0145-6008/93/1706-1188$3.00/0
ALCOHOLISM: CLINICAL AND EXPERIMENTAL RESEARCH

Vol. 17, No. 6
November/December 1993

Use of the TWEAK Test in Screening for Alcoholism/ Heavy Drinking in Three Populations

Arthur W. K. Chan, Edward A. Pristach, John W. Welte, and Marcia Russell

TWEAK is an acronym for *Tolerance* (T_1 number of drinks to feel high; T_2, number of drinks one can hold), *Worry* about drinking, *Eye-opener* (morning drinking), *Amnesia* (blackouts), and *Cut* down on drinking (K/C). In this study, two versions (T_1 and T_2) of the TWEAK were part of a questionnaire used to detect alcoholism or heavy alcohol intake in three populations, namely, alcoholics in treatment, patients in two outpatient clinics, and the general population. Similar to the CAGE and the 10-item brief MAST, the TWEAK identified most known alcoholics, but the TWEAK had a higher sensitivity and specificity than the CAGE and B-MAST in detecting alcoholism/heavy drinking in the clinical and general populations. Different cut-off values for tolerance (T_1 and T_2) are recommended for screening different populations.

Key Words: TWEAK, CAGE, Brief-MAST, Clinical Outpatients, General Population.

ALCOHOLICS EXPERIENCE many drinking problems before seeking professional help.[1] Unfortunately, doctors who are uniquely placed to detect problem drinking in their patients, often fail to identify the majority of them.[2-6] One possible reason is that doctors may be reluctant to spend time administering lengthy questionnaires or to personally conduct a structured interview.[7] Therefore, the availability of brief and reliable screening instruments to detect alcohol problems will increase the likelihood of doctors using these tools in their clinical practices.

Currently, there are several short screening tests that medical practitioners can use to detect hazardous drinking/probable alcoholism in their patients. One example is the Cyr and Wartman[8] 2-item questionnaire ["Have you ever had a drinking problem?" "When was your last drink?" (24-hr cut-off)], which was reported to have a sensitivity of 91.5% and a specificity of 89.7%. Using this instrument, Goldberg et al.[7] reported that 35.6% of the patients in an academic, general medical clinic were screened positive. More importantly, over four times as many patients screened this way (10.8% vs. 2.3%) accepted referrals for counseling, as did patients seen by doctors providing standard care. Studies of elderly veterans and women have indicated that this test was less useful in these populations.[2,9]

CAGE is an acronym for: *Cut* down on drinking, *Annoyed* by criticism of drinking, feeling *Guilty* about drinking, and *Eye-opener* (morning drinking), with two or more positive responses being indicative of problem drinking or probable alcoholism.[10] The advantages of the CAGE are its brevity and high clinical validity.[11] Thus, many studies have shown that the CAGE identified most alcoholics and excessive drinkers.[10-16] On the other hand, some studies have reported relatively low sensitivity for the CAGE in other selected populations, such as college students,[17] DWI offenders,[18] emergency room patients,[19] and elderly medical patients.[20] The disadvantages of the CAGE are that it focuses on lifetime rather than current problems, because of the phrasing "Have you ever. . .," and that the questions about "cutting down" and "feeling guilty" are often answered "YES" by current light drinkers or abstainers. These may lead to an overestimation of alcohol-related problems by the CAGE.[21] There have only been two studies that used the CAGE in a general population survey of drinking.[11,22]

