Sources of Misspecification Bias in Assessments of Risks Related to Alcohol Use

PAUL J. GRUENEWALD, PH.D.,^{a,*} MEME WANG-SCHWEIG, PH.D.,^a & CHRISTINA MAIR, PH.D.^b

^aPrevention Research Center, Oakland, California

^bDepartment of Behavioral and Community Health Sciences, University of Pittsburgh Graduate School of Public Health, Pittsburgh, Pennsylvania

ABSTRACT. Objective: Many different measures of alcohol use are applied in survey-based epidemiological studies of alcohol-related risks. Differences in the selection of drinking measures and alternative specifications of quantitative relationships of these measures to problem outcomes limit researchers' abilities to compare and assess alcohol effects across studies. We used a quantitative definition of *drinking patterns* to identify relationships among drinking measures and uncover sources of bias in assessments of drinking risks. **Method:** A census of drinking measures from studies published in four leading journals in the first half of 2013 were mapped onto a "drinking patterns table," quantitatively relating each measure to every other. Relationships among these measures and in relation to two problem outcomes, physiological problems and sexual risks, were examined using data from 41,352 undergraduate college student drinkers in California. **Results:** Twenty-nine sets of drinking

measures appeared across 74 published studies; no common statistical procedure was used to assess drinking risks. Empirically observed distributions of heavy drinking ($R^2 = .887, p < .001$) and variances in drinking quantities ($R^2 = .645, p < .001$) were predicted from the drinking patterns table. Heteroscedasticity in drinking measures also biased estimates of physiological risks related to drinking quantities ($z = -5.159, p < .001$), volume ($z = 4.592, p < .001$), and heavy drinking ($z = -5.431, p < .001$). **Conclusions:** Relationships between drinking measures can be formally identified and related to one another using drinking patterns tables. Biases related to selections of different drinking measures and unobserved heteroscedasticity can be identified and controlled through formal quantitative assessments of relationships between drinking measures and observed outcomes. (*J. Stud. Alcohol Drugs*, 77, 802–810, 2016)

IN THIS ARTICLE, WE ADDRESS a major theoretical concern with survey-based assessments of health risks related to alcohol use: The great diversity of drinking measures used to examine risks leads to incomparable estimates of alcohol effects. Many sources of error arise in epidemiologic studies of alcohol use and problems. Although survey-based measures of drinking patterns have greatly improved since the 1940s (O'Malley, 2014), these improvements have largely focused on reducing measurement error. In addition to measurement error are specification biases that result from the application of different measures of use to assessments of risks for the same or different problem outcomes. This leads to different estimates of risks related to "drinking." Not only do metrics differ between measures (e.g., daily quantity vs. monthly volume), but also correlations among drinking measures (e.g., frequency and volume) lead each to mark for the other in assessments of drinking risks. Although alcohol epidemiology can establish nominally significant correlations between all measures of use and problems, theoretical and practical advancements in the field are limited by the use of many different drinking measures and models for the assessment of alcohol risks.

Measuring drinking and problems

A sense of the scope of this problem is provided by a census of publications in four leading journals in the first half of 2013 (*Addiction*, *Alcoholism: Clinical and Experimental Research*, the *Journal of Studies on Alcohol and Drugs*, and *Prevention Science*). As shown in Figure 1, 74 studies collected one or more drinking measures from human subjects.

Among the 67 studies that collected interval measures, 28 gathered a single measure and 39 collected two or more measures (indicated in associated boxes, see Figure 1). The most common single measure was whether the subject had used alcohol in his or her lifetime or the past year ("Use," 10 studies). The second most common measure was volume used over some period of time ("V," 8 studies). The most common pair of measures was drinking frequency ("F") and average or typical quantity ("Q," 10 studies). Six studies included *F*, *Q*, and a measure of the frequency or prevalence of heavy drinking, " $F(Q > k)$," with the heavy drinking criterion, *k*, variously defined (e.g., $k = 3$ for women and $k = 4$ for men, 6 studies). The remaining studies included up to six different drinking measures along with assessments of maximum drinking levels ("Max *Q*"). Seven studies categorized drinkers into groups: four used a priori guidelines provided by the National Institute on Alcohol Abuse and Alcoholism (NIAAA), two used latent class analyses to statistically define a posteriori drinking classes, and one used average quantity to identify heavy drinkers. In all, 29 different collections

Received: September 15, 2015. Revision: February 26, 2016.

