Reliability of Alcohol Recall After 15 Years and 23 Years of Follow-Up in the Johns Hopkins Precursors Study*

AUDREY Y. CHU, M.H.S.,[†] LUCY A. MEONI, SC.M.,[†] NAE YUH WANG, PH.D.,[†] KUNG-YEE LIANG, PH.D.,[†] DANIEL E. FORD, M.D., M.P.H.,[†] AND MICHAEL J. KLAG, M.D., M.P.H.[†]

Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, 615 N. Wolfe Street, W6507, Baltimore, Maryland 21205

ABSTRACT. Objective: Recall of past alcohol intake is used in many studies of chronic disease, but few studies have been able to examine its long-term reliability. **Method:** We sought to assess the reliability of recalled alcohol intake assessed at an average age of 70 years in 2001, after 15 and 23 years of follow-up, in a prospective study of medical students in classes 1948 to 1964. **Results:** Average reported alcohol intake 15 years and 23 years prior were 6.3 and 7.4 drinks per week, respectively. Recall of alcohol intake overestimated the concurrently reported intake after 15 years by a mean of 0.47 (95% CI [0.10, 0.85]) drinks per week and underestimated intake after 23 years by a mean of 0.79 (95% CI [-1.27, -0.30]) drinks per week, mostly driven by differences between

concurrently reported and recalled distilled spirits consumption. Characteristics associated with underestimation of alcohol recall were age of 71 years or older in 2001, self-report of memory difficulties, and self-report of difficulties in physical functioning. In multivariate regression analyses combining 15- and 23-year recall, subjects who reported consumption of more than 14 alcoholic drinks per week in 2001 marginally overestimated recall by slightly more than 1 drink per week ($M = 1.18$ drinks/week, 95% CI [-0.03, 2.40]). **Conclusions:** Although significant differences were detected, recalled alcohol intake after 15 and 23 years of follow-up is remarkably reliable. (*J. Stud. Alcohol Drugs*, 71, 143-149, 2010)

VALID AND RELIABLE ASSESSMENT of alcohol intake is important in many studies of chronic disease and health behaviors. Studies that characterize past or lifetime alcohol use assume that participants can reliably recall alcohol intake before the beginning of the study, sometimes decades before (Klipstein-Grobusch et al., 1999). The validity of self-reported alcohol intake (Del Boca and Darkes, 2003; Embree and Whitehead, 1993; Grant et al., 1997; Poikolainen, 1985) and the reliability of short-term recall (Ekholm, 2004; Grant et al., 1997; Longnecker et al., 1992; Sobell and Sobell, 1978) are well described, but long-term recall of alcohol intake has not been studied as extensively (Czarnecki et al., 1990; Dwyer et al., 1989; Lee et al., 1992; Liu et al., 1996). In general, studies have found that alcohol intake can be reliably assessed through recall. Only four of these studies have examined recall reliability over 10 years or longer, and only one (Liu et al., 1996) has examined respondent characteristics that may be related to differences

in recall. None of the studies has used longitudinal data analysis techniques to examine the difference in recall over two recall periods.

The Johns Hopkins Precursors Study has followed a cohort of former medical students since 1948 (Thomas, 1951), with periodic mailed questionnaires requesting information on health status and risk factors for cardiovascular disease, including alcohol intake. The purpose of the present analysis was to assess the reliability of recall of total alcohol intake across 15 and 23 years of follow-up and to determine factors associated with disagreement between recalled and concurrently reported intake. We hypothesized that poorer self-assessed memory, older age, and incident illness between concurrent assessment of alcohol drinking and recall 15 and 23 years later were associated with greater differences between concurrently reported and recalled alcohol intake. We also hypothesized that the difference in recall did not differ by level of consumption at the time of recall.

Received: January 14, 2009. Revision: June 25, 2009.

*This research was supported by National Institutes of Health grants NIH RO1 AG01760, NIH T32-HL007024, and NIH-NIDDK K24 DK02856.

