Screening properties of questionnaires and laboratory tests for the detection of alcohol abuse or dependence in a general practice population

B Aertgeerts, F Buntinx, S Ansoms and J Fevery

SUMMARY

Background: *Early identification of alcohol abuse or dependence is important in general practice because many diseases are influenced by alcohol. General practitioners, however, fail to recognise most patients with alcohol problems.*

Aim: *To assess the diagnostic performance of the CAGE and AUDIT questionnaires, their derivatives, and laboratory tests in screening for alcohol abuse or dependence in a primary care population (male and female patients), attending their general practitioner (GP).*

Design of study: *A diagnostic cross-sectional study.*

Setting: *A random sample of patients who were over 18 years of age (*n *= 1992) attending 69 general practices situated in the same region in Belgium.*

Method: *Alcohol questionnaires (CIDI 1.1, section I, CAGE, AUDIT, AUDIT-C, Five-Shot, and AUDIT Piccinelli) were completed, demographic information was recorded, and patients underwent conventional blood tests, including mean corpuscular volume, liver function tests, the gamma-glutamyl transferase test, and carbohydrate-deficient transferrin (CDT, estimated using %CDT). Calculations of sensitivity, specificity, positive predictive value, negative predictive value, odds ratios with their 95% CIs, and receiver operating characteristic (ROC) curves for different scores of the questionnaires and laboratory tests, using DSM-III-R as the reference standard.*

Results: *The past-year prevalence of alcohol abuse or dependence in this population was 8.9% (178/1992) of which there were 132 male and 45 female patients attending a general practice. The GPs identified 33.5% of patients with alcohol abuse or dependence. Among male patients, all questionnaires had reasonable sensitivities between 68% and 93% and hence at lower cut-points than recommended. Only the sensitivity of the CAGE, even at its lowest cut-point of 1 was lower (62%). In female patients the sensitivities were lower; however, odds ratios were higher for different questionnaires. The receiver operating characteristic (ROC) curves did not differ between the questionnaires. The laboratory tests had low diagnostic accuracy with areas under the ROC curves (AUCs) between 0.60 and 0.67 for female patients and 0.57 and 0.65 for male patients.*

Conclusions: *This is one of the largest known studies on alcohol abuse or dependence among family care practices. We confirm earlier results that the AUDIT questionnaire seems equally appropriate for males and females; however, screening properties among male patients are higher. Nevertheless, the Five-Shot questionnaire is shorter and easier to use in a general practice setting and has nearly the same diagnostic properties in male and female general practice patient populations. We confirm that conventional laboratory tests are of no use for detecting alcohol abuse or dependence in a primary care setting. Also, the %CDT cannot been used as a screening instrument in this general practice population.*

Keywords: *cross-sectional study, alcohol abuse; alcohol dependence; screening questonnaire.*

B Aertgeerts, MD, PhD, associate professor; F Buntinx, MD, PhD, associate professor, Department of General Practice, Clinical Epidemiology Unit; S Ansoms, MD, psychiatrist; and J Fevery, MD, PhD, professor, Department of Liver Diseases, University Hospital Gasthuisberg, Catholic University of Leuven, Belgium.

Address for correspondence

Dr Bert Aertgeerts, Department of General Practice Katholieke Universiteit Leuven, Kapucijnenvoer 33 - Blok J, 3000 Leuven, Belgium. E-mail: bert.aertgeerts@med.kuleuven.ac.be

Submitted: 21 December 1999, Editor's response: 23 May 2000, final acceptance: 15 November 2000.

Introduction

THE detection of alcohol problems in a general practice population remains an important issue as, in general, physicians fail to recognise most patients with alcohol problems. In previous studies, only 6–36% of the patients with alcohol abuse or dependence were correctly identified.[1-6] Nevertheless, the early identification of alcohol-related problems is important because these problems are prevalent and pose serious health risks to patients and their families.[7-8] Additionally, recent research shows that a brief intervention by the physician may be very helpful in decreasing alcohol-related problems or amounts of drinking, especially in an early phase of the disease.[9-12] Primary care has been recognised as a potentially effective setting for such brief interventions that offer a mixture of advice, information, and health promotional literature to the targeted patient.[13]

While several studies tried to improve the assessment of alcohol abuse or dependence using educational interventions intended for use by general practitioners (GPs),[14,15] it is important that the GP has access to a screening instrument that is not only powerful but also practical and short, with a high sensitivity and an acceptable specificity.