Research and preparation of this manuscript were supported by National Institute on Alcohol Abuse and Alcoholism Research Center Grant Number P60-AA06282 to Paul J. Gruenewald.

*Correspondence may be sent to Paul J. Gruenewald at the Prevention Research Center, 180 Grand Avenue, Suite 1200, Oakland, CA 94612, or via email at: paul@prev.org.

FIGURE 1. A review of drinking measures used in 74 articles published from 1/2013 through 6/2013 from four major scientific journals. *F* = drinking frequency; *Q* = quantity consumed on a drinking occasion; *V* = drinking volume; *F(Q > k)* = frequency of heavy drinking; *Max Q* = maximum drinking quantity. NIAAA = National Institute on Alcohol Abuse and Alcoholism.

of drinking measures were used, making it very difficult to compare alcohol effects across studies.

This situation substantively limits our ability to interpret what we label as “alcohol effects.” Thus, the claim that more alcohol-related problems may arise because “males drink more than females” might mean (a) more men drink, (b) men drink more frequently, (c) men drink greater volumes of alcohol, (d) more men exceed heavy drinking guidelines, or (e) more men are high-risk drinkers according to NIAAA guidelines. Using an example from our census, we statistically linked measures of lifetime, annual, and frequent alcohol use to risks for unprotected sex, but the formal relationships of these measures to one another and to incidents of unprotected sex remain unknown (Berger et al., 2013; Snipes & Benotsch, 2013).

What is a “drinking pattern”?

When different studies use different drinking measures to predict drinking outcomes, these different specifications lead to different estimates of alcohol effects. This can lead to some confusion as to the degree to which alcohol is involved in personal and social problems. One obvious way to resolve this issue is to enforce a common measurement sys-

tem across studies. A second more useful way is to formally relate drinking measures one to the other to facilitate comparisons across studies. This can be achieved by developing a theoretical definition of *drinking patterns* that identifies formal relationships among drinking measures with testable implications for statistical models of drinking risks.

For this purpose, we define *drinking patterns* as the number of different ways in which a volume of alcohol (*V*) can be consumed over drinking occasions (*F*) (ignoring inter-temporal intervals and the order of drinking events). *V* and *F* are measured by total number of drinks (*V*) and drinking days (*F*). A drinker who consumed *V* = 4 drinks over *F* = 1, 2, 3, or 4 drinking days could do so in five different ways: (a) 4 drinks on 1 day, (b) 3 drinks on 1 day and 1 drink on another, (c) 2 drinks on each of 2 days, (d) 2 drinks on 1 day and 1 drink on each of 2 days, and (e) 1 drink on each of 4 days. These five patterns are the integer partitions of 4; $4 = 3 + 1 = 2 + 2 = 2 + 1 + 1 = 1 + 1 + 1 + 1$. The number of integer partitions for any values of *V* and *F* can be computed and the number of drinking patterns for all possible combinations of *V* and *F* tabled, and this used to represent all the drinking measures gathered in our census of studies. Figure 2 shows a subset of this table for *V* = 1 to 14 drinks and *F* = 1 to 14 drinking days. The number of drinking patterns increases

