

SCREENING FOR HAZARDOUS OR HARMFUL DRINKING USING ONE OR TWO QUANTITY–FREQUENCY QUESTIONS

ANDREA CANAGASABY and DANIEL C. VINSON*

Department of Family and Community Medicine, University of Missouri-Columbia, Columbia, MO 65212, USA

(Received 3 November 2004; first review notified 16 December 2004; in revised form 8 March 2005; accepted 9 March 2005;
Advance Access publication 29 March 2005)

Abstract — Aims: To address the accuracy of quantity–frequency (QF) questions in screening for hazardous or harmful drinking. **Methods:** Three groups were interviewed: patients presenting to emergency departments for care of an acute injury ($n = 1537$) or a medical illness ($n = 1151$), and community controls interviewed by telephone ($n = 1112$). The first question about alcohol was a single alcohol screening question (SASQ), ‘When was the last time you had more than X drinks in one day?’, where $X = 4$ for women and 5 for men, with any time in the past 3 months considered a positive screen (1 drink = 14 g ethanol). The subsequent alcohol questions were a calendar-based review of recent drinking and the alcohol questions from the diagnostic interview schedule (DIS), which included questions about usual frequency and average quantity. Hazardous drinking was defined as drinking >4 drinks in 1 day or >14 drinks in 1 week for men (women 3 and 7) (Guidelines of the US National Institute on Alcohol Abuse and Alcoholism). Current alcohol use disorders were defined using DSM-IV criteria. The areas under the receiver operating characteristic (ROC) curves in identifying hazardous drinking or current alcohol use disorder were compared. **Results:** The area under the ROC curves in the three samples combined were 0.81 for SASQ (95% confidence interval (CI) 0.79–0.82), 0.80 for a question about average quantity alone (0.79–0.82) and 0.85 for the product of usual frequency times average quantity (0.84–0.86). The QF product and the question about average quantity performed consistently across the three groups. **Conclusions:** In clinical settings, one way to put these findings into practice is to screen first with a single question, such as the SASQ, a single question about typical quantity, or a question about the frequency of heavy drinking such as the third item of the alcohol use disorders test (AUDIT).

INTRODUCTION

Hazardous and harmful alcohol use are common. In the US, 17.6 million adults abuse alcohol or are alcohol dependent, an increase from 13.8 million only a decade ago (Grant *et al.*, 2004), and 85 000 deaths in the US can be attributed to alcohol annually (Mokdad *et al.*, 2004).

Up to one-fourth of patients in primary care settings may be hazardous or harmful drinkers (Whitlock *et al.*, 2004). Using brief interventional measures, primary care clinicians can help 40% of the harmful drinkers (compared with 20% in control groups) to reduce their drinking to safe levels (Wallace *et al.*, 1988; Fleming *et al.*, 1997), with documented improvements in health outcomes in one study shown even 4 years later (Fleming *et al.*, 2002). Hazardous and harmful drinkers, however, often go undetected (Spandorfer *et al.*, 1999; Vinson *et al.*, 2000; McGlynn *et al.*, 2003; Rush *et al.*, 2003). Screening is therefore necessary to identify the individuals who would benefit from intervention (U.S. Preventive Services Task Force, 2004).

Several studies have examined the sensitivity and specificity of various screening instruments (Fiellin *et al.*, 2000). The 4-item cutting down, annoyance by criticism, guilty feeling and eye-openers (CAGE) and the 10-item alcohol use disorders identification test (AUDIT) are widely recommended. With the CAGE, the area under a receiver operating characteristic (ROC) curve is lower in outpatient settings (0.60–0.71) than among hospitalized patients (0.87) (Aertgeerts *et al.*, 2004). In another study, when the CAGE was augmented to include quantity–frequency (QF) questions, the area under the ROC curve increased from 0.73 to 0.78 (Bradley *et al.*, 2001). In identifying hazardous drinking or

current alcohol abuse or dependence, the AUDIT outperformed both the CAGE and the augmented CAGE (areas under the ROC curves of 0.86, 0.72 and 0.77, respectively) (Bradley *et al.*, 1998).