The T-ACE questionnaire[23] retained three CAGE items, namely, A, C, and E. T stands for tolerance to alcohol that was defined as needing two or more drinks to make a female subject feel "high." When used in the prenatal detection of risk-drinking in gravid women, it correctly identified 69% of the risk drinkers.[23] In another study of screening for risk-drinking in pregnant women,[24] the sensitivity of the T-ACE was 60–89%, and the specificity was 80–86%, depending on which one of the two definitions of tolerance (T_1 or T_2) was used: number of drinks needed to get high ($T_1 \geq 3$ drinks) or the number of drinks one could hold ($T_2 \geq 5$ drinks). Other variations of the T-ACE and CAGE are the 3-item NET (N, whether one is a normal drinker; E, eye-opener; T, tolerance) and the 5-item TWEAK [T, tolerance; W, worry about drinking; E, eye-opener; A, amnesia (blackouts); and K/C, cut down].[24] When applied to the screening of risk-drinking in pregnant women, the TWEAK (T_1 high or T_2 hold) had a sensitivity of 68–79% and a specificity of 83–93%. In contrast, the NET (high) had a 100% sensitivity but only 22% specificity.[24] There has been no report dealing with the use of the T-ACE, NET, or TWEAK in detecting problem drinking in the general population.

In the present study, the TWEAK is part of a questionnaire used to detect alcoholism or heavy alcohol intake in

From the Research Institute on Addictions, New York State Office of Alcoholism and Substance Abuse Services, Buffalo, New York.

Received for publication February 25, 1993; accepted June 1, 1993

This work was supported in part by Public Health Service Grant AA07794.

Reprint requests: Arthur W. K. Chan, Ph.D., Research Institute on Addictions, 1021 Main Street, Buffalo, NY 14203-1016.

three populations, namely, alcoholics in treatment, patients in two outpatient clinics in a county hospital, and a general household population. The usefulness of the TWEAK is compared with that of the CAGE, and the 10-item brief MAST.[25]

METHODS

Subject Recruitment

Alcoholics admitted to the Alcoholism Treatment Center in a County Medical Center in Buffalo were randomly recruited for the study by the staff of the treatment center. To be eligible, they must have had their last drink no longer than 4 days before the interview. This eligibility requirement was not expected to affect the generalizability of the screening test results in this study, but was intended to result in a uniform control group for the determination of serum levels of carbohydrate-deficient transferrin (CDT; data to be presented elsewhere); levels of CDT usually decrease during abstinence. Outpatients from the Primary Health Care Center and the Family Care Center in the same County Medical Center were recruited via telephone by the project staff at least several days before their scheduled appointments at the clinics. These patients mostly belonged to lower- or lower middle-class households. A random-digit dial telephone interview was conducted on household occupants residing in the Buffalo metropolitan area. We used a computer-assisted telephone interviewing (Sawtooth Software Incorporated, 1991) system for the random selection and dialing of telephone numbers. The procedures for oversampling Blacks and heavy drinkers were the same as previously described.[26] Briefly, Blacks were oversampled by using a higher proportion of phone numbers in telephone districts with high Black populations. Blacks were oversampled to permit adequate numbers for analysis of possible ethnic differences in blood chemistry items or in responses to questionnaire items. Because there were far fewer female heavy drinkers (those consuming two or more drinks/day) than male drinkers, female drinkers were oversampled by using biased screening tables such that the average probability of selecting a female from a household was 6–8 times higher than that of selecting a male. To produce unbiased samples of the general population in data analysis, each case was weighted inversely proportional to its probability of selection, taking into account these factors.

Informed consent was obtained from all three groups of subjects before their participation in a face-to-face interview, during which a questionnaire was completed. The age limit was 18–65 years, inclusive. There were 1,635 subjects, 252 alcoholics in treatment (ALC), 390 clinical outpatients (CL), and 993 general population (GP) subjects. The demographic characteristics of these subjects are summarized in Table 1. The GP subjects were significantly younger than the CL and ALC subjects ($p < 0.001$). The male:female ratio for the ALC was ~3:1, and that for the CL was ~1:2, but for the GP it was close to 1:1, due to the oversampling of female drinkers. The distribution of White and Black subjects was similar in the GP and ALC samples, being 63.4% White in the former and 64.3% in the latter. In contrast, the CL sample was composed of 41.5% White and 57.4% Black subjects. Other ethnic groups (American Indians and Orientals) constituted between 1.1–2.1% of each of the subject groups. The educational levels of the CL and ALC subjects were very similar, with 34–39% having some college education; but 60.4% of the GP subjects had some college education. Among the CL and GP subjects, 14.1% and 11.7%, respectively, were heavy drinkers.