		DRINKING FREQUENCY													
DRINKING VOLUME		1	2	3	4	5	6	7	8	9	10	11	12	13	14
	1	1													
	2	1	1												
	3	1	1	1											
	4	1	2	1	1										
	5	1	2	2	1	1									
	6	1	3	3	2	1	1								
	7	1	3	4	3	2	1	1							
	8	1	4	5	5	3	2	1	1						
	9	1	4	7	6	5	3	2	1	1					
	10	1	5	8	9	7	5	3	2	1	1				
	11	1	5	10	11	10	7	5	3	2	1	1			
	12	1	6	12	15	13	11	7	5	3	2	1	1		
	13	1	6	14	18	18	14	11	7	5	3	2	1	1	
	14	1	7	16	23	23	20	15	11	7	5	3	2	1	1
		Q = 4			Q = 3			Q = 2			Q = 1				

FIGURE 2. Drinking patterns table with cell entries indicating numbers of drinking patterns for combinations of V and F ; lines connect cells with common average drinking quantities, Q .

(roughly exponentially) with V and increases then decreases with F . Summing across F , there are seven ways of drinking $V = 5$ drinks, 42 ways of drinking $V = 10$ drinks, and 135 ways of drinking $V = 14$ drinks. For each V , a single drinking pattern occurs when $F = V$ and $F = 1$, with greater numbers of patterns appearing between these extremes. Other drinking measures can be represented as follows:

(a) *Average drinking quantities.* Each cell represents a combination of V and F such that $Q = V / F$ (connected by solid lines in Figure 2).

(b) *Maximum drinking quantity.* Each cell represents a combination of V and F such that $\text{Max } Q = V - F + 1$ (not shown).

(c) *Heavy or "binge" drinking.* Each cell contains a number of drinking patterns. These patterns can be listed, and patterns in which heavy drinking occurs, $Q > k$, can be enumerated. Assuming that choices to drink on every occasion are independent of one another (i.e., random), the probability of heavy drinking is proportional to the number of heavy drinking patterns weighted by the probability that each will occur. For example, the probability that more than 4 drinks will be consumed when drinking $V = 9$ drinks over $F = 5$ days is $(1 / F)^{V-F} \approx 0.002$. This small probability arises because the one pattern in which heavy drinking occurs, $9 = 1 + 1 + 1 + 1 + 5$, is unlikely to occur by chance alone (the most likely pattern is $9 = 2 + 2 + 2 + 2 + 1$). The complexity

FIGURE 3. Drinking patterns table showing expected probability of heavy drinking (left) and expected variance in drinking quantities (right).

of this calculation increases with F and V , so solutions are best obtained through simulation. Predicted probabilities of heavy drinking are shown to the left in Figure 3 for $k = 4$.

(d) *Variance in drinking quantities.* The quantities consumed on each of F drinking events will likely differ from $Q = V / F$, the average for all events. If $V = 9$ and $F = 5$, a high variance pattern would be $9 = 5 + 1 + 1 + 1 + 1$; a low variance pattern would be $9 = 2 + 2 + 2 + 2 + 1$. Variance can be computed for combinations of V and F using the same procedures as those for heavy drinking. These estimates are to the right in Figure 3.

(e) *Categorical measures.* Every categorical measure can be represented by a subset of cells in the drinking patterns table. A simplified version of the NIAAA guidelines might identify high-risk drinkers as those who drink an average of two or more drinks per drinking day with at least one heavy drinking event. This category is represented by cells below the $Q = 2$ isoline in Figure 2 with a probability of heavy drinking greater than 0.5 in Figure 3.

Two crucial facets of drinking measures are represented in Figures 2 and 3. First, average drinking quantities, heavy drinking, variances in drinking quantities, maximum drinking levels (not shown), and probabilities of high-risk drinking (e.g., from NIAAA guidelines, also not shown), all increase from top to bottom and right to left across the table. These

measures are correlated and will mark for one another in statistical analyses of drinking risks. Second, the determinate and stochastic aspects of drinking are distinguished by row/column labels and cell contents, respectively. For example, when $V = 9$ and $F = 5$, an average of $Q = 1.8$ drinks will be consumed on each of five occasions (determinate aspects). But some drinkers will consume about 2 drinks every time they drink (low variability), whereas others will consume about 4 drinks or more on some occasion (high variability). Greater numbers of patterns are thereby related to greater variations in drinking levels and drinking risks, a systematic source of risk heterogeneity related to F and V (and Q) and, therefore, also a potential source of bias related to heteroscedasticity in most statistical analyses of drinking problems. If error variances in estimates of drinking problems are systematically related to drinking measures (heteroscedasticity), then biased estimates of effects will be obtained from most statistical models (with the lone exception of ordinary least squares methods; Greene, 2012).