The first three AUDIT items, called AUDIT-C, ask about the usual frequency of drinking, typical quantity and frequency of heavy drinking. It performed as well as the full AUDIT in detecting heavy drinking, with an area under the ROC curve of 0.89 vs 0.88, but had a somewhat lower area under the ROC curve for detecting active alcohol abuse and dependence alone (0.79) than the full AUDIT (0.81) (Bush *et al.*, 1998). These findings have been confirmed in several European studies. In Germany, the AUDIT-C had an area under the ROC curve of 0.88 in detecting hazardous or harmful drinking (Rumpf *et al.*, 2002); in Belgium, the areas under the ROC curves were 0.83 for men and 0.82 for women in detecting hazardous or harmful drinking; and in Spain, the areas under the ROC curves were 0.91 for men and 0.96 for women in detecting physician-diagnosed hazardous drinking (Gual *et al.*, 2002).

Although it was asked in the research interviews in the context of the full AUDIT, the third AUDIT question by itself, ‘How often do you have 6 or more drinks on one occasion,’ has also been examined as a single screening question. It had an area under the ROC curve of 0.83 in detecting current alcohol use disorder or hazardous drinking in men (Bush *et al.*, 1998) and 0.76 in women (Bradley *et al.*, 2003). With a modified version of the question (‘4 or more drinks’), the area under the ROC curve in women increased to 0.86 (Bradley *et al.*, 2003).

Using the same dataset examined here, we previously reported (Williams and Vinson, 2001) a study of a single alcohol screening question (SASQ), ‘When was the last time you had more than X drinks in 1 day,’ with $X = 4$ for women and 5 for men (one standard drink in the US contains 14 g ethanol). Based on previous work (Taj *et al.*, 1998), the

*Author to whom correspondence should be addressed at: Daniel C. Vinson, Department of Family and Community Medicine, University of Missouri-Columbia, MO 65212, USA. Tel.: +1 573 882 3184; Fax: +1 573 884 5734; E-mail: VinsonD@health.missouri.edu

threshold numbers were set one drink higher than usually recommended (National Institute on Alcohol Abuse and Alcoholism, 2003) to balance sensitivity and specificity. Open-ended answers were coded as never, >12 months ago, 3–12 months ago, and within the last 3 months, with the last category considered a positive screen. Both sensitivity and specificity were 0.86 in identifying the past-month hazardous drinking, current alcohol use disorders, or both. The area under the ROC curve was 0.90 (Williams and Vinson, 2001).

Questions about quantity or frequency of drinking may have reasonable sensitivity and specificity, but not all studies have supported their use. In their systematic review, Fiellin *et al.* (2000) concluded that the CAGE and AUDIT 'consistently performed better than other methods, including QF questions'. Using data from a population-based US study, Dawson concluded that 'single- or dual-item indicators of alcohol consumption have limited value as screeners of alcohol dependence' (Dawson, 1994).

Finding time to screen for alcohol problems in a busy primary care clinic is challenging, and only a third of primary care doctors regularly screen patients for alcohol problems (The National Center on Addiction and Substance Abuse at Columbia University, 2000; Aertgeerts *et al.*, 2001). When doctors do screen patients, they generally use QF questions, not the CAGE or the AUDIT (Spandorfer *et al.*, 1999). We therefore decided to examine more closely the utility of two standard QF questions in detecting hazardous or harmful drinking, comparing them with the SASQ.

SUBJECTS AND METHODS

Data are from a case-control study of alcohol and injury, which was approved by the Institutional Review Boards of all participating institutions. Cases were recruited from among patients presenting for care to one of the three hospital emergency departments in Columbia, Missouri, within 48 hours of an acute injury. Patients were eligible for the study if they were at least 18 years old, able to converse in English, cognitively intact, not in police custody, and if the injury did not occur in a controlled environment (e.g. a nursing home or jail, where access to alcohol is limited). Research staff trained in the use of the structured interview covered each day of the week and hour of the day in a systematically representative fashion over the course of data collection from February 1998 through March 2000. Among patients approached, 87% participated.

Cases were matched with two separate groups of control participants by age group, sex and rural versus urban. Inclusion criteria were the same as for cases. One group of controls was patients presenting to the same emergency departments for care of a medical illness; 88% of those approached participated. The second group was recruited by random-digit dialing using the telephone exchanges of Boone County and contiguous counties, with a response rate of 47%.