[†]Correspondence may be sent to Audrey Y. Chu at the above address or via email at: achu@jhsph.edu. She is also with the Welch Center for Prevention, Epidemiology and Clinical Research, Johns Hopkins University, Baltimore, MD. Lucy A. Meoni is with the Department of Biostatistics, Bloomberg School of Public Health; the Department of Medicine, Division of General Internal Medicine, School of Medicine; and the Welch Center for Prevention, Epidemiology and Clinical Research, Johns Hopkins University, Baltimore,

MD. Nae Yuh Wang is with the Department of Medicine, Division of General Internal Medicine, School of Medicine; and the Welch Center for Prevention, Epidemiology and Clinical Research, Johns Hopkins University, Baltimore, MD. Kung-Yee Liang is with the Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD. Daniel E. Ford and Michael J. Klag are also with the Department of Health Policy and Management, Bloomberg School of Public Health; the Department of Medicine, Division of General Internal Medicine, School of Medicine; and the Welch Center for Prevention, Epidemiology and Clinical Research, Johns Hopkins University, Baltimore, MD.

Method

Study population

The Precursors Study was designed and implemented by Dr. Caroline B. Thomas (Thomas, 1951) in 1947 and enrolled medical students at the Johns Hopkins School of Medicine in the graduating classes of 1948 through 1964. Baseline medical examinations and questionnaires concerning personal and family medical history and health habits were completed during medical school. Questionnaires have been mailed annually since graduation to assess current health status as well as presence of risk factors for cardiovascular disease and other morbidities.

In 2001, 646 (73%) persons of 886 who completed questionnaires in 1978 or 1986 and who were not lost to follow-up by the end of 2001 answered questions about recall of alcohol intake consumed in 1978 and 1986. Of these, 501 (78%) and 611 (95%) concurrently reported their alcohol drinking behavior in 1978 and 1986, respectively. Seventy-two met the criteria of a positive history of alcohol misuse defined as the presence of any of four criteria: (a) self-report of alcoholism or alcohol misuse on annual questionnaires since the beginning of the study, (b) a diagnosis of alcohol abuse in medical records since the beginning of the study, (c) self-report of 28 drinks per week or more in 2001, or (d) two or more positive responses to the CAGE questionnaire in 2000. The CAGE survey consists of four questions that assess presence of alcohol abuse and dependence (Ewing, 1984). Because of the relatively small number of participants meeting the criteria for alcohol misuse and the resultant limited power, we excluded subjects with a history of alcohol misuse from our analyses. Our final analytical sample includes those subjects who completed the recall questionnaire in 2001, at least one of the questionnaires in 1978 or 1986, and were not identified as alcohol misusers ($n = 574$). The study protocol was approved by our institutional research ethics committee. All subjects gave written informed consent for participation in the Precursors Study.

Alcohol intake

Questionnaires sent in 1978, 1986, 1989, 1993, 1997, and 2001 inquired about frequency of concurrent total and beverage-specific alcohol consumption; the 2001 questionnaire included additional questions regarding recall of alcohol intake in 1978 and 1986. Questions were asked in the form of "How many servings of a specific type of alcohol (beer, wine, or hard liquor) were consumed in a typical week?" for all years. Total alcohol consumption was determined by the sum of the beverage-specific responses.

Covariates

Factors thought to affect recall of alcohol intake were included as covariates in the analysis. We chose variables

obtained by questionnaire between 2000 and 2001 and from medical records, attempting to mimic the type of data obtained in cross-sectional and case-control studies.

Body mass index (BMI) was calculated from concurrently reported weight in 2001 (a measure shown to be valid; Klag et al., 1993b) and measured height from the baseline physical examination at cohort entry. Physical functioning scores from the Short Form-36 Health Survey (SF-36) questionnaire administered in 2001 were standardized with a range from 0 to 100, with 100 representing ability to perform all types of physical activity without limitation due to health (Ware and Sherbourne, 1992).