Table 1. Demographic Characteristics of Three Groups of Subjects

Variable	GP	CL	ALC
n	993	390	252
Age (years)*	34.7 ± 3.8†	41.3 ± 0.61	37.6 ± 0.60
Male	48.3%	33.6%	73.0%
Female	51.7%	66.4%	27.0%
White	63.4%	41.8%	64.3%
Black	35.0%	57.2%	33.3%
Others	1.6%	1.0%	2.4%
Years of schooling*	13.2 ± 0.07‡	12.2 ± 0.24	12.1 ± 0.12

* Mean ± SE.
† Significantly different from the other two groups ($p < 0.01$).
‡ Significantly different from the other two groups ($p < 0.001$).

Questionnaire

All subjects received a face-to-face interview by research interviewers who were trained to administer a standardized questionnaire. The questionnaire elicited information on demographic characteristics such as age, sex, race, occupation, and education; beverage-specific quantity-frequency questions for beer, wine, and liquor in the past year, 5 years ago, and 10 years ago; age of first drink and time (date) of last drink; use of nonprescription drugs in the last 2 weeks and in the past year; current use of prescription drugs; family history of alcoholism; the 4-item CAGE questionnaire;[10,15] the 10-item Brief MAST;[25] the TWEAK;[24] age of onset of smoking; quantity and frequency of smoking for last year; lifetime record of diseases such as hepatitis, cirrhosis, and other liver diseases, seizures, etc.; recent significant health problems; depression[27]; and anxiety disorders.[28] Some subjects in each of the three groups (ALC, $n = 53$; CL, $n = 84$; GP, $n = 369$) were also asked questions pertaining to lifetime and current (past year) DSM-III-R criteria for alcohol dependence.

The TWEAK contained the following questions:

1. How many drinks does it take before you begin to feel the first effects of the alcohol?
2. How many drinks does it take before the alcohol makes you fall asleep or pass out? Or, if you never drink until you pass out, what is the largest number of drinks you have?
3. Have your friends or relatives *Worried* or complained about your drinking in the past year?
4. Do you sometimes take a drink in the morning when you first get up? (*Eye-opener*)
5. Are there times when you drink and afterwards you can't remember what you said or did? (*Amnesia*)
6. Do you sometimes feel the need to *Cut* down on your drinking? (K or C)

Questions 1 (T_1, high) and 2 (T_2, hold) were the two versions of tolerance described in the introduction. For initial data analysis, we used a cut-off of ≥3 drinks for T_1 and ≥ 5 drinks for T_2, giving us two versions of TWEAK, namely, T_1WEAK and T_2WEAK. As described in "Results," we also experimented with other cut-offs for T_1 and T_2. As suggested by Russell et al.,[24] both T and W were scored two points, whereas each of the other items (E, A, K) was scored 1 point. A total score of 3 or more was considered positive for T_1WEAK or T_2WEAK.

Drinker Definitions

Based on their self-reports of drinking for the past year, the clinical outpatients and general population subjects were stratified into the following drinker categories: Heavy drinkers were males consuming an average of six or more drinks/day and females consuming four or more drinks/day. Individuals drinking less than these amounts were considered nonheavy drinkers. Abstainers were included as nonheavy drinkers in the data analysis.