Instruments

The first question of the structured interview asked about tobacco use in all the three groups of participants. The second Question was the SASQ, quoted above. The subsequent questions about alcohol were a timeline follow-back

interview, a retrospective calendar-based review of day-by-day consumption (Sobell and Sobell, 1992). We reviewed the past 28 days for injured cases and telephone controls and the past 8 days for the medical controls (keeping their interview shorter to avoid missing cases).

Following the timeline, the interviewer asked questions from the structured diagnostic interview schedule (DIS) (Robins *et al.*, 1996). The first DIS question was, 'In the last year, have you had 6 or more drinks?' Participants who twice denied drinking ≥ 6 drinks in the entire previous year were not asked the remaining DIS questions. The next DIS question addressed the frequency of drinking, with answers on a 5-point ordinal scale from 'less than once a month' to 'almost every day.' The third DIS question asked the 'average' number of drinks per occasion, with the answer recorded as the number given by the respondent. Using these questions, we created a QF measure by multiplying the usual frequency of drinking times average quantity. We also examined the utility of the single question about average quantity.

We identified current (past 12 months) alcohol abuse and dependence using the criteria in the 4th edition of the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 1994), based on answers from the DIS. Hazardous drinking was defined as drinking >4 drinks in 1 day or >14 drinks in 1 week for men, >3 drinks in 1 day or 7 drinks in 1 week for women (National Institute on Alcohol Abuse and Alcoholism, 2003).

Statistical analysis

We used Stata (Stata Corporation, 2003) to estimate the area under the ROC curve (Jaeschke *et al.*, 1994) with 95% confidence intervals (CIs), directly comparing the ability to identify hazardous drinking and/or alcohol use disorders of three screening approaches: the single alcohol screening question, the QF product calculated from the DIS questions and the question about average quantity alone. Analyses took into account the correlations caused by applying the three screening approaches to the same respondents.

RESULTS

Of 2517 cases, 1537 (61%) acknowledged drinking at least a total of 6 drinks in the previous 12 months and provided complete DIS data. Table 1 shows the number of men and women in each sample group with hazardous drinking and/or alcohol use disorders. Because injuries are more common among younger adults and men, and because those demographic groups are also more likely to have hazardous drinking and alcohol use disorders, the prevalence of alcohol problems was also high among the controls who were matched by age group and sex to the injured cases.

In identifying those with either a current alcohol use disorder, recent hazardous drinking, or both, the area under the ROC curve for the QF product was higher (0.85) than for the SASQ (0.81) or the quantity question alone (0.80) (Table 2). The ROC curves overlapped in all three samples (Figure 1) at regions of high sensitivity and low specificity.

The QF questions were significantly better than SASQ and the quantity question in two groups (cases and medical controls), and SASQ was better than QF among telephone

controls, but these differences were small (details available from the corresponding author). For each screening approach, the similarities across the three samples were remarkable considering they were in different places (emergency departments or at home) and interviewed in different ways (face to face or by telephone). Differences among subgroups defined by gender and age were generally minor.

The numbers of participants from ethnic groups other than African-American and Caucasian were too small to allow meaningful comparisons. Combining all three samples and comparing African-Americans ($n = 330$) and Caucasians ($n = 3262$), the areas under the ROC curves were almost identical for the QF product (0.84 and 0.85, respectively) and for the quantity question (0.80 and 0.81, respectively), but different for the SASQ: 0.74 (95% CI 0.69–0.79) among African-Americans and 0.81 (0.80 to 0.83) among Caucasians.

Using all three samples combined, Table 3 shows the sensitivity and specificity of each screening test at selected threshold values for the SASQ, the QF product and the quantity question.