Memory functioning was assessed through the Memory Functioning Questionnaire in 2000. The Memory Functioning Questionnaire consists of 19 self-reported questions that assess daily memory functioning (Gilewski and Zelinski, 1988). Responses are on a 7-point Likert scale, with higher scores indicating higher opinion of one's memory. Two domains of memory functioning are evaluated: general frequency of forgetting (15 questions) and mnemonics usage (4 questions). The general frequency of forgetting domain measures self-opinion of how often forgetting occurs and self-ratings of general memory performance (Bunch et al., 2004; Zelinski et al., 1990); mnemonics usage is a measure of adaptation techniques to losses in memory (Bunch et al., 2004; Zelinski et al., 1990). Both measures of memory were categorized into quartiles.

Incidence of diseases—cardiovascular disease, diabetes, and hypertension—postulated to influence level of alcohol intake was assessed between 1978 and recall of intake in 2001 by annual questionnaires. Diagnoses were assigned by a committee of physicians after thorough review of all available information, including medical records (Klag et al., 1993a).

We considered other covariates thought to be confounders in recall of alcohol intake, such as smoking and socioeconomic status. However, cigarette smoking in 2000 was not associated with recalled alcohol intake, and was not included in further analyses. Additionally, adult levels of socioeconomic status were similar for the subjects included in this analysis as they are all medical doctors. Thus, the relationship of socioeconomic status to recalled alcohol intake could not be assessed.

Ethics

Procedures were followed in accordance with ethical standards of The Johns Hopkins School of Medicine Office of Human Subjects Research and Institutional Review Board.

Statistical analysis

To assess the reliability of long-term recall of total and beverage-specific alcohol intake, we compared recall as-

essed in 2001 with alcohol intake determined in 1986 (15 years previously) and in 1978 (23 years previously). Recall difference was defined as the previous alcohol intake (number of drinks/week) recalled in 2001 minus the alcohol consumption reported concurrently, for both 1986 and 1978. Our main exposure variable was total alcohol intake reported in 2001 (drinks/week). Spearman rank correlations between concurrently reported alcohol intake and recalled past intake were determined for total alcohol and beverage-specific intake (beer, wine, distilled spirits). Recall differences were graphed as a function of the average of recalled and concurrently reported intake (Bland and Altman, 1986, 1995, 1999). This method (Bland-Altman plots) reveals small differences and effects of outliers that may be masked in scatter plots and avoids autocorrelation. Bias may be induced by unequal variances of the variables in Bland-Altman analysis; therefore, variances of recalled intake and concurrent intake in this analysis were equalized before examination of the plots. We applied generalized estimating equations (GEE) developed by Liang and Zeger (1986) to examine factors associated with reliability of recall of alcohol intake over time. Within-subject correlation of alcohol recall over the recall period can be accounted for by GEE, which allows valid inferences from longitudinal data.

To test the hypothesis that recall of alcohol consumption did not differ by level of alcohol drinking in 2001, the difference between recalled and concurrent alcohol intake was regressed on the level of alcohol consumption in 2001 using GEE clustered on individual subject response for both 1978 and 1986. For these analyses, alcohol consumption in

2001 was categorized into four groups: 0 drinks per week (nondrinkers), 1-7 drinks per week, 8-14 drinks per week, and more than 14 drinks per week. The same approach was used to test whether other variables, based on a priori hypotheses, assessed in 2000 or 2001 were associated with recall difference. These variables included age, BMI, SF-36 physical functioning, memory function, and incident illness between 1978 and 2001 (cardiovascular disease, diabetes, or hypertension). The sum of the general frequency of forgetting questions and the sum of mnemonics usage were entered into the model separately as they represent different aspects of memory (general forgetting and adaptation techniques, respectively). Indicators of incident disease were cumulatively summed and entered as an ordinal variable.

Variables were introduced into bivariate analysis with level of 2001 alcohol intake (the main exposure); if the covariate was not significant in its association with difference in alcohol recall or if it did not change the estimated 2001 alcohol intake coefficient by more than 10%, it was not considered to be a confounder and was dropped from analysis. Covariates kept in the analysis include level of 2001 alcohol intake (entered as indicator variables for the four categories), age (continuous), sex (indicator for female), general frequency of forgetting (quartiles), and SF-36 physical functioning (quartiles). All analyses were carried out using the Stata (Release 9.0; StataCorp LP, College Station, TX). Significance is indicated at the .05 two-sided α level.