Statistical Analysis

Data were analyzed by the Statistical Package for the Social Sciences (SPSS-X) using one-way ANOVA, χ^2, multiple range test (Neuman-Keuls), and correlation analysis procedure, where appropriate. The level of significance was defined at $p < 0.05$ throughout. Sensitivity and specificity were computed using standard methodology.[29] A Receiver

Operating Characteristic (ROC) curve was constructed as described by Hsiao et al.[30]

RESULTS

Either version of the TWEAK (T_1WEAK and T_2WEAK) correctly identified 96.2% of the ALC. In comparison, the CAGE and B-MAST identified 98.3% and 99.2% of the ALC, respectively. The sensitivities and specificities of T_1WEAK and T_2WEAK in the CL and GP groups are compared with those of the CAGE and B-MAST in Table 2. Based on the DSM-III-R gold standard (A), the two versions of TWEAK had higher sensitivity and specificity than the CAGE or B-MAST, except for the higher sensitivity for CAGE in CL and the higher specificity of B-MAST in GP. With heavy drinking as the gold standard (B), the two versions of TWEAK correctly identified more heavy drinkers than the CAGE or B-MAST, and still had the lowest false positive rates except for the higher specificity of B-MAST in GP. The sensitivity of T_2WEAK in the CL and GP samples (Table 2) approached that for the ALC and its specificity was reasonably good.

Data shown in Table 2 are based on the criteria of $T_1 \geq 3$ drinks and $T_2 \geq 5$ drinks as the positive indications of tolerance. Because it is possible that CL and GP subjects, as well as males and females, might need different cut-offs for T_1 and T_2, we have experimented with different values for T_1 and T_2 without changing the total cut-off scores for T_1WEAK and T_2WEAK (Table 3). As seen in Table 3, increasing the T_1 cut-off stepwise from 2 to 6 only decreased slightly the sensitivity in CL males, but there was a gradual increase in specificity as the T_1 value increased. However, in GP males, the sensitivity was highest when $T_1 = 2$ and was lowest when $T_1 = 6$, and the specificity increased as T_1 value increased at the expense of a corresponding loss in sensitivity. There was no change in the 100% sensitivity in CL males as T_2 was increased from 5 to 8, with only a slight gain in specificity. Similarly, increasing T_2 from 5 to 8 only changed slightly the sensitivity in GP males, with very little change in specificity. Therefore, it appears that for males, the T_1(3)WEAK would be the test of choice for CL subjects if the objective were to maximize sensitivity, but T_1(6)WEAK would be the measure of choice if a high specificity were desired. The T_2(5 or 6)WEAK or T_1(2)WEAK would be good for screening GP subjects if a high sensitivity is desired. In contrast, the data in Table 3 indicate that for females, the T_1(2)WEAK would have the better sensitivity and specificity for both CL and GP subjects.

Table 2. Sensitivity and Specificity of the TWEAK, CAGE, and B-MAST

	Gold standard A*		Gold standard B†	
	Sensitivity/specificity (%)		Sensitivity/specificity (%)	
Test	CL	GP	CL	GP
T_1WEAK	94.4/95.5	88.7/82.2	85.4/83.0	86.1/72.4
T_2WEAK	94.4/89.4	98.6/74.5	93.8/80.4	92.1/64.6
CAGE	100/68.2	84.5/71.1	72.9/65.2	75.5/65.3
B-MAST	77.8/80.3	47.9/84.9	72.9/76.9	55.0/84.4

* DSM-III-R past year alcohol dependence (n = 84 and 369, for CL and GP, respectively).

† Heavy drinking as defined in text (n = 390 and 992, for CL and GP, respectively).

Table 3. Gender Differences in Sensitivity and Specificity of TWEAK

	Male		Female	
	Sensitivity/specificity† (%)		Sensitivity/specificity† (%)	
Tolerance cut-off*	CL	GP	CL	GP
T_1 (2)	100/75.0	95.5/69.9	100/96.0	96.3/81.7
T_1 (3)	100/81.3	88.6/74.8	83.3/100	88.9/87.4
T_1 (4)	91.7/87.5	70.5/83.7	83.3/100	81.5/91.4
T_1 (5)	91.7/87.5	63.6/88.6	83.3/100	81.5/94.9
T_1 (6)	91.7/93.8	61.4/90.6	83.3/100	81.5/95.4
T_2 (5)	100/68.8	100/67.5	83.3/96.0	96.3/79.4
T_2 (6)	100/81.3	100/67.5	83.3/100	96.3/81.7
T_2 (7)	100/81.3	97.7/71.5	83.3/100	88.9/84.0
T_2 (8)	100/81.3	95.5/72.4	83.3/100	88.9/84.6

* Number in parentheses is the cut-off value for Tolerance (T_1 or T_2). Total cut-off score (≥3) remained unchanged for T_1WEAK or T_2WEAK.