Table 1. Numbers of participants in the three samples with hazardous drinking and/or current alcohol use disorder

	Hazardous drinking by gender-specific per-day or per-week thresholds			
	No		Yes	
	Current alcohol use disorder (abuse or dependence)		Current alcohol use disorder (abuse or dependence)	
	No	Yes	No	Yes
Cases				
Men	412	93	228	265
Women	261	47	146	85
Medical controls				
Men	311	93	141	138
Women	310	60	43	55
Telephone controls				
Men	312	67	122	135
Women	286	39	87	64

Table 2. Area under the receiver operating characteristic curves (AUROC) in identifying current alcohol use disorder and/or recent hazardous drinking in three samples of respondents

	<i>n</i>	SASQ		Usual QF		Quantity alone	
		AUROC	95% CI	AUROC	95% CI	AUROC	95% CI
Cases	1537	0.79	0.77–0.82	0.86	0.84–0.88	0.81	0.79–0.83
Medical controls	1151	0.77	0.75–0.80	0.85	0.83–0.87	0.81	0.78–0.83
Telephone controls	1112	0.85	0.83–0.87	0.84	0.82–0.86	0.81	0.78–0.83
All three samples	3800						
In detecting hazardous drinking		0.82	0.81–0.83	0.84	0.83–0.85	0.79	0.77–0.80
In detecting alcohol abuse or dependence		0.70	0.68–0.71	0.79	0.78–0.81	0.75	0.73–0.77
In detecting either		0.81	0.79–0.82	0.85	0.84–0.86	0.80	0.79–0.82

SASQ is the single alcohol screening question ('When was the last time you had more than X drinks in one day?', where X = 4 for women, 5 for men). Usual QF is the product of a question about usual frequency of drinking on a 5-point ordinal scale times average quantity consumed in standard drinks. Quantity alone is the question about average quantity.

DISCUSSION

Both standard QF questions and the SASQ may be effective screening tools in detecting alcohol use disorders. The major strength of this study was the use of well-validated criterion standards, the DIS for detecting current alcohol abuse and dependence and the timeline follow-back interview for identifying recent hazardous drinking. Furthermore, the three samples were >1000 and totaled 3800. Despite methodological differences in the way the data were collected, results in the three samples were very similar.

Our previous study found that the SASQ performed well in identifying persons with either hazardous drinking or a current alcohol use disorder, with an area under the ROC curve of 0.90 (95% CI 0.88–0.91) (Williams and Vinson, 2001). The analyses reported then assumed that cases who did not answer the DIS questions (because they twice denied drinking ≥ 6 drinks in the entire previous 12 months) did not have an alcohol use disorder. The value for the area under the ROC curve presented here is lower because non-drinkers were excluded from the analyses.

Using QF questions in screening for alcohol problems in clinical settings is feasible, though not straightforward. First, when using the quantity question alone, one must use a threshold of ≥ 3 drinks/occasion to have a sensitivity of 77%. At a threshold of ≥ 4 , sensitivity is substantially lower (Table 3). Second, the frequency question used an ordinal scale, and remembering that scale would make screening with the QF product harder. Third, some QF results can be produced in different ways and thus indicate different patterns of drinking. For example, a QF score of 6 could be derived from 6 drinks less than once a month, 3 drinks 1–3 days a month, or 2 drinks once or twice a week. This ambiguity may dissuade clinicians from using it.

Other screening tests also use questions about quantity, frequency or maximal quantity. The 10-item AUDIT (Saunders *et al.*, 1993) starts with three questions about frequency, typical quantity and frequency of drinking ≥ 6 drinks on one occasion. Used as a separate scale, these three questions also have good psychometric properties in both men (Bush *et al.*, 1998) and women (Bradley *et al.*, 2003). The fast alcohol screening test (Hodgson *et al.*, 2002) starts with the

