CLINICAL PRACTICE

Comparison of CAGE questionnaire and computer-assisted laboratory profiles in screening for covert alcoholism

THOMAS P. BERESFORD FREDERIC C. BLOW ELIZABETH HILL
KATHLEEN SINGER MICHAEL R. LUCEY

To identify the most effective method of screening for covert alcoholism Ewing's CAGE questionnaire was compared with several computer-assisted laboratory data profiles in a prospectively gathered, random sample of 915 adults admitted to a general hospital. Whether a subject was alcohol dependent (n = 244) or not (n = 671), as defined by DSM-III-R, was determined on the basis of a structured interview. The CAGE questionnaire was highly sensitive (76%) and specific (94%) for recognition of alcohol dependence (positive predictive power 87%). None of the discriminant laboratory functions gave recognition rates greater than chance alone. Until the sensitivities, specificities, and positive predictive powers of computer-assisted methods improve, brief interview alone remains the best screening method for general hospital populations.

Lancet 1990; **336**: 482–85.

Introduction

Alcoholic patients are often admitted to general hospitals with medical complaints,[1,2] but their alcohol dependence is seldom recognised, so the opportunity for treatment is missed.[3-5] Recognition and treatment of alcohol-dependent patients leads to much cost savings by lessening the use of hospital and emergency services by these patients.[6,7] The problem for clinicians has been one of recognition.

Screening for alcoholism has been based on a brief interview or on biochemical markers. Ewing[8] pioneered the first approach with the development of the four brief CAGE questions—Have you ever felt you should Cut down on your drinking? Have other people Annoyed you by criticising your drinking? Have you ever felt Guilty about drinking? Have you ever taken a drink in the morning to steady your nerves or get rid of a hangover? (Eye opener). The CAGE questionnaire has been found to be both valid and reliable for clinical use, with a high positive predictive power of 50–82%.[9-12] However, it has not been validated systematically against the DSM-III-R criteria for alcohol dependence.[13]

As to the second approach, statistical associations between specific clinical laboratory variables (such as the mean corpuscular volume or serum bilirubin) and alcohol-related disease have long been recognised. Although previous studies[9,14] have shown that single laboratory tests give low positive predictive values, there remains the possibility that the computer-assisted statistical construction of groups of tests might improve the predictive power enough for them to be clinically useful.

The effectiveness of any screening procedure rests on the statistical probability with which a case identified by screening can be confirmed clinically. A large, prospective study of randomly selected hospital patients was designed to assess this probability for putative laboratory data profiles. Several such profiles were constructed, and their value in screening for alcohol dependence as defined by DSM-III-R was compared with that of the CAGE questionnaire. For clarity of analysis, the diagnosis of alcohol abuse was not used; the study sample was dichotomised, on the basis of the study interview, as fulfilling or not fulfilling criteria for alcohol dependence.

Patients and methods

Subjects

1262 patients selected at random, from random number tables, from the University of Michigan Hospital daily admission census between Jan 1, 1987, and April 30, 1989, inclusive, were invited to take part in the study. The 915 who agreed to participate were aged 19 to 86 years (mean 48·8, SD 17); 52% were men; 89% were white, 10% black, and 2% of other races. Patients remained in hospital for a mean of 9·8 days and, on average, were interviewed 3 days after admission. They were drawn from a variety of medical and surgical general and specialty units. Patients from obstetric and gynaecological, psychiatric, and paediatric services were excluded, as were those aged under 18 years or those judged by the interviewer (research nurse) to be too ill to take part. The 915 participants represented 1·2% of the total hospital admissions (n = 76 737) over the duration of the study.