This study tests (a) whether self-reported distributions of heavy drinking and variances in drinking quantities among college drinkers reflect predictions from the drinking patterns table (i.e., comparing predicted and observed distributions) and (b) whether heteroscedasticity and bias in estimates of drinking risks follow the general pattern

predicted from the table, increasing in V and increasing then decreasing in F

Method

Data sources

Survey data on alcohol use and related problems come from the Safer California Universities evaluation of alcohol risk management strategies to reduce campus drinking problems (Saltz et al., 2010). Web-based survey data were collected from randomly sampled undergraduate students attending eight University of California and six California State University campuses in nine consecutive fall semesters (2003–2011). The final sample included 60,889 students, with 37,762 past-semester drinkers providing data on drinking behaviors, alcohol-related problems, and demographics. Response rates for the entire sample ranged from 50% in 2003 to 39% in 2008. Year-specific post hoc sample weights for each university were developed based on the gender and racial/ethnic composition of the target sample at each university relative to the gender and racial/ethnic composition of survey respondents from each university (Paschall & Saltz, 2007).

Drinking data were collected to estimate relative frequencies of use of different quantities of alcohol over 28 days (Gruenewald et al., 1996). Respondents indicated the number of times they consumed 1+, 2+, 3+, 6+, and 9+ drinks of alcohol in the previous 28 days or since the beginning of the semester (for low-frequency drinkers). A drink was defined as a 12-oz. glass or bottle of beer, 5-oz. glass of wine, or 1-oz. shot of distilled spirits. A log-logistic function was used to model these responses and obtain an estimate of the number of days on which $n = 1, 2, 3, \dots, 15$ drinks were consumed. Model fits were good, with an average R^2 value of .98. Drinking measures derived from this procedure exhibit good reliability (Gruenewald & Johnson, 2006). These procedures provided estimates of F , Q , V , $F(Q > k)$, and variance in drinking quantities for every drinker in the sample. Two measures of heavy drinking (defined as $k > 3$ or $k > 4$ for women and men, respectively) were computed: a dichotomous measure that identified whether one or more heavy drinking events occurred in the previous 28 days and frequencies of heavy drinking. A simplified version of the NIAAA guidelines for high-risk drinking identified those respondents who had consumed an average of 1 or 2 or more drinks per day over the previous 28 days and had more than 0.5 heavy drinking events during that time (again distinguishing between women and men).

Students reported the number of times they experienced problems related to drinking from the beginning of the current semester to time of the survey, a common question in college drinking studies (Wechsler et al., 1994). These reports were converted to rates per 28 days. Problems in-

cluded five physiological outcomes related to alcohol use (hangovers, forgetting, medical treatment for overdose, nausea/vomiting, and passing out) that frequently occur among college students (Gruenewald & Mair, 2015). To assess heteroscedasticity for dichotomous measures of rare events (Mair et al., 2016), self-reported dichotomous outcomes related to “unprotected sex” and “forced sex” were also included.

Statistical approaches

Predicted values for probabilities of heavy drinking and variances in drinking quantities were calculated for all combinations of V (1–56) and F (1–28) across the drinking patterns table. By definition, at least one drink must be consumed on each of F drinking days for all patterns. We assumed that the remaining $V - F$ drinks would be consumed on randomly selected days with probability $1 / F$ (i.e., any additional drinking independent of other drinks and days). A total of 20,000 simulations were executed for each combination of V and F . The proportion of iterations on which one or more heavy drinking events occurred provided an estimate of the probability of heavy drinking. Variances in drinking quantities across F drinking days about the mean $Q = V / F$ were averaged across iterations. Pearson correlations were then calculated between predicted and observed values for those cells of the drinking patterns table for which $n \geq 10$ cases were observed among college drinkers. Because heavy drinking is differently defined for women and men, these comparisons were made only among males (17,140 drinkers).