† The gold standard was DSM-III-R past-year alcohol dependence (CL, n = 28 males and 56 females; GP, n = 167 males and 202 females).

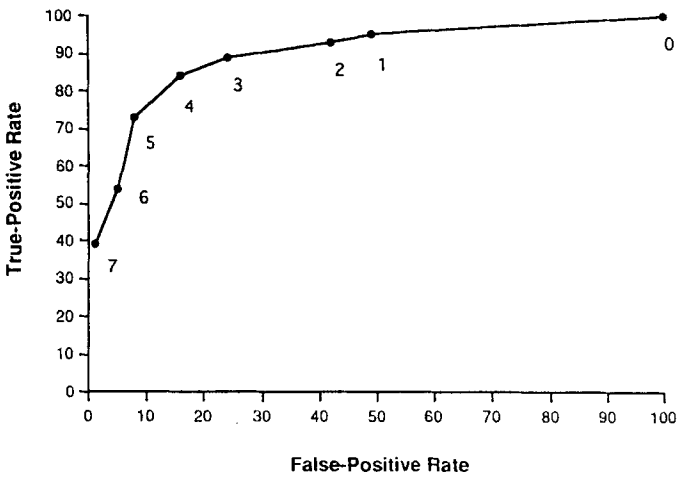

Fig. 1. ROC curve for T_1WEAK. Computations for sensitivity and specificity (varying total cut-off scores) were based on the following group definitions: ALC subjects plus all heavy drinkers in CL and GP as "true-positives" and all the nonheavy drinkers in CL and GP as the "true-negatives."

Figure 1 shows an ROC curve for T_1(3)WEAK. "True-positives" were defined as all the ALC subjects plus the heavy drinkers in CL and GP, and "true-negatives" consisted of all the nonheavy drinkers in CL and GP. The ROC curve plots the true-positive rate (sensitivity) against the false-positive rate (1 minus specificity),[30] using different total cut-off score for T_1WEAK. Steep parts of the ROC curve indicate large gains in sensitivity, with small loss of specificity, and horizontal parts indicate small gains in sensitivity for large loss of specificity. It is seen from Fig. 1 that a total score of 3 or 4 is an efficient cut-off for the TWEAK. A similar result was obtained (data not shown) when the "true-positives" included only the heavy drinkers in CL and GP but not the ALC subjects.

DISCUSSION

The clinical validity of the two versions of TWEAK has been confirmed in this study in that either test identified

nearly all of the known alcoholics. Both the T_1WEAK and T_2WEAK also have fairly high sensitivity and specificity in detecting alcohol dependence or heavy alcohol consumption in CL and GP subjects (Table 2). Phrasing of the TWEAK questions has been deliberately cast in the present tense, and in the "Worry" question, a past-year criterion has been built in. These are aimed to avoid the major disadvantage of the CAGE and B-MAST, namely, inability to distinguish between lifetime and current alcohol problems because of the phrasing, "Have you ever ..." in the questions. It should be emphasized that the TWEAK is only a screening instrument rather than a diagnostic instrument. For screening purposes, a high sensitivity of the test is desired, which the TWEAK can achieve. The less-than-ideal specificity will inevitably result in some false-positives that can be ruled out by supplementary tests, such as physical symptoms, biochemical tests, or a more in-depth diagnostic interview.