Results

Table 1 describes characteristics of the 574 participants included in the analysis who reported alcohol intake in 1978, 1986, and 2001. The majority of the study population was White and male, with an average age of 70 years in 2001 and a mean BMI in the overweight range. The average number of alcoholic drinks concurrently reported in each of the three questionnaires was about seven drinks per week. Less than half the study population developed cardiovascular disease, diabetes, or hypertension between 1978 and 2001.

TABLE 1. Characteristics of participants who responded to alcohol intake questionnaires ($n = 574$)

Variable	<i>M</i>	<i>SD</i>
Age in 2001, in years	70.1	5.2
Body mass index in 2001, kg/m ²	25.0	3.6
SF-36 physical functioning in 2001	84.3	20.6
Concurrently reported alcohol intake, drinks/week		
2001 ($n = 574$)	6.3	6.3
15 years before ($n = 543$)	6.3	6.4
23 years before ($n = 445$)	7.4	7.0
Recall of past intake, drinks/week		
15 years before	6.7	6.5
23 years before	6.3	6.2
	<i>n</i>	%
Male	524	91.3
White	555	97.0
Incident disease, 1978-2001		
Cardiovascular disease	118	20.6
Diabetes	36	6.3
Hypertension	183	31.9
Category of alcohol intake in 2001, drinks/week		
0	143	24.9
1-7	236	41.1
8-14	119	20.7
>14	76	13.2

Notes: SF-36 = Short Form-36 Health Survey

TABLE 2. Spearman rank correlations and differences between recalled and concurrently reported alcohol intake over 15 and 23 years of follow-up, by type of alcohol

Variable	Recall, mean drinks/week	Concurrent, mean drinks/week	<i>r</i> ²	Difference, mean drinks/week	<i>p</i>
15 years ($n = 543$)					
Total	6.73	6.26	.81	0.47	.01
Beer	1.32	1.20	.67	0.11	.14
Wine	2.86	2.69	.69	0.17	.19
Distilled spirits	2.56	2.40	.74	0.19	.16
23 years ($n = 445$)					
Total	6.58	7.36	.73	-0.79	<.01
Beer	1.47	1.53	.60	-0.07	.57
Wine	2.48	2.47	.57	0.01	.93
Distilled spirits	2.63	3.36	.68	-0.73	<.01

FIGURE 1. Difference in recalled and concurrent total alcohol intake after 15 years of follow-up, by average of recalled and concurrent alcohol intake, in 541 participants. Dashed lines: Difference between recalled and concurrent total alcohol intake over 15 years of follow-up (drinks/week) ($SD = 1.96$). Solid line: Linear regression of difference in recall and concurrent intake on average intake.

FIGURE 2. Difference in recalled and concurrent total alcohol intake after 23 years of follow-up, by average of recalled and concurrent alcohol intake, in 443 participants. Dashed lines: Difference between recalled and concurrent total alcohol intake over 23 years of follow-up (drinks/week) ($SD = 1.96$). Solid line: Linear regression of difference in recall and concurrent intake on average intake.

Table 2 presents the concurrently reported (in 1978 and 1986) and recalled (in 2001) mean alcohol intake, correlation coefficients, and mean recall difference by type of alcohol consumed (total, beer, wine, or distilled spirits). Recall overestimated total intake 15 years previously (0.5 drinks/week) but underestimated total intake 23 years previously (-0.8 drinks/week). Although these differences were statistically significant, the absolute magnitude was small and correlations between recalled and concurrently reported values were high. Concurrently reported and recalled intake for distilled spirits at the 23-year follow-up followed the same pattern as for total intake, but that for beer and wine did not differ for either the 15- or 23-year follow-up.

In Bland-Altman analysis, the difference between recalled and concurrently reported intake did not differ by the average of recalled and concurrently reported alcohol intake for either 1986 ($b = 8.4 \times 10^{-8}$, $p = 1.00$) or 1978 ($b = -5.9 \times 10^{-8}$, $p = 1.00$) intake. As expected, given that a wider range of differences is possible at higher levels of alcohol intake, variability in the difference between recall of total alcohol intake and concurrent intake appeared to be greater at higher levels of drinking for both 15- and 23-year follow-up data (Figure 1: 15 years, 1986; Figure 2: 23 years, 1978).