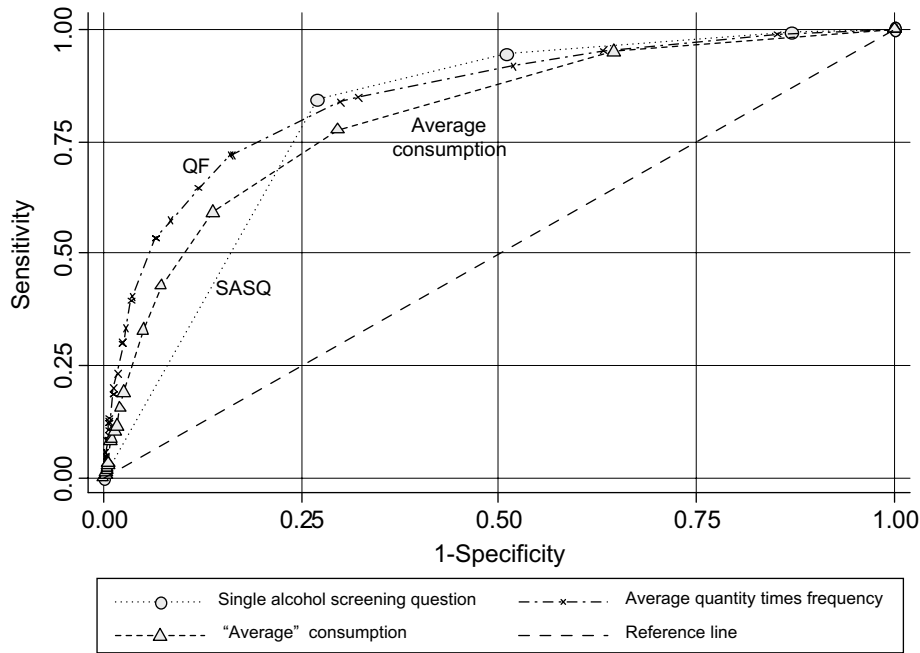


Fig.1. ROC curves in identifying current alcohol use disorder, recent hazardous drinking, or both in all three samples combined comparing the SASQ, the QF, and the quantity question alone.

Table 3. Sensitivity and specificity of the three screening approaches at selected threshold values in all three samples combined (*n* = 3800)

	Men		Women	
	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
SASQ				
Had more than X drinks ever	99	10	99	16
Had > X drinks in the last 12 months	96	47	92	51
Had > X drinks in the last 3 months	85	70	82	77
QF product				
QF ≥ 2	99	13	98	17
QF ≥ 4	94	42	88	56
QF ≥ 6	87	64	78	78
QF ≥ 8	78	78	61	92
QF ≥ 10	65	88	42	96
QF ≥ 12	61	90	38	97
Quantity question alone				
≥2 drinks	96	32	94	40
≥3 drinks	81	66	71	76
≥4 drinks	64	82	50	92

The SASQ is ‘When was the last time you had more than X drinks in one day?’, where X = 4 for women, 5 for men. The QF product multiplies the ‘average’ number of drinks consumed on one occasion by the usual frequency of drinking expressed on an ordinal scale from 1 (‘less than once a month’) to 5 (‘almost every day’).

third AUDIT question. That question alone was able to correctly classify about half of the patients as hazardous or non-hazardous drinkers in their study (Hodgson *et al.*, 2002). The Paddington alcohol test (Patton *et al.*, 2004) asks patients

who acknowledge any drinking, ‘what is the most you will drink in any one day?’, then asks about the frequency of heavy drinking. Using the AUDIT as the criterion standard, it had a sensitivity of 97% and specificity of 88% (Patton *et al.*, 2004).

Limitations

We studied the utility of these three screening approaches in identifying current alcohol use disorders or recent hazardous drinking. All three would be less effective in identifying patients with a past alcohol use disorder who are not currently drinking hazardous amounts.

The structure of the interview may have influenced the results. The detailed review of recent consumption came immediately before the questions about average quantity and usual frequency. That may have enabled respondents to recall their drinking more accurately when asked the QF questions than when asked the SASQ, which was asked at the beginning of the interview. However, the results were similar for the medical controls, who had an 8-day timeline, and the cases and telephone controls who had a 28-day timeline, suggesting that the differences seen here were not owing solely to the effect of doing a timeline right before the QF questions. Alternatively, an 8-day timeline may be sufficient to enhance the utility of subsequently asked QF questions.

The amount of data used in applying the criterion standard to identify hazardous drinkers was limited and varied by group. This may have led to true-positive screening results among individuals whose hazardous drinking was not confirmed because the timeline covered only the last 28 days or, even more a problem, 8 days.

Of the individuals who answered both the SASQ and QF questions, 9.2% were African-American, which is similar to the 8.5% African-American population in central Missouri, but lower than the 12.3% US average. Future research could

develop and validate other rapid screening approaches based on quantity and/or frequency that have greater utility among ethnicities other than Caucasians.

The non-participation rate varied among the three samples. The effects of non-response bias are uncertain, particularly among the telephone controls, but the similarity of the findings across all three groups suggest that the non-response did not bias the results substantially.