The 347 subjects declining participation in the study were interviewed briefly for demographic data and reasons for their refusal to enter the study, and, with their verbal permission, the interviewer examined the medical record of the current admission for information on alcohol (positive for 14%) or drug use (8%) and alcohol-related illnesses (3%). Over half of those refusing consent for study (53%) were 60 years or older, 11% were aged 18–29, 8·6% 30 to 39, 12% 40 to 49, and 16% 50 to 59. 56% of those refusing to enter the study were male. 3% had cirrhosis and 3% had hepatitis. None of these data differentiated participants from non-participants. Reasons for non-participation were: wanting to avoid venesection (21%); feeling too ill or too tired to complete the interview (24%); no alcohol use (6%); excessive interviewing in hospital (5%); immediate discharge (6%); no interest in the project (7%); information considered too personal (3%); suspicion of consent forms (2%); heavy alcohol use in the family (2%); no

ADDRESS. **Alcohol Research Center** (T. P. Beresford, MD, F. C. Blow, PhD, Kathleen Singer, RN, E. Hill, PhD), **and Department of Internal Medicine** (M. R. Lucey, FRCPI), **University of Michigan, USA**. Correspondence to Dr T. P Beresford, University of Michigan Alcohol Research Center, 400 East Eisenhower Parkway, Suite A, Ann Arbor, Michigan 48104, USA

explanation (6%); and other, such as never volunteering for anything, not wanting to discuss alcohol use, never participating in activities irrelevant to present clinical care, already participating in other research projects, and inability to discuss alcohol use because a family member had recently been killed by drunk driver (18%).

Methods

The study was approved by our institutional review board and informed consent was obtained from all participants. A small quantity of blood for serum chemistry and haematological assessment was obtained from all participants at the time of the interview.

The study interview, which lasted about an hour, was based on Vaillant's interview schedule for alcohol use[15] and included the CAGE questionnaire. From the resultant data we can construe a series of published scales measuring historical variables pertinent to alcohol use. In addition to the above, we recorded each subject's quantity and frequency of alcohol use, state of cognition, and demographic characteristics.

Patients were classified as alcohol dependent when they satisfied at least three of the nine DSM-III-R criteria plus at least one criterion representing each of the following three domains: impairment of control of drinking (criteria 1–3), impairment of social functioning because of drinking (criteria 4–5), and physical tolerance to alcohol or withdrawal symptoms on abstaining from alcohol (criteria 7–9). For comparability with previous studies,[9-12] a positive answer to two or more of the CAGE questions was taken as an indication of alcohol dependence.

All blood samples for the study were analysed by the same commercial clinical laboratory. To assess the effect of the interval between admission and phlebotomy on these data, laboratory findings on admission were obtained from our own hospital laboratory retrospectively for a subset of participants (n = 451).

A library of 38 possible variables were used for the discriminant analysis. Most values were used without mathematical transformation: serum calcium, phosphorus, uric acid, total protein, sodium bicarbonate, globulin, blood urea nitrogen/creatinine ratio, red cell volume, haemoglobin, packed cell volume, mean corpuscular volume (MCV), mean corpuscular haemoglobin (MCH), mean corpuscular haemoglobin concentration (MCHC), and cell counts of platelets, lymphocytes, monocytes, and eosinophils. Variables whose distributions were very abnormal were transformed into the square (serum albumin, chloride, and the neutrophil count) or the reciprocal (triglycerides, blood urea nitrogen, cholesterol, total bilirubin, direct bilirubin, alkaline phosphatase, lactate dehydrogenase, aspartate aminotransferase, alanine aminotransferase, creatinine, potassium, albumin/globulin ratio, indirect bilirubin, glucose, and the white blood cell and basophil counts). γ-glutamyl transpeptidase was not included because of its lability in relation to acute alcohol use. For the multivariate discriminant analysis, missing values were imputed by inserting the mean value for the subject's sex and age group; imputations were required for an average of 121 observations per laboratory variable.

TABLE I—PERCENT OF SUBJECTS RESPONDING POSITIVELY ON SPECIFIC DSM-III-R DIAGNOSTIC CRITERIA FOR ALCOHOL DEPENDENCE

Criteria	Dependent* (n = 244)	Non-dependent (n = 671)
Impaired control		
1. Regular use of amounts greater than intended	19·3	3·0
2. Attempts to cut down/control use	96·7	25·5
3. Time and effort to get/take/recover from alcohol	34·4	3·7
Social/physical decline		
4. Use despite social obligations or hazards	57·9	6·3
5. Use replaces activities	58·2	7·3
6. Use despite recurrent problems	93·4	30·3
Tolerance/withdrawal		
7. Tolerance	13·1	2·7
8. Withdrawal symptoms	93·4	23·0
9. Use to treat withdrawal	43·4	1·3

*Dependent/non-dependent frequency comparisons differ throughout, Yates-corrected χ^2 test, $p < 0\ 0001$

Statistical analysis

Univariate analyses of covariance (two-way analysis of variance by alcohol diagnosis and gender, with age as a covariate) were used for differences between those who met criteria for alcoholism and those who did not.