Censored regression models with and without multiplicative corrections for heteroscedastic errors (TOBITHEM procedure in Stata; Shehata, 2011) were used to assess the impacts of heteroscedasticity on risk estimates relating drinking measures to rates of self-reported physiological problems. Censored regression models are appropriate when dependent measures are left- or right-censored (i.e., all values that fall below or above the scale as measured are assigned the top or bottom value of that scale). In the current case, problem measures are treated as left-censored (no values less than zero) and, for this class of analysis methods, heteroscedasticity can lead to biased estimates of effects. Separate censored regression models were run for each of six different drinking measures: F , Q , V , H (dichotomous measure of heavy drinking), H_f (frequency of heavy drinking), and the NIAAA guidelines for high-risk drinking. As characterized in Table 1, heteroscedasticity was assumed to be a direct function of V and a quadratic function of F (increasing then decreasing). Differences in coefficients between models with and without corrections for heteroscedasticity tested the extent to which heteroscedasticity biased estimates of effects.

We also related three problem indices to drinking risks

TABLE 1. Bivariate regression coefficients for physiological problems without and with heteroscedasticity ($n = 37,762$)

Model	Coefficients		Multiplicative heteroscedasticity						$b - b(\text{Het})$	
	b	z	$b(F)$	$z(F)$	$b(F^2)$	$z(F^2)$	$b(V)$	$z(V)$	Bias	$z(\text{Bias})$
F	0.636	19.37								
$F(\text{Het})$	0.578	12.86	0.052	10.75	-0.001	-6.43	0.006	8.14	0.058 (9%)	1.005 [†]
Q	1.798	19.69								
$Q(\text{Het})$	1.202	16.99	0.076	15.66	-0.002	-9.32	0.005	15.86	0.596 (33%)	5.159
V	0.127	33.06								
$V(\text{Het})$	0.154	35.16	0.031	5.57	-0.001	-3.26	0.008	20.61	-0.027 (-21%)	-4.592
H	7.331	15.39								
$H(\text{Het})$	4.413	17.71	0.058	11.08	-0.002	-7.30	0.008	20.73	2.918 (40%)	5.431
Hf	1.033	32.15								
$Hf(\text{Het})$	1.111	31.37	0.072	15.55	-0.002	-9.01	0.005	10.23	-0.078 (-8%)	-1.632 [†]
NIAAA	9.054	20.59								
NIAAA(Het)	7.652	18.94	0.061	12.36	-0.002	-7.73	0.007	15.00	1.402 (15%)	2.347

Notes: Het = heteroscedasticity; F = number of drinking days (per 28 days); Q = average quantity consumed per drinking day; V = total number of drinks (per 28 days); H = dichotomous measure of heavy drinking; Hf = frequency of heavy drinking; NIAAA = National Institute on Alcohol Abuse and Alcoholism guidelines for high-risk drinking.

[†]not significant, $p > .05$.

assessed using a dose–response model applied in previous work. Measured rates of physiological problems were regressed over F , the difference $V - F$, and the interaction of $V - F$ with $1 / Q$, representing attenuation in dose–response among heavier drinkers (see Gruenewald & Mair, 2015). Binary indices of unprotected and forced sex were regressed over the same measures using heteroscedastic logistic regressions (OGLM procedure in Stata; Williams, 2009, 2010). To further estimate the contributions of one stochastic aspect of drinking patterns within cells, we included frequencies of heavy drinking in the analyses of physiological problems and forced sex. Collinearity among drinking measures with respect to the prevalence of unprotected sex precluded including this measure in the final analysis model. Population weights helped ensure generalizability of results to all students attending University of California and State campuses. All models corrected for loss of unit independence related to nesting of students within schools by including fixed campus effects.