On paper, two items of the TWEAK, namely, "cut-down" and "eye-opener," (see "Method" for phrasing of these two TWEAK questions) appear to overlap with the two similarly worded items in the CAGE. The phrasing of the two CAGE questions are: "Have you ever felt you ought to cut down on your drinking?" and "Have you ever had a drink first thing in the morning to steady your nerves or get rid of a hangover or for an eye-opener?" However, results of correlation analysis and logistic regression analysis indicate that these similarly worded items are not redundant. Correlation analysis between the TWEAK and CAGE items in both the GP and CL samples showed no correlation coefficients >0.62. In other words, no one TWEAK item explains >37% of the variance in any CAGE item. A logistic regression analysis of the GP data in which the CAGE items were entered first followed by insertion of the TWEAK items showed that the *F*-to-enter values of the similarly worded (i.e., "cut-down" and "eye-opener") items were usually higher than those of the nonsimilarly worded items. These data demonstrate that the similarly worded items in the CAGE and TWEAK are not redundant. Perhaps the different phrasing of these similarly worded items contributes to the extra discriminatory power of the TWEAK items over the CAGE items.

The ROC curve shown in Fig. 1 indicates that a TWEAK score of 3 or 4 is an efficient cut-off. This is in agreement with the initial choice of a cut-off score of 3 by Russell et al.[24] in their study of a special clinical population, namely, pregnant women.

Based on the results, we recommend T_1(3)WEAK or T_1(6)WEAK for CL males, T_2(5 or 6)WEAK or T_1(2)WEAK for GP males, and T_1(2)WEAK for CL and GP females. Replications of these findings are necessary before our recommendations can be extended to, or modified for use in, other clinical populations or other selected populations such as college students, DWI offenders, emergency room patients, etc. Data in Table 3 also suggest that the TWEAK appears to have a higher specificity for women than for men. However, it should be noted that the number of subjects in the CL sample was small, and this finding needs to be replicated with a larger sample size. There is also a need to determine whether the TWEAK will be equally effective if it is administered by doctors themselves rather than by research personnel. Plans are underway to conduct some of these studies. Another challenge is to convince more doctors and health professionals to adopt the screening of alcohol problems in their patients as part of their clinical routine.

ACKNOWLEDGMENTS

We thank Janet Berg for her skillful typing. We gratefully acknowledge the cooperations of Dr. Charles Hershey, Dr. Barbara Majeroni, Dr. Douglas Moffat, Mitch Lopez, and the staff of the Erie County Medical Center Primary Health Care Clinic, Family Care Clinic, and Alcoholism Treatment Center for their help in the recruitment of subjects. The following individuals provided assistance in the recruitment and interviewing of research subjects: Jorge Antonetti, Kevin Dees, Juan Figueroa, Carol Marx, Colleen Marx, Florian Penetrante, Maria Penetrante, Sheila Pilc, Catherine Rebmann, Richard Topolski, and Robin Truesdale.

REFERENCES

1. Bucholz KK, Homan SM, Helzer JE: When do alcoholics first discuss drinking problems? J Stud Alcohol 53:582–589, 1992

2. Bradley KA: Screening and diagnosis of alcoholism in the primary care setting. West J Med 156:166–171, 1992

3. Moore RD, Bone LR, Geller G, Mamon JA, Stokes EJ, Levine DM: Prevalence, detection, and treatment of alcoholism in hospitalized patients. JAMA 261:403–407, 1989

4. Schuckit MA, Irwin M: Diagnosis of alcoholism. Med Clin North Am 72:1133–1153, 1988

5. Skinner HA, Holt S, Sheu WJ, Israel Y: Clinical versus laboratory detection of alcohol abuse: The alcohol clinical index. Br Med J 292:1703–1708, 1986

6. Babor TF, Kadden R: Screening for alcohol problems: Conceptual issues and practical considerations, in Chang NC, Chao HM (eds): Early Identification of Alcohol Abuse (NIAAA Research Monograph No. 17). Washington, D.C., U.S. Government Printing Office, 1985, pp 1–30

7. Goldberg HI, Mullen M, Ries RK, Psaty BM, Ruch BP: Alcohol counseling in a general medicine clinic. Med Care 29:JS49–JS56, 1991