Multivariate analyses were used to determine whether the computer could derive laboratory data profiles that characterised the two clinical patient groups and to assess the effectiveness of the various profiles as screening strategies. The first stage of the discriminant analysis was a test of whether the equation generated in our pilot study[9] could be validated in this larger, more heterogeneous sample. Because the laboratory profile derived from our pilot sample did not discriminate well in the present sample, we wished to ensure that time of collection did not act as a confounding variable. To account for the differences in prediction studies, day-of-admission laboratory data were examined to rule out time as inpatient as a confounding variable. We ran a repeated measures analysis of variance to test for possible pattern differences between admission and commercial laboratory findings for serum creatinine, aspartate aminotransferase, uric acid, lactate dehydrogenase, and the mean corpuscular volume. Although there were consistent differences between the laboratories, there were no interactions by diagnosis ($p > 0·05$)—that is, there were no differences at admission that dissipated three days later.

In the second stage only the variables found to be important in our pilot study (blood urea nitrogen, creatinine, uric acid, mean corpuscular volume, total bilirubin, aspartate aminotransferase, and lactate dehydrogenase) were used for the discriminant analysis. We next created a new function from all available laboratory tests. The DSM-III-R diagnostic groups of probable alcoholics and non-alcoholics were divided randomly into halves by use of the random number generating function in the SAS program. One half of the sample was used to create the discriminant function and became the criterion sample; the other random half was used to test the efficacy of the created function and is referred to as the validation sample.

Stepwise discriminant analysis was used to reduce the number of laboratory tests included in subsequent analyses. Significance levels were 0·10 for entering and 0·15 for leaving the equation. Prior probabilities were set at 0·5. We entered these reduced sets of variables into a discriminant analysis to test the accuracy of discrimination. This process was repeated with age, gender, and CAGE score in the variable set used for stepwise regression.

Each analysis was done twice, with both linear and quadratic methods. The equality of within-group covariances for the alcoholic group and the non-alcoholic group was tested by use of a likelihood ratio.[16] In this sample, the linear method is subject to bias because the within-group covariance matrices are not equal and should not be pooled in two cases: the function using the variables derived from previous work ($\chi^2 = 199·735$, $p < 0·001$) and the present analysis including CAGE score ($\chi^2 = 80·206$, $p < 0·001$). The quadratic form is appropriate for the discriminant functions created from all available markers ($\chi^2 = 17·856$, $p = 0·177$) and that with the addition of age and sex ($\chi^2 = 43·258$, $p = 0·189$). Nonetheless, we included results from both methods to test their predictive accuracy; discriminant analysis is considered fairly robust to violation of variance assumptions.[17,18]

Results

Univariate differences

On the basis of the DSM-III-R drinking history variables 244 subjects were classified as alcohol dependent (27% of the total number) (table I). There were significant differences in laboratory findings between alcohol dependent and non-dependent subjects (table II). Results were collapsed over three age groups (19–39, 40–59, 60 and

over) since an interaction of age and diagnosis occurred only for alanine aminotransferase ($p = 0.034$). This occurred because alanine aminotransferase was higher for alcoholics only among the youngest age-group. Significant differences occurred for serum uric acid and sodium concentrations, and for MCV and MCH (experiment wise $p < 0.05$ using the Bonferroni correction for multiple tests; nominal $p < 0.006$). Suggestive differences occurred for total bilirubin, serum lactate dehydrogenase, creatinine, potassium, chloride, blood urea nitrogen/creatinine ratio, haemoglobin, haematocrit, and the monocyte count. In view of the large number of comparisons, these differences (nominal $p < 0.05$) should be considered suggestive only.