CONCLUSIONS

A growing body of research findings and the current study suggest that QF questions may have an utility in detecting individuals with hazardous drinking or alcohol use disorders. In clinical settings, one way to put these findings into practice is to screen first with a single question (the SASQ, a single question about typical quantity, or a question about the frequency of heavy drinking such as the third AUDIT item). Using a low and therefore sensitive threshold, patients who screen positive would then be asked further screening questions (the full AUDIT, for example), followed by exploration of alcohol-related consequences and symptoms of dependence.

Acknowledgements — Malcolm Maclure and Gordon Smith, helped greatly with the design, analysis, and interpretation of the parent study. Nancy Mabe helped in the design of the preliminary study on which the current work built. Interviews of cases were conducted by Carol Reidinger, Carey Smith, Ciprian Crismaru, Amelia Devera-Sales, Indira Gujral, Kari Gilmore, and Lindsay Wiles, Aneesh Tosh, Stephen Griffith, Darin Lee, Greg Morlin, and Rebecca Shumate, who were medical students at the time. Data management was by Sandy Taylor, Darla Horman, Robin Kruse and Carol Reidinger. Telephone interviews of controls were conducted by Research Triangle Institute, Inc., Research Triangle Park, N.C. 27709. We are very grateful to our colleagues in the emergency departments who helped make this study possible. The study was funded by a grant (R01 AA11078) from the National Institute on Alcohol Abuse and Alcoholism.

REFERENCES

Aertgeerts, B., Buntinx, F., Ansoms, S. *et al.* (2001) Screening properties of questionnaires and laboratory tests for the detection of alcohol abuse or dependence in a general practice population. *British Journal of General Practice* **51**, 206–217.

Aertgeerts, B., Buntinx, F. and Kester, A. (2004) The value of the CAGE in screening for alcohol abuse and alcohol dependence in general clinical populations: a diagnostic meta-analysis. *Journal of Clinical Epidemiology* **57**, 30–39.

American Psychiatric Association (1994) Substance-related disorders. In *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn, pp. 175–204. American Psychiatric Association, Washington, DC.

Bradley, K. A., Bush, K. R., Mcdonell, M. B. *et al.* (1998) Screening for problem drinking: comparison of CAGE and AUDIT. *Journal of General Internal Medicine* **13**, 379–388.

Bradley, K. A., Kivlahan, D. R., Bush, K. R. *et al.* (2001) Variations on the CAGE alcohol screening questionnaire: strengths and limitations in VA general medical patients. *Alcoholism: Clinical and Experimental Research* **25**, 1472–1478.

Bradley, K. A., Bush, K. R., Epler, A. J. *et al.* (2003) Two brief alcohol-screening tests from the Alcohol Use Disorders Identification Test (AUDIT): validation in a female veterans affairs patient population. *Archives of Internal Medicine* **163**, 821–829.

Bush, K., Kivlahan, D. R., Mcdonell, M. B. *et al.* (1998) The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. *Archives of Internal Medicine* **158**, 1789–1795.

Dawson, D. A. (1994) Consumption indicators of alcohol dependence. *Addiction* **89**, 345–350.

Fiellin, D. A., Reid, M. C. and O'Connor, P. G. (2000) Screening for alcohol problems in primary care: a systematic review. *Archives of Internal Medicine* **160**, 1977–1989.

Fleming, M. F., Barry, K. L., Manwell, L. B. *et al.* (1997) Brief physician advice for problem alcohol drinkers: a randomized controlled trial in community-based primary care practices. *Journal of the American Medical Association* **277**, 1039–1044.

Fleming, M. F., Mundt, M. P., French, M. T. *et al.* (2002) Brief physician advice for problem drinkers: long-term efficacy and benefit-cost analysis. *Alcoholism: Clinical and Experimental Research* **26**, 36–43.

Grant, B. F., Dawson, D. A., Stinson, F. S. *et al.* (2004) The 12-month prevalence and trends in DSM-IV alcohol abuse and dependence: United States, 1991–1992 and 2001–2002. *Drug and Alcohol Dependence* **74**, 223–234.

