LETTERS TO THE EDITOR

RE: "RELIABILITY OF ALCOHOL INTAKE AS RECALLED FROM 10 YEARS IN THE PAST"

The recent article of Liu et al. (1) describes the reliability of alcohol intake as recalled from 10 years in the past in a cohort of 2,907 US adults. Individuals were categorized into "very light," "light," "moderate," or "heavy" drinkers according to their self-reported baseline and recalled alcohol intake. The authors estimate the accuracy of recalled alcohol intake using the intake at baseline as the "gold standard." Yet, in the Discussion, they briefly note that reported intake at baseline may not be free of reporting error.

Indeed, an error-free gold standard for self-reported alcohol intake is badly lacking. Instead of relying on the baseline reports, a latent class approach could be used (2-4) to explore reporting errors of both recalled intake and intake reported at baseline. By this approach, individuals are assumed to belong to an unknown true category of very light, light, moderate, or heavy drinkers. By estimating the misclassification rates (error probabilities) within these true latent drinking categories, the accuracy of recalled intake and intake at baseline can be assessed.

We have taken the liberty to use one of the authors' tables (table 2, "Correspondence between alcohol intake reported at baseline ... and at follow-up" in men (totals)) to explore the data by two latent class models, herein presented as tables 1 and 2. Because we have only two replicates per individual, constraints are required for the estimation of the error probabilities (2). The analyses were performed using the CDAS© software (5).

Model I

In model I (table 1), we assume that individuals always classify themselves into the correct or an adjacent category. This model leaves unexplained some gross discrepancies in reported intake (very light at baseline and heavy at 10 years past and vice versa, representing 1.9 and 0.6 percent of the individuals, respectively). Nevertheless, 15 probabilities (prevalences of the true latent categories and error rates) should be estimated in this model. We obtain a $\chi^2$ of 0.709, but we do not have any degree of freedom left to formally test the fit to the data.

Model II

In model II (table 2), we assume that the individuals always understate their drinking habits. The amount of underreporting may be different at baseline and at follow-up. Moreover within latent classes, the probabilities of understating more than one category are considered to be equal at baseline and after 10 years, respectively. This model exhibits an acceptable fit to the data ($\chi^2 = 8.61$, 4 df).

The models above are two examples from a vast area of possible latent class approaches to the data. As a matter of fact, the analysis in the original publication (using the gold standard) could also be considered a latent class model, albeit with even more rigid restrictions than ours. Neither of our two models should be considered "appropriate" without additional discussion of the potential error mechanism and the assumptions underlying the latent class approach. Although both models indicate that moderate and heavy drinkers understate their alcohol consumption, we get conflicting results for the error rates of light drinkers.

As mentioned in the original publication (1), factors such as age and current drinking habits may influence accuracy. It would certainly be worthwhile to include these factors in a more thorough latent class analysis of true alcohol intake. In our opinion, the description of errors in the reporting of alcohol intake could be enhanced by using a latent class approach, which does not necessarily assume that one of the reports is error free.

---

**TABLE 1.** Estimates of latent probabilities and conditional probabilities (model I)

<table>
<thead>
<tr>
<th>Baseline</th>
<th>Very light</th>
<th>Light</th>
<th>Moderate</th>
<th>Heavy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latent class</td>
<td>0.249</td>
<td>0.249</td>
<td>0.179</td>
<td>0.323</td>
</tr>
<tr>
<td>Very light</td>
<td>0.906</td>
<td>0.185</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Light</td>
<td>0.095</td>
<td>0.056</td>
<td>0.449</td>
<td>0</td>
</tr>
<tr>
<td>Moderate</td>
<td>0</td>
<td>0.759</td>
<td>0.486</td>
<td>0.482</td>
</tr>
<tr>
<td>Heavy</td>
<td>0</td>
<td>0</td>
<td>0.065</td>
<td>0.518</td>
</tr>
</tbody>
</table>