Multivariate analyses

The equation generated in our pilot study[9] did not discriminate well; it identified only about 4% of the entire sample as being positive for alcoholism (table III).

Discrimination by use of variables previously found useful

In the pilot sample, the variables MCV, blood urea nitrogen, creatinine, total bilirubin, aspartate aminotransferase, uric acid, and lactate dehydrogenase discriminated between alcoholics and non-alcoholics. When these variables were applied to a randomly split half sample for quadratic discriminant analysis, only 21% of the alcohol-dependent subjects were so classified, whereas 86% of the non-dependent subjects were correctly classified. For the linear discriminant analysis, only 62% of alcoholics and 59% of the non-alcoholics were correctly classified.

TABLE II—UNIVARIATE DIFFERENCES IN LABORATORY VALUES BETWEEN DSM-III-R ALCOHOL DEPENDENT AND NON-DEPENDENT SUBJECTS

Variable	Dependent (n = 244)	Non-dependent (n = 671)	Difference (p)
Serum			
Uric acid (mg/dl)	5·5	5·2	0·0016
LDH (U/l)	230·0	202·0	0·089
Creatinine (mg/dl)	1·5	1·3	0·04
Sodium (mmol/l)	139·0	140·0	0·014
Potassium (mmol/l)	4·43	4·35	0·03
Chloride (mmol/l)	104·0	105·0	0·034
Blood			
MCV (fl)	92·8	91 6	0·001
MCH (pg)	31·0	30·5	0·0006
Hb (g/dl)	12·4	12·1	0·012
PCV (%)	37·3	36·3	0·017
Monocytes ($10^3/\mu l$)	6·3	6·0	0·023
Platelets ($10^3/\mu l$)	289·5	299 7	0 077

LDH = lactic dehydrogenase; MCV = mean corpuscular volume; MCH = mean corpuscular haemoglobin; Hb = haemoglobin, PCV = packed cell volume

TABLE III—COMPARISON BETWEEN BRIEF HISTORY AND LABORATORY TESTS IN PREDICTION OF DSM-III-R ALCOHOL DEPENDENCE

—	Sensitivity	Specificity	Predictive power
CAGE questionnaire	76%	94%	87%
Linear function, pilot study	5%	96%	28%
Laboratory variables, pilot study			
Linear	62%	59%	50%
Quadratic	21%	86%	33%
Exploratory function, present study			
Linear	57%	57%	31%
Quadratic	55%	55%	29%
Exploratory function plus age and sex			
Linear	58%	66%	36%
Quadratic	56%	69%	36%
Exploratory function plus age, sex, and CAGE			
Linear	82%	91%	76%
Quadratic	80%	90%	72%

Discrimination by use of all available variables

When all available laboratory variables were entered into a stepwise discriminant analysis, the function incorporated only haemoglobin, sodium, blood urea nitrogen/creatinine ratio, uric acid, and MCH. The first variable entered was haemoglobin, the second was sodium. All had shown univariate differences between the two groups. Variables that showed univariate differences but did not enter into the discriminant were MCV and to some extent lactate dehydrogenase, creatinine, potassium, chloride, haemoglobin, packed cell volume, and the monocyte count. Neither the linear nor quadratic discriminant functions were better than chance (table III).

Incorporation of age, gender, and CAGE scores

When age and gender were allowed to compete for entry into the stepwise discriminant function, they were the first variables to be entered. The stepwise discriminant function then incorporated sodium, MCH, total protein, alanine aminotransferase, alkaline phosphatase, and bicarbonate. Sodium and MCH had shown univariate differences between the two groups, but total protein, alanine aminotransferase, alkaline phosphatase, and bicarbonate had not. Alanine aminotransferase had shown an interaction between age and diagnosis, however. The linear and quadratic discriminant functions performed slightly better in classifying non-alcoholics correctly.