Gual, A., Segura, L., Contel, M. *et al.* (2002) AUDIT-3 and AUDIT-4: effectiveness of two short forms of the Alcohol Use Disorders Identification Test. *Alcohol and Alcoholism* **37**, 591–596.

Hodgson, R., Alwyn, T., John, B. *et al.* (2002) The Fast Alcohol Screening Test. *Alcohol and Alcoholism* **37**, 61–66.

Jaeschke, R., Guyatt, G. H. and Sackett, D. L. (1994) Users' guides to the medical literature: III. How to use an article about a diagnostic test: B. What are the results and will they help me in caring for my patients? *Journal of the American Medical Association* **271**, 703–707.

McGlynn, E. A., Asch, S. M., Adams, J. *et al.* (2003) The quality of health care delivered to adults in the United States. *The New England Journal of Medicine* **348**, 2635.

Mokdad, A. H., Marks, J. S., Stroup, D. F. *et al.* (2004) Actual causes of death in the United States, 2000. *Journal of the American Medical Association* **291**, 1238–1245.

National Institute on Alcohol Abuse and Alcoholism (2003) *Helping Patients with Alcohol Problems: A Health Practitioner's Guide*. U.S. Department of Health and Human Services, Washington, DC.

Patton, R., Hilton, C., Crawford, M. J. *et al.* (2004) The Paddington alcohol test: a short report. *Alcohol and Alcoholism* **39**, 266–268.

Robins, L., Cottler, L., Bucholz, K. *et al.* (1996) *Diagnostic Interview Schedule for DSM-IV*. Washington University School of Medicine, Department of Psychiatry, St Louis, MI.

Rumpf, H. J., Hapke, U., Meyer, C. *et al.* (2002) Screening for alcohol use disorders and at-risk drinking in the general population: psychometric performance of three questionnaires. *Alcohol and Alcoholism* **37**, 261–268.

Rush, B. R., Urbanoski, K. A. and Allen, B. A. (2003) Physicians' enquiries into their patients' alcohol use: public views and recalled experiences. *Addiction* **98**, 895–900.

Sobell, L. C. and Sobell, M. B. (1992) Timeline Follow-back: a technique for assessing self-reported alcohol consumption. In *Measuring Alcohol Consumption: Psychosocial and Biochemical Methods*, Litten, R. and Allen, J. eds, pp. 41–72. Humana Press, Totowa, NJ.

Spandorfer, J. M., Israel, Y. and Turner, B. J. (1999) Primary care physicians' views on screening and management of alcohol abuse: inconsistencies with national guidelines. *Journal of Family Practice* **48**, 899–902.

Stata Corporation (2003) *Stata statistical software: Release 8.1*. College Station. Stata Corporation, TX.

Saunders, J. B., Aasland, O. G., Babor, T. F. *et al.* (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. *Addiction* **88**, 791–804.

Taj, N., Devera-Sales, A. and Vinson, D. C. (1998) Screening for problem drinking: does a single question work? *Journal of Family Practice* **46**, 328–335.

The National Center on Addiction and substance Abuse at Columbia University (2000) *Missed Opportunity: National Survey of Primary Care Physicians and Patients on Substance Abuse*. Last accessed November 2, 2004, from <http://www.casacolumbia.org/pdshopprov/shop/item.asp?itemid=9>

U.S. Preventive Services Task Force (2004) Screening and behavioral counseling interventions in primary care to reduce alcohol misuse: recommendation statement. *Annals of Internal Medicine* **140**, 554–556.

- Vinson, D. C., Elder, N. C., Werner, J. J. *et al.* (2000) Alcohol-related discussions in primary care: a report from ASPN. *Journal of Family Practice* **49**, 28–33.
- Wallace, P., Cutler, S. and Haines, A. (1988) Randomised controlled trial of general practitioner intervention in patients with excessive alcohol consumption. *British Medical Journal* **297**, 663–668.
- Whitlock, E. P., Polen, M. R., Green, C. A. *et al.* (2004) Behavioral counseling interventions in primary care to reduce risky/harmful alcohol use by adults: a summary of the evidence for the U.S. Preventive Services Task Force. *Annals of Internal Medicine* **140**, 557–568.
- Williams, R. H. and Vinson, D. C. (2001) Validation of a single question screen for problem drinking. *Journal of Family Practice* **50**, 307–312.