The association seen in the analysis of total alcohol intake 15 years ago was mostly driven by differences between concurrently reported and recalled beer consumption. Differences between recalled and concurrently estimated beer intake 15 years previously were greater at higher levels of beer intake ($b = 0.17$ drink/weekly intake difference, $p < .001$, for each one beer increase in concurrently estimated intake). However, the association of total alcohol intake 23 years ago was driven by differences in concurrently reported and recalled distilled spirits consumption; recall of intake 23 years ago underestimated ($b = -0.43$ drink/weekly intake difference, $p < .001$) previous intake. Recall of wine did not differ by level of average alcohol intake for either 15 or 23 years of recall.

We examined the mean difference in recalled alcohol intake by sex and found there was no difference in recalled alcohol intake 15 years ago or 23 years ago. There was also no difference by sex in linear regression analysis of recalled alcohol intake by alcohol consumption in 2001 for either of the two recall periods.

In univariate GEE analysis of alcohol consumption in 2001 and recall, those in the highest category of drinking overestimated concurrently reported alcohol intake by 1.21 drinks per week ($p = .01$) and those who did not consume alcohol in 2001 underestimated concurrently reported alcohol intake by 0.96 drinks per week ($p = .007$). Recall did not differ, however, when examined over the other two alcohol intake categories (1-7 and 8-14 drinks/week, $p > .05$). We found a significant linear trend of greater overestimation of recall by increasing the alcohol intake category ($p = .002$).

TABLE 3. Estimated mean differences (drinks per week) between recalled and concurrently reported alcohol intake among subgroups of participants adjusted for 2001 alcohol intake, sex, age, number of incident diseases, quartile of general frequency of forgetting, and quartile of SF-36 physical functioning ($n = 494$)

Variable	Mean diff.	[95% CI]
Alcohol intake		
2001, drinks/week		
0	-0.75*	[-1.79, 0.28]
1-7	-0.41*	[-1.32, 0.50]
8-14	0.34*	[-0.67, 1.34]
>14	0.97*	[-0.26, 2.21]
Sex		
Male	-0.75	[-1.79, 0.28]
Female	-0.93	[-2.53, 0.67]
Age, in years, in 2001		
<66	-0.74*	[-1.92, 0.43]
66-70	0.32*	[-0.88, 1.52]
71-75	-1.30*	[-2.45, -0.14]
>75	-1.50*	[-2.84, -0.16]
No. of incident diseases ^a		
0	-0.75	[-1.79, 0.28]
1	-0.27	[-1.43, 0.88]
2	-0.36	[-1.92, 1.20]
3	2.67	[-1.07, 6.41]
General freq. of forgetting 2000, percentiles		
>75th	-0.75	[-1.79, 0.28]
51st-75th	-0.62	[-1.66, 0.41]
26th-50th	0.10	[-1.02, 1.22]
≤25th	-1.39	[-2.51, -0.27]
SF-36 physical functioning 2001, percentiles		
>75th	-0.75	[-1.79, 0.28]
51st-75th	-1.97	[-3.31, -0.64]
26th-50th	-0.36	[-1.66, 0.95]
≤25th	-1.36	[-2.65, -0.07]

Notes: Diff. = difference; SF-36 = Short Form-36 Health Survey; freq. = frequency. ^aTotal number of incident diseases, including cardiovascular disease, hypertension, and diabetes.

*Significant for trend across categories ($p < .05$).

We examined associations with mean differences in recall using GEE in the total population (Table 3). BMI in 2001 was considered for inclusion in the multivariate analysis but was dropped because BMI was not associated with recall differences, and inclusion in the multivariate model did not affect the relationship between recall and level of alcohol intake in 2001. As shown in Table 3, those who were more likely to underestimate previous alcohol intake were older, reported worse memory function as evidenced by a general frequency of forgetting score in the lowest quartile, and reported worse physical functioning scores from the SF-36. The largest difference between concurrently reported and recalled intake, however, was less than three drinks per week, just more than a third of a drink per day.