When CAGE score was also entered, it was the first predictor; next came sodium (removed in a further step), gender, age, alkaline phosphatase, uric acid, and MCH. As mentioned above, alkaline phosphatase had not shown any univariate differences. For the linear and quadratic discriminant analyses, over 80% of the alcoholics and almost 90% of the non-alcoholics were correctly classified. Addition of laboratory values to the CAGE score did not, however, significantly increase its power to predict a diagnosis of alcoholism in this sample. It significantly decreases the positive predictive power of the CAGE questions by virtue of including a greater number of screening variables.

Discussion

Unlike previous studies,[9-12] the present work validates the CAGE screening questionnaire against standard diagnostic criteria in a large number of randomly selected patients admitted to general hospital wards. In our pilot study[9] we examined patients admitted consecutively, rather than those chosen at random, and the sample size was one-tenth that reported here. In Bush and co-workers' sizeable study[12] of non-randomly selected patients the criteria for diagnosis were questionable—the criteria were a measure of quantity and frequency of alcohol use, the presence of an "alcohol-related diagnosis", or the subject's response to the Michigan Alcohol Screening Test (MAST), a screening test rather than a diagnostic instrument. We trust that our statistical and methodological safeguards are sufficient for us to conclude that the CAGE questionnaire is a powerful screening tool for clinical use.

Some will argue that the CAGE questions overlap significantly with the more extensive DSM-III-R criteria,

so that the comparison between CAGE screening and DSM-III-R diagnosis remains unclear. In our view, the overlap of four easily used screening questions with the more extensive diagnostic criteria indicates the precise strength of the CAGE screening examination. Until better diagnostic schemes emerge, possibly incorporating physical examination or clinical laboratory data, brief screening examinations based on the patient's history remain the best method for early recognition of alcohol dependence.

Despite promising early studies, the data presented here show no real advantage of using statistical manipulations of clinical laboratory variables to assist physicians in recognising the presence of covert alcoholism. Although mean differences for groups of alcohol dependent versus non-dependent patients reached significance, individual differences were not sufficiently robust as to allow for provisional diagnosis in individual patients through either linear or quadratic discriminant procedures. This is contrary to the results of earlier studies done with smaller, non-randomly sampled, and homogeneous subject pools. It is clear now that a brief clinical interview is far more efficient than computer-generated profiles in identifying alcoholic subjects.

Although one strength of our study is its size and the wide range of medical conditions represented, there is the possibility that smaller subpopulations (for example, patients admitted for gastrointestinal complaints) may show sufficient homogeneity in their clinical laboratory variables so as to make computer-assisted provisional diagnosis possible in that subgroup. We are exploring this possibility.

Approximately one-quarter of those approached refused to take part in the study. Refusal to enter a study related to alcoholism is a complex issue.[19] Inconvenience was the principal reason for refusal cited by out subjects. Although the gender distribution was similar among participants and non-participants, there was a striking preponderance of persons aged 60 years and over among the non-participant group. Interestingly, alcoholism was more prevalent among the younger participants than among those aged 60 years and over. It is possible that we have underestimated the true prevalence of alcoholism in the older age group. The case records revealed some history of alcohol abuse in 14% of the non-participants. Obtaining information from case-records is a less stringent means of diagnosing alcoholism than the detailed structured interview. We believe, therefore, that our conclusions about screening for covert alcoholism should be tempered with the caveat that the value of the CAGE questionnaire is not so well established among older patients who are reticent about discussing alcohol use.

Our study is cross-sectional in nature, with little to say about either the validity of the CAGE examination over time or the possibility that serial clinical laboratory mesures may be useful in screening for alcohol dependent patients. Nonetheless, the data substantiate the effectiveness of the CAGE screening examination when measured against standard diagnostic yardstick. If the questionnaire were widely used in routine clinical practice there could be substantial gains in the morbidity, mortality, and health care costs associated with alcoholism. It must be remembered that the CAGE examination is only a screening examination and cannot replace more detailed patient assessment and diagnosis. Its high positive predictive power is directly related to the high base rate of alcohol dependence among hospital populations, so it may not be as effective in other settings. We recommend that house officers and those who train them incorporate the CAGE questions into their admission evaluation for all adult patients admitted to hospital.

We thank Dr Kirk Brower for his help in preparing this report.