8. Cyr MG, Wartman SA: The effectiveness of routine screening questions in the detection of alcoholism. JAMA 259:51–54, 1988

9. Moran MB, Naughton BJ, Hughes SL: Screening elderly veterans for alcoholism. J Gen Intern Med 5:361–364, 1990

10. Ewing JA: Detecting alcoholism. The CAGE questionnaire. JAMA 252:1905–1907, 1984

11. Smart R, Adlaf EM, Knoke D: Use of the CAGE scale in a population survey of drinking. J Stud Alcohol 52:593–596, 1991

12. Beresford TP, Blow FC, Hill E, Singer K, Lucey MR: Comparison of CAGE questionnaire and computer-assisted laboratory profiles in screening for covert alcoholism. Lancet 336:482–485, 1990

13. Buchsbaum DG, Buchanan RG, Centor RM, Schnoll SH, Lawton MJ: Screening for alcohol abuse using CAGE scores and likelihood ratios. Ann Intern Med 115:774–777, 1991

14. Bush B, Shaw S, Cleary P, Delbanco TL, Aronson MD: Screening for alcohol abuse using the CAGE questionnaire. Am J Med 82:231–235, 1987

15. Mayfield D, McLeod G, Hall P: The CAGE questionnaire: Validation of a new alcoholism screening instrument. Am J Psychiatry 131:1121–1123, 1974

16. Pileire B, Brendent-Bangou J, Valentino M: Comparison of

questionnaire and biochemical markers to detect alcohol abuse in a West Indian population. Alcohol Alcohol 26:353–359, 1991

17. Heck EJ, Lichtenberg JW: Validity of the CAGE in screening for problem drinking in college students. J Coll Stud 31:359–364, 1990

18. Mischke HD, Venneri RL: Reliability and validity of the MAST, Mortimer-Filkins Questionnaire and CAGE in DWI assessment. J Stud Alcohol 48:492–501, 1987

19. Barrett TG, Williams CHV: Use of a questionnaire to obtain an alcohol history from those attending an inner city accident and emergency department. Arch Emerg Med 6:34–40, 1989

20. Mangion DM, Platt JS, Syam V: Alcohol and acute medical admission of elderly people. Age Ageing 21:362–367, 1992

21. Rydon P, Redman S, Sanson-Fisher RW, Reid ALA: Detection of alcohol-related problems in general practice. J Stud Alcohol 53:197–202, 1992

22. Chan AWK, Pristach EA, Welte JW: Detection of alcohol problems and heavy drinking in three populations. Alcohol Clin Exp Res 15:375, 1991

23. Sokol RJ, Martier SS, Ager JW: The T-ACE Questions: Practical prenatal detection of risk-drinking. Am J Obstet Gynecol 160:863–870, 1989

24. Russell M, Martier SS, Sokol RJ, Jacobson S, Jacobson J, Bottoms S: Screening for pregnancy risk-drinking: Tweaking the tests. Alcohol Clin Exp Res 15:368, 1991

25. Pokorny AD, Miller BA, Kaplan HB: The brief MAST: A shortened version of the Michigan Alcoholism Screening Test. Am J Psychiatry 129:118–120, 1972

26. Chan AWK, Pristach EA, Welte JW, Hershey CO: Use of the CAGE in the detection of alcoholism or heavy drinking. J Subst Abuse (submitted)

27. Radloff LS: The CES-D scale: A self-report depression scale for research in the general population. Appl Psych Meas 1:385–501, 1977

28. Robins L, Helzer J, Cottler L, Goldring E: NIMH Diagnostic Interview Schedule (ed 3, revised), 1989, pp 13–14

29. Sackett DL: A primer on the precision and accuracy of the clinical examination. JAMA 267:2638–2644, 1992

30. Hsiao JK, Bartok JJ, Potter QZ: Diagnosing diagnoses: Receiver operating characteristic methods and psychiatry. Arch Gen Psychiatry 46:664–667, 1989