As seen in Figure 3, although we observed significant differences in recall in both those who abstained from alco-

FIGURE 3. Difference between recalled and concurrently reported alcohol intake in drinks per week (and 95% confidence interval) by category of weekly total alcohol intake in 2001 adjusted for age, incident disease, and sex.

hol in 2001 and in those who drank 14 drinks per week or more in 2001 in univariate analyses, the reliability of recall of alcohol consumption was good across all levels of 2001 alcohol intake adjusting for age, sex, quartiles of general frequency of forgetting, and quartiles of physical functioning. When we limited the stratified analysis to men only, we obtained results similar to the combined analysis with both men and women.

Discussion

This study demonstrated a high level of reliability of recalled alcohol intake 15 and 23 years in the past. We found relatively small differences between recalled and concurrently reported intake of total and beverage-specific alcohol, with the largest difference slightly greater than one drink per week. Recall reliability was higher for wine than beer or distilled spirits. Overall, older age, difficulty with memory, and difficulty with physical functioning were associated with underestimation of alcohol intake.

The current study is one of the larger cohorts to assess reliability of recalled alcohol intake and has a longer duration of follow-up than previous studies. Our study is in agreement with previous studies that have demonstrated good overall reliability of long-term alcohol recall in epidemiological studies (Czarnecki et al., 1990; Dwyer et al., 1989; Lee et al., 1992; Liu et al., 1996). Only one previous study (Liu et al., 1996) has examined difference in recall by current drinking categories. This study was conducted in a nationally representative population sample from the First National Health and Nutrition Examination Survey and found that higher levels of current alcohol intake were associated with greater underestimation of recall 10 years earlier (0.5-0.75 drinks/day). In our study, higher levels of current alcohol consumption had no influence on recall of intake 15 and 23 years earlier, independent of age, sex, memory function, and physical function.

The major strengths of this study are the long duration of follow-up—up to 23 years—and multiple measures of alcohol recall over that duration. A limitation is the lack of

statistical power to analyze subjects with a history of alcohol misuse. We did not perform formal analyses of those with a history of alcohol misuse because our sample was limited to 72 individuals. Other limitations in this analysis include the unique characteristics of our study population (mostly older White men with high socioeconomic status and high levels of education) and a lack of generalizability to more demographically diverse populations. In addition, the Precursors Study participants, compared with the general population, have a lower prevalence of alcohol abstinence and a higher average level of alcohol intake among those who do consume alcohol. As former medical students, the participants of the Precursors Study may not only be more aware of the risks of alcohol consumption, but they may also be more aware of the benefits, factors that might affect both consumption and recall. Lastly, individuals who are part of a longitudinal study may be more reliable in their recall of alcohol use than the general population.

These results strengthen the conclusion of studies of alcohol recall in other populations; recalled alcohol intake, even over long periods, is reliable and can be used in epidemiological studies (Czarnecki et al., 1990; Dwyer et al., 1989; Liu et al., 1996). Ideally, concurrently obtained alcohol consumption data would be collected in such studies, but large prospective studies of this kind are expensive, require years to complete, and frequently are not feasible. In lieu of prospective cohort studies, case-control study designs based on past self-reported alcohol use have a role in examining the effects of lifetime alcohol use.

Acknowledgments

We thank all the participants of the Johns Hopkins Precursors Study.

