58.
- World Health Organization. *International Guide for Monitoring Alcohol Consumption and Related Harm*. World Health Organization, Department of Mental Health and Substance Dependence, Noncommunicable Diseases and Mental Health Cluster; Copenhagen, Denmark: 2000.

Figure 1.

Table 1Sample Characteristics, N=220^a

Variable	Analyzed TAS Cases (n=30) ^b	Analyzed Diary Cases (n=220) ^c	All Diary cases (n=250) ^d
Gender			
Female	.31	.42	.41
Male	.69	.58	.59
Ethnicity			
White	.86	.70	.69
Black	.00	.04	.04
Hispanic	.07	.15	.16
Other	.07	.11	.11
Age			
18-29	.21	.25	.26
30-39	.28	.22	.23
40-49	.17	.20	.20
50+	.34	.33	.30
Marital Status			
Not Married	.60	.50	.52
Married/ Living with Partner	.28	.36	.34
Missing	.12	.15	.16
Education			
< 4 Year Degree	.30	.37	.35
≥ 4 Year Degree	.70	.63	.65
Employment Status			
Not Employed	.22	.24	.22
Employed	.65	.61	.60
Missing	.13	.15	.18
Income			
30K	.24	.22	.22
> 30K	.70	.73	.71
Missing	.06	.05	.07
Current Alcohol Problems			
2+ Dependence Symptoms	.35	.28	.30
2+ Consequences	.10	.11	.12
Drinking Pattern 12 Month Means:			
GF Usual Frequency/Year	275	263	264
GF Usual Quantity/Past Year	3.07	2.92	3.00
GF Days 5+/Year	48	43	46
GF Days 8+/Year	22	21	22

*Tests for significant differences between groups defined in ^b and ^c above were conducted but none were significant at the .05 level

^aBased on screener data; drinking patterns use 12-Month Graduated Frequency (GF) measure

^b Comparisons between n=30 with high quality TAS and n=125 with lower quality TAS units

^c Comparisons between n=220 with ≥ 26 diary days and n=30 with < 26 diary days

Table 2

Graduated Frequency (GF) Drink Quantity Bands: Arithmetic, Empirical and Ethanol Adjusted Means, and Adjustment Ratio (Adjusted Mean/Empirical Mean), N=220^a

GF Quantity Level (drinks)	Arithmetic Mean	Empirical Mean	Adjusted Mean ^a	Ratio of Adjusted to Empirical Mean
1-2	1.5	1.47	1.90	1.30
3-4	3.5	3.38	4.26	1.26
5-7	6.0	5.79	7.36	1.27
8-11	9.5	9.02	10.08	1.12
12+	-	15.26	15.25	1.00

^a Adjusted Mean takes account of drink ethanol based on size and strength data in diaries (expressed in the metric of 14 g US standard drinks)

Table 3

Summary of Regression Models Predicting Mean Drink Ethanol Adjustment Ratio from Demographics,
N=178

Factor	B ^a	SE	Beta ^b	t	p value
<u>Full Model:</u>					
Constant	1.13	.17	--	6.70	<<.001
GF Usual Quantity/Past Year	.05	.03	.27	1.63	.095
GF Usual Frequency/Year	.003	.001	.09	1.86	.063
GF Days 5+/Year	.001	.001	.32	1.81	.072
Male Gender	.03	.06	.04	.47	.64
Age	.002	.002	.07	.85	.42
Hispanic vs Other	-.05	.08	-.04	-.57	.57
Married/Living with Partner	-.10	.06	-.13	-1.67	.09
Employed	.01	.07	.01	.13	.89
Income \$30,000 or less	.21	.08	.22	2.76	.006
Education < 4Y Degree	-.01	.06	-.11	-.14	.89
F(10,167) = 2.73 p = .004		R ² = .140			
<u>Parsimonious Model:</u>					
Constant	1.22	.11	--	11.31	<<.001
GF Usual Quantity/Past Year	.05	.03	.24	1.51	.13
GF Usual Frequency/Year	.001	.001	.09	1.20	.23
GF Days 5+/Year	.001	.001	.27	1.65	.11
Married/Living with Partner	-.10	.06	-.12	-1.67	.10
Income \$30,000 or less	.19	.07	.20	2.75	.007
Education < 4Y Degree	-.02	.06	-.03	-.38	.71
F(6,173) = 2.85 p = .005		R ² = .112			

^a Unstandardized regression coefficient

^b Standardized Regression Coefficient

Table 4

Summary of Regression Models Predicting (a) Alcohol Dependence Symptoms and (b) Alcohol-related Consequences: Effect of Drink-Ethanol Adjustment in Each. N=178

Factor	B ^b	SE	Beta ^c	t	p value
<i>Alcohol Dependence Symptoms</i>					
Constant	1.65	.51	--	3.23	.001
Male Gender	.10	.22	.03	.46	.65
Age	-.03	.01	-.25	-3.59	<<.001
Hispanic vs Other	.20	.29	.05	.71	.48
Married/Living with Partner	-.03	.22	-.01	-.13	.89
Employed	-.38	.25	-.10	-1.54	.13
Income \$30,000 or less	.57	.28	.14	2.03	.04
Education < 4Y Degree	.19	.23	.06	-.85	.40
Raw Mean Drinks/Day	.012	.002	.352	5.30	<<.001
Drink-Ethanol Delta ^a	.011	.004	.177	2.66	.009
F(9,178) = 9.87 p <<.001		R ² = .345			
<i>Alcohol-related Consequences</i>					
Constant	.41	.28	--	1.49	.14
Male Gender	.27	.12	.16	2.33	.02
Age	-.02	.01	-.15	-1.92	.06
Hispanic vs Other	.03	.16	.01	.18	.86
Married/Living with Partner	.13	.12	.07	1.06	.29
Employed	-.18	.13	-.10	-1.32	.19
Income \$30,000 or less	.27	.15	.14	1.76	.08
Education < 4Y Degree	-.07	.12	-.04	-.59	.56
Raw Mean Drinks/Day	.003	.001	.201	2.77	.006
Drink-Ethanol Delta ^a	.006	.002	.186	2.56	.011
F(9, 178) = 5.28 p <<.001		R ² = .178			

^a Ethanol-adjusted number of 14 g drinks/day equivalents minus raw reported number of drinks/day.

^b Unstandardized regression coefficient

^c Standardized Regression Coefficient