**TABLE 2.** Estimates of latent probabilities and conditional probabilities (model II)

<table>
<thead>
<tr>
<th>Baseline</th>
<th>Very light</th>
<th>Light</th>
<th>Moderate</th>
<th>Heavy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latent class</td>
<td>0.168</td>
<td>0.104</td>
<td>0.314</td>
<td>0.415</td>
</tr>
<tr>
<td>Very light</td>
<td>1</td>
<td>0.500</td>
<td>0.102</td>
<td>0.048</td>
</tr>
<tr>
<td>Light</td>
<td>0</td>
<td>0.500</td>
<td>0.144</td>
<td>0.048</td>
</tr>
<tr>
<td>Moderate</td>
<td>0</td>
<td>0</td>
<td>0.754</td>
<td>0.461</td>
</tr>
<tr>
<td>Heavy</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.443</td>
</tr>
</tbody>
</table>

* Probabilities assumed to be equal in the model.
† Probabilities assumed to be equal in the model.
In their recent article on the "reliability of alcohol intake as recalled from 10 years in the past," Liu et al. (1) compared alcohol consumption reported in the First National Health and Nutrition Examination Survey (NHANES I, 1971–1975) with that reported in the NHANES I Epidemiologic Follow-up Study conducted 10 years later. Their description of questions on alcohol intake asked in NHANES I omitted the instruction (2) that after question 18b on frequency of drinking, respondents were asked, "Which do you most frequently drink—beer, wine, or liquor?" (18c). The quantity question (18d) followed and asked, "When you drink (beer/wine/liquor), how much do you usually drink over 24 hours? [Enter an amount only for the one marked in 18c.]" Accordingly, their estimate of baseline alcohol consumption is based on a quantity-frequency (QF) measure in which the quantity represents the number of drinks respondents usually had of their favorite beverage (favored beverage QF), rather than the global QF measure used in the recall of drinking 10 years previously, which asked about the usual amount drunk when any alcoholic beverage was drunk.

The influence of this difference in the quantity questions on estimations of alcohol consumption was investigated in 1991 (3), using data from the New York State (NYS)-wide survey of alcohol and other drug use in which both beverage-specific and global QF questions had been asked. The NYS survey was conducted by telephone in a sample taken using random digit dialing techniques supplemented by samples of homeless people, college students, and individuals not having telephones (n = 6,364). Favored beverage estimations of mean daily alcohol consumption were based on the following: (drinking days per year X drinks per drinking day of beverage drunk)/365 days = mean ounces of ethanol per day. Global estimations were analogous. Quantity (ounces) of ethanol/drink was based on drinking patterns in NYS, for the global QF had been asked. The NYS survey was conducted by telephone in a sample taken using random digit dialing techniques supplemented by samples of homeless people, college students, and individuals not having telephones (n = 6,364).

Liu et al. (1) pointed out that a major difference between their study and theirs was that the former included abstinence and abstainers were included in their sample. However, we pointed out in our paper (1) that an examination of the data showed that only 8.72 percent of the respondents are classified into the same categories by both the favored beverage and global estimates, 7.1 percent are classified into a higher category by the favored beverage estimate, and 5.7 percent are classified into a higher category by the global estimate. Thus, despite the somewhat greater mean estimate of consumption derived from the favored beverage, there is little evidence to suggest that it would have greatly affected categorization of baseline drinking in NHANES I.

Liu et al. found that heavy drinkers recalled having an average of 0.57 drinks per day less than they originally reported 10 years earlier. However, the present analysis suggests that using the favored beverage QF may overestimate original intake among heavy drinkers by as much as 1.02 drinks per day. If this potential overestimation is taken into consideration, it appears that heavy drinkers tended to recall drinking more per day than originally reported, rather than less.

**REFERENCES**


Irene Guggenmoos-Holzmann
Richard Vonk
Institut für Medizinische Statistik und Informationsverarbeitung
Universitätsklinikum Benjamin Franklin, Freie Universität Berlin
Hindenburgdamm 30
12200 Berlin, Germany

**THE AUTHORS’ REPLY**

We appreciate Guggenmoos-Holzmann and Vonk’s suggestion regarding the use of latent class models to assess misclassification between “true” alcohol intake and recalled alcohol intake. It would be ideal if we had a “gold standard” measure of alcohol intake with which recalled intake could be compared. However, as we pointed out in our paper (1), no such measurement of alcohol intake was available, and we thus had to rely on self-reported alcohol intake at baseline as an indicator.

If we had had multiple records of recalled alcohol intake, some latent class modeling techniques could have helped to assess the validity of recalled intake (2). Because only one measure of recalled alcohol intake per individual was available in our data, we decided to assess the reliability of recalled intake. Using our data, Guggenmoos-Holzmann and Vonk pro-