References

- Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, *1*, 307-310.
- Bland, J. M., & Altman, D. G. (1995). Comparing methods of measurement: Why plotting difference against standard method is misleading. *Lancet*, *346*, 1085-1087.
- Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, *8*, 135-160.
- Bunch, T. J., White, R. D., Smith, G. E., Hodge, D. O., Gersh, B. J., Hammill, S. C., ... Packer, D. L. (2004). Long-term subjective memory function in ventricular fibrillation out-of-hospital cardiac arrest survivors resuscitated by early defibrillation. *Resuscitation*, *60*, 189-195.
- Czarnecki, D. M., Russell, M., Cooper, M. L., & Salter, D. (1990). Five-year reliability of self-reported alcohol consumption. *Journal of Studies on Alcohol*, *51*, 68-76.
- Del Boca, F. K., & Darkes, J. (2003). The validity of self-reports of alcohol consumption: State of the science and challenges for research. *Addiction*, *98* (Supplement No. 2), 1-12.
- Dwyer, J. T., Gardner, J., Halvorsen, K., Krall, E. A., Cohen, A., & Valadian, I. (1989). Memory of food intake in the distant past. *American Journal of Epidemiology*, *130*, 1033-1046.
- Ekholm, O. (2004). Influence of the recall period on self-reported alcohol intake. *European Journal of Clinical Nutrition*, *58*, 60-63.
- Embree, B. G., & Whitehead, P. C. (1993). Validity and reliability of self-reported drinking behavior: Dealing with the problem of response bias. *Journal of Studies on Alcohol*, *54*, 334-344.
- Ewing, J. A. (1984). Detecting alcoholism: The CAGE Questionnaire. *Journal of the American Medical Association*, *252*, 1905-1907.
- Gilewski, M. J., & Zelinski, E. M. (1988). Memory Functioning Questionnaire (MFQ). *Psychopharmacology Bulletin*, *24*, 665-670.
- Grant, K. A., Arciniega, L. T., Tonigan, J. S., Miller, W. R., & Meyers, R. J. (1997). Are reconstructed self-reports of drinking reliable? *Addiction*, *92*, 601-606.
- Klag, M. J., Ford, D. E., Mead, L. A., He, J., Whelton, P. K., Liang, K.-Y., & Levine, D. M. (1993a). Serum cholesterol in young men and subsequent cardiovascular disease. *New England Journal of Medicine*, *328*, 313-318.
- Klag, M. J., He, J., Mead, L. A., Ford, D. E., Pearson, T. A., & Levine, D. M. (1993b). Validity of physicians' self-reports of cardiovascular disease risk factors. *Annals of Epidemiology*, *3*, 442-447.
- Klipstein-Grobusch, K., Becker, N., Kroke, A., & Boeing, H. (1999). Patterns of past alcohol consumption in the EPIC-Germany cohorts: European Investigation into Cancer and Nutrition. *Annals of Nutrition and Metabolism*, *43*, 258-265.
- Lee, M. M., Whittemore, A. S., & Lung, D. L. (1992). Reliability of recalled physical activity, cigarette smoking, and alcohol consumption. *Annals of Epidemiology*, *2*, 705-714.
- Liang, K.-Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*, 13-22.
- Liu, S., Serdula, M. K., Byers, T., Williamson, D. F., Mokdad, A. H., & Flanders, W. D. (1996). Reliability of alcohol intake as recalled from 10 years in the past. *American Journal of Epidemiology*, *143*, 177-186.
- Longnecker, M. P., Newcomb, P. A., Mittendorf, R., Greenberg, E. R., Clapp, R. W., Bogdan, G., ... MacMahon, B. (1992). The reliability of self-reported alcohol consumption in the remote past. *Epidemiology*, *3*, 535-539.
- Poikolainen, K. (1985). Underestimation of recalled alcohol intake in relation to actual consumption. *British Journal of Addiction*, *80*, 215-216.
- Sobell, L. C. & Sobell, M. B. (1978). Validity of self-reports in three populations of alcoholics. *Journal of Consulting and Clinical Psychology*, *46*, 901-907.
- Thomas, C. B. (1951). Observations on some possible precursors of essential hypertension and coronary artery disease. *Bulletin of the Johns Hopkins Hospital*, *89*, 419-441.
- Ware, J. E., Jr., & Sherbourne, C. D. (1992). The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual framework and item selection. *Medical Care*, *30*, 473-483.
- Zelinski, E. M., Gilewski, M. J., & Anthony-Bergstone, C. R. (1990). Memory Functioning Questionnaire: Concurrent validity with memory performance and self-reported memory failures. *Psychology and Aging*, *5*, 388-399.