Results

Figure 4 shows relationships between predicted and observed variances in drinking quantities for $F = 1$ –28 and $V = 10, 15, 20, 25$, and 56 drinks. Across the 64 cells for which comparisons were possible, the correlation between predicted and observed variances was significant ($r = .621, p < .001$). However, variance in drinking quantities was far greater than expected in the mid-range of drinking frequencies. This overdispersion was captured in a significant quadratic effect relating variance to F (ordinary least squares regression, $V = a + bF + cF^2$; $b = 4.418, t = 9.044, p < .001$; $c = -0.556,$

$t = -6.724, p < .001$), yielding a much-improved correlation between predicted and observed values ($r = .8031, p < .001$). The correlation between predicted and observed probabilities for heavy drinking was also significant ($n = 38, r = .934, p < .001$). Infrequent drinking was associated with greater-than-expected probabilities of heavy drinking. Observed and predicted values were similar at higher drinking frequencies and volumes.

Table 1 presents the results of separate regression models relating physiological problems to six different drinking measures. Each coefficient was estimated with and without corrections for heteroscedasticity. All coefficient estimates were significantly different from zero. All coefficient estimates related to heteroscedasticity followed the predicted pattern. Bias between models, assessed by differencing coefficients estimated with and without corrections for heteroscedasticity, was significant for all but the measures of drinking and heavy drinking frequencies. Uncorrected coefficient estimates were biased between -8% and 40%.

Table 2 presents results from the censored and logistic regression analyses of rates of physiological problems and incidents of unprotected and forced sex. Controls for heteroscedasticity were again significant in all cases and followed the predicted pattern in all but one case. It was expected that heteroscedasticity would increase with respect to $V - F$ but instead significantly declined for forced sex. Effects estimates for coefficients of the dose–response model of physiological problems were similar to those obtained in previous studies; greater risks were associated with $V - F$ moderated by $1 / Q$ (reciprocal of average drinking levels) (Gruenewald & Mair, 2015). Risks for unprotected sex increased with F but were unrelated to $V - F$ or the

FIGURE 4. Predicted versus observed variance in drinking quantities for selected drinking volumes, V , across 28 day drinking frequencies; males only. White circles = predicted values, black circles = observed values with $\pm 95\%$ SE bars.

TABLE 2. Results of heterogeneous dose–response censored and logistic regressions including corrections for heteroscedasticity ($n = 37,762$)

Variable	Physiological problems			Unprotected sex			Forced sex		
	<i>b</i>	<i>z</i>	<i>p</i>	<i>b</i>	<i>z</i>	<i>p</i>	<i>b</i>	<i>z</i>	<i>p</i>
Independent measures									
Constant	-0.556	-17.70	<.001	4.154	31.78	<.001	4.144	23.95	<.001
<i>F</i>	-0.052	-8.74	<.001	0.092	4.82	<.001	-0.473	-4.27	<.001
<i>V</i> – <i>F</i>	0.003	1.19	N.S.	-0.019	-1.89	N.S.	0.016	2.54	.011
(<i>V</i> – <i>F</i>)(1 / <i>Q</i>)	0.335	25.02	<.001	0.052	1.80	N.S.	0.353	3.26	.001
Heavy drinking	0.137	12.64	<.001	—	—	— ^b	0.077	2.07	.039
Heteroscedasticity									
<i>F</i>	0.043	15.70	<.001	0.038	5.55	<.001	0.098	8.39	<.001
<i>F</i> ²	-0.001	-7.52	<.001	-0.003	-8.23	<.001	-0.002	-7.02	<.001
<i>V</i> – <i>F</i>	0.008	31.13	<.001	0.016	6.63	<.001	-0.005	-2.55	.011
Het $\Delta\chi^2$ ($\Delta df = 3$) ^a	—	6,643.22	<.001	—	374.08	<.001	—	94.31	<.001
σ	1.144	87.08	<.001	—	—	—	—	—	—
Model χ^2	—	28,804.31	<.001	—	2,309.15	<.001	—	544.44	<.001