This work was supported by grants R01-AA-07236 and P50-AA-07378 from the National Institute on Alcohol Abuse and Alcoholism.

REFERENCES

1. Beresford TP. Alcoholism consultation and general hospital psychiatry. *Gen Hosp Psychiatry* 1979; **1:** 233–300.
2. Moore RD, Bone LR, Geller G, et al. Prevalence, detection and treatment of alcoholism in hospitalized patients. *JAMA* 1989; **261:** 403–07.
3. Beresford TP, Low D, Adduci R, Goggans F. Alcoholism assessment on an orthopaedic surgery service. *J Bone Jt Surg* 1982; **64A:** 730–33.
4. Geller G, Levine DM, Mamon J, et al. Knowledge, attitudes, and reported practices of medical students and house staff regarding the diagnosis and treatment of alcoholism. *JAMA* 1989; **261:** 3115–20.
5. West LJ, Maxwell DS, Noble EP, Solomon DH. Alcoholism. *Ann Intern Med* 1984; **100:** 405–16.
6. Jones KR, Vischi TR. Impact of alcohol, drug abuse and mental health on treatment of medical care utilization. *Med Care* 1979; **17** (suppl): 1–81.
7. Reiff S, Griffith B, Forsythe AB, Sherman RM. Utilization of medical services by alcoholics participating in health maintenance organization outpatient treatment program: three year follow-up. *Alcohol Clin Exp Res* 1981; **5:** 559–62.
8. Ewing JA. Detecting alcoholism, the CAGE questionnaire. *JAMA* 1984; **252:** 1905–07.
9. Beresford TP, Low D, Hall RCW, Adduci R, Goggans F. A computerized diagnostic biochemical profile for the detection of alcoholism. *Psychosomatics* 1982; **23:** 713–20.
10. Beresford TP. Screening for alcohol abuse using the CAGE questionnaire. *Am J Med* 1987; **84:** 805.
11. Bernadt MW, Mumford J, Taylor C, Smith B. Comparison of questionnaire and laboratory tests in the detection of excessive drinking and alcoholism. *Lancet* 1982; i: 325–28.
12. Bush B, Shaw S, Cleary P, Delbanco TL, Aronson MD. Screening for alcohol abuse using the CAGE questionnaire. *Am J Med* 1987; **82:** 231–25.
13. American Psychiatric Association: Diagnostic and Statistical Manual, 3rd ed, revised. Washington, DC: American Psychiatric Press, 1987.
14. Eckhardt MJ, Ryback RS, Rawlings RR, Graubard BI. Biochemical diagnosis of alcoholism. *JAMA* 1980; **246:** 2707–10.
15. Valliant GE. The natural history of alcoholism. Cambridge, Massachusetts: Harvard University Press, 1983: 329–30.
16. Freeland KE, Frankel MT, Evenson RC. Biochemical diagnosis of alcoholism in men psychiatric patients. *J Stud Alcohol* 1985; **46:** 103–06.
17. Anderson TW. An introduction to multivariate statistical analysis. New York: Wiley, 1958.
18. SAS user's guide: statistics version 5. Cory, North Carolina: SAS Institute Inc., 1985.
19. Beresford TP, Low D, Hall RCW. Patients who refuse study: a bias factor in alcoholism prevalence research. *Curr Controv Alcohol* 1983; **2:** 17–22.

From the Lancet

Fitness for service

Boys mostly leave the elementary schools in good condition. Physical training under the Board of Education is an everyday part of their lives, well organised and well taught. After the age of fourteen they have to fend for themselves and get what exercise they can. Many of them, lacking the stimulus and competition of companions, become physically lazy, and as Lord Dawson has pointed out . . . "the promise of childhood too often fades away into weediness and futility." It seems deplorable that we need a war to drive this lesson home—that we, who pride ourselves on human dealing, should never have ensured for our younger citizens the chance to achieve full growth. If a system of widespread physical training for young people can be arranged now perhaps it is not too much to hope that it may be maintained when peace comes. Fitness for service is an ideal worth keeping before us in peace as well as war.

(July 20, 1940)