Notes: Het = heteroscedasticity; — = effect not estimated or reported; N.S. = not significant. ^aStatistical contribution of heteroscedasticity terms to each model;

^bheavy drinking effects could not be estimated (see text).

moderating effects of $1 / Q$. Risks for forced sex decreased as drinking was spread across more drinking occasions but increased with $V - F$ moderated by $1 / Q$. Frequencies of heavy drinking were associated with greater physiological problems and forced sex. An examination of the impacts of heteroscedasticity on the 11 coefficient estimates showed that, without these controls, the moderating effect of $1 / Q$ was underestimated for physiological problems ($\Delta z = -2.60$, $p = .003$), and all estimates were biased for the measure of unprotected sex ($\Delta z = -4.72$, $p < .001$ for F ; $\Delta z = 2.51$, $p = .006$ for $V - F$; $\Delta z = 2.01$, $p = .022$ for $1 / Q$) and forced sex ($\Delta z = 4.18$, $p < .001$ for F ; $\Delta z = -2.81$, $p = .003$ for $V - F$; $\Delta z = -2.47$, $p = .007$ for $1 / Q$).

Discussion

The theoretical and empirical analyses of drinking patterns provided here show that different measures of alcohol use can be assembled into a drinking patterns table and directly related one to the other and to distributions of heavy drinking and variances in drinking quantities. It appears that the number and variety of drinking patterns that arise for combinations of V and F is a crucial source of heteroscedasticity and potential bias in estimates of risks related to alcohol use. We show how probabilities of heavy drinking and variances in drinking quantities are predictable from the table and demonstrate that the overall predictive performance was generally good but with some overdispersion of variances in drinking quantities at moderate frequencies for any volume (Figure 4). The extremes of these distributions represent episodic heavy drinking (high-volume, low-frequency—to the left of the drinking patterns table) and regular daily drinking (volume is a small multiple of frequency—to the right of the table). In between, many different drinking patterns arise (see Figure 2). Greater observed than predicted probabilities of heavy drinking were also observed at lower drinking frequencies, suggesting that drinkers tend to cluster

drinking on specific occasions such as weekend nights and holidays (an observation supported by time series data; e.g., Gruenewald et al., 2005). Observed probabilities of heavy drinking were less than those expected by chance alone at the greatest volumes ($V = 56$) and indicate that the highest volumes are achieved among frequent regular users.

Implications for the measurement of drinking behaviors

There is little evidence of any coordinated application of drinking measures across studies in the empirical literature (Figure 1). As shown here, measures of F and V , used together, capture some of the breadth of alcohol effects. But only 8 of 67 studies (12%) used both to predict problems, suggesting limited awareness of the importance of distinguishing impacts of these two measures. The drinking patterns table also reminds us that Q is nonlinearly related to F and V , as are many other drinking measures. For this reason we suspect that relationships between drinking patterns and problems are nonlinear (Table 2), and linear models of these relationships are inappropriate. Finally, it should be obvious that the seven studies that used categorical measures cannot inform the nature of these quantitative risk relationships because they conflate rather than distinguish measures of drinking patterns and dose–response relationships. These categorical assessments may be of great value for the communication of effective and useful public health messages. However, the science on which such recommendations are based should be sought in quantitative analyses of dose–response relationships.

It appears that researchers must critically assess the measures they use for statistical analyses of drinking risks. Otherwise, interpretations of “drinking effects” will remain incomparable and uninformative to alcohol epidemiology. The attribution of an alcohol-related effect must be carefully made with all alternative drinking measures in mind or with some formal model of dose–response relationships in

hand. In addition, consideration of the determinate aspects of drinking patterns alone is not enough; stochastic aspects of drinking events have important and predictable implications for assessments of risks related to use. This observation has long been recognized in studies that include measures of heavy drinking, but the quantitative relationships between these and other measures of alcohol use are poorly understood.

There are two final questions that emerge from this study: What are appropriate measures of use relevant to any alcohol-related problem, and how do researchers identify these measures? The first question is answered through a consideration of drinking patterns as defined here. For standard survey-based drinking measures, all such patterns can be identified and measures developed that reflect aspects of these patterns relevant to any drinking problem (heavy drinking measures are one example). The second question bears upon the development and testing of quantitative models relating measures of drinking patterns to problems, the formulation of explicit dose-response models of alcohol-related harms, and explorations of the interactions of these dose-response effects with ecological conditions of use.

References

- Berger, L., Fendrich, M., & Fuhrmann, D. (2013). Alcohol mixed with energy drinks: Are there associated negative consequences beyond hazardous drinking in college students? *Addictive Behaviors*, 38, 2428–2432. doi:10.1016/j.addbeh.2013.04.003
- Greene, W. H. (2012). *Econometric analysis* (7th ed.). New York, NY: Prentice Hall.
- Gruenewald, P. J., & Johnson, F. W. (2006). The stability and reliability of self-reported drinking measures. *Journal of Studies on Alcohol*, 67, 738–745. doi:10.15288/jsa.2006.67.738
- Gruenewald, P. J., & Mair, C. (2015). Heterogeneous dose-response and college student drinking: Examining problem risks related to low drinking levels. *Addiction*, 110, 945–954. doi:10.1111/add.12887
- Gruenewald, P. J., Searles, J., Helzer, J., & Badger, G. J. (2005). Exploring drinking dynamics using interactive voice response technology. *Journal of Studies on Alcohol*, 66, 571–576. doi:10.15288/jsa.2005.66.571
- Gruenewald, P. J., Treno, A. J., & Mitchell, P. R. (1996). Drinking patterns and drinking behaviors: Theoretical models of risky acts. *Contemporary Drug Problems*, 23, 407–440.
- Mair, C., Ponicki, W. R., & Gruenewald, P. J. (2016). Reducing risky sex among college students: Prospects for context-specific interventions. *AIDS and Behavior*, 20, Supplement 1, 109–118. doi:10.1007/s10461-015-1147-2
- O'Malley, P. M. (2014). A review of studies of drinking patterns in the United States since 1940. *Journal of Studies on Alcohol and Drugs, Supplement 17*, 18–25. doi:10.15288/jsads.2014.s17.18
- Paschall, M. J., & Saltz, R. F. (2007). Relationships between college settings and student alcohol use before, during and after events: A multi-level study. *Drug and Alcohol Review*, 26, 635–644. doi:10.1080/09595230701613601
- Saltz, R. F., Paschall, M. J., McGaffigan, R. P., & Nygaard, P. M. (2010). Alcohol risk management in college settings: The safer California universities randomized trial. *American Journal of Preventive Medicine*, 39, 491–499. doi:10.1016/j.amepre.2010.08.020
- Shehata, E. A. E. (2011). *TOBITHEM: Stata module to estimate Tobit Multiplicative Heteroscedasticity Regression*. Retrieved from <http://ideas.repec.org/c/boc/bocode/s457323.html>
- Snipes, D. J., & Benotsch, E. G. (2013). High-risk cocktails and high-risk sex: Examining the relation between alcohol mixed with energy drink consumption, sexual behavior, and drug use in college students. *Addictive Behaviors*, 38, 1418–1423. doi:10.1016/j.addbeh.2012.07.011
- Wechsler, H., Davenport, A., Dowdall, G., Moeykens, B., & Castillo, S. (1994). Health and behavioral consequences of binge drinking in college. A national survey of students at 140 campuses. *JAMA*, 272, 1672–1677. doi:10.1001/jama.1994.03520210056032
- Williams, R. (2009). Using heterogeneous choice models to compare logit and probit coefficients across groups. *Sociological Methods & Research*, 37, 531–559. doi:10.1177/0049124109335735
- Williams, R. (2010). Fitting heterogeneous choice models with oglm. *Stata Journal*, 10, 540–567. Retrieved from <http://www.stata-journal.com/article.html?article=st0208>