Most laboratory tests or test combinations (mean cell volume [MCV], aspartate aminotransferase [AST], alanine aminotransferase [ALT], gamma-glutamyltransferase [GGT], and uric acid) seem to be inappropriate for screening alcohol abuse or dependence.[16] A recent study by Meerkerk and colleagues concerning the diagnostic accuracy of carbohydrate-deficient transferrin (CDT) in a general practice population concluded that CDT seems to be the best alcohol marker available although the difference between CDT and MCV is small.[17] To confirm the diagnosis of alcohol abuse or dependence the %CDT (an enzyme immunoassay that measured the relative amount of CDT in proportion to total transferrin), the successor of the CDT, also showed interesting results in populations with a high — or even a low — prevalence of alcohol problems.[18,19]

Compared with laboratory tests, most questionnaires produce better results in screening situations, although physicians have difficulty with their use during a routine consultation,[2,20] especially the CAGE questionnaire, developed by Ewing, and the Alcohol Use Disorder Identification Test (AUDIT).[21,22] The CAGE's brevity and non-intimidating approach make it a useful screening and case-finding tool, particularly for the busy primary care physician.[23-25] Most studies show that implementation of the CAGE may improve

HOW THIS FITS IN

What do we know?
According to several studies, GPs can only identify one-third of their patients with alcohol abuse of dependence. The literature demonstrates that in many cases a short intervention by the GP about alcohol problems can be sufficient to reduce the patient's consumption of alcohol. Hence early identification of alcohol problems is important.

What does this paper add?
Questionnaires are much better for screening of alcohol problems and provide the opportunity to talk about alcohol consumption with patients. This study investigates the diagnostic properties of several questionnaires and laboratory markers in general practice. Laboratory tests are of no use for screening for alcohol abuse or dependence.

the identification of alcohol problems in different settings.[24-26] Screening with the AUDIT is used not only for case finding but also for the detection of risky behaviour, such as hazardous levels of alcohol consumption.[27] The AUDIT may be used orally, in writing or via computer and may be included as part of a general health risk assessment. However, the 10-item AUDIT is not as easy to administer as the CAGE and is consequently less acceptable in clinical practice.[28] Therefore, several studies have been published examining the screening properties of different short versions of the AUDIT, or of a combination of some CAGE and AUDIT questions. Piccinelli *et al* recommended the use of five out of the 10 AUDIT items using the original AUDIT scoring procedure.[29] Seppa *et al* developed a questionnaire that combines two questions from the AUDIT and three from the CAGE with a different scoring system.[30] Lastly, Bush *et al* described the usefulness of the first three AUDIT questions only.[31]

Much research has been published describing the diagnostic characteristics of single instruments regarding alcohol abuse or dependence in a general practice population. Nevertheless, there is still need to compare several screening instruments in the same population. The diagnostic accuracy, in male as well as in female patients, requires further investigation.

The aim of this study was therefore to assess and compare the diagnostic accuracy of the CAGE, the AUDIT, and their derivatives in a general practice population according to DSM-III-R criteria of alcohol abuse or dependence. In addition, we compared the diagnostic accuracy of conventional laboratory tests (GGT, ASAT, ALAT, uric acid, MCV) and the %CDT in the same population and with the same criterion.

Method

Data collection

Patients aged over 18 years and attending their GP during a three-week period (Monday to Friday) were either consecutively or randomly included. Ninety-six GPs working in the same region invited 2262 patients to take part in the study, of which 189 refused. Most patients refused because they did not want a blood sample taken (n = 92), they were too ill (n = 35), had no time (n = 27), or for other reasons (n = 35). The mean age and sex of the non-participants did not differ from the study group. Complete data from the GP, the patient, and the laboratory were available from 1992 patients. Data were gathered only for those patients who agreed to participate and signed an informed consent form. The Ethical Committee of the University's Medical School of Leuven approved this study.

Instruments

The GP recorded the patient's answers to the CAGE questionnaire and his opinion of whether the patient had an alcohol problem. A blood sample was then taken from each patient. After the encounter with the doctor, the patient completed an auto-questionnaire that included demographic information, the Composite International Diagnostic Interview (CIDI 1.1, section I), the CAGE, and the AUDIT. Demographic information included age, sex, marital status, and employment status. The Composite International Diagnostic Interview is a standardised diagnostic interview for assessing mental disorders according to the criteria of the Diagnostic and Statistical Manual of Mental Disorders, third revised version (DSM-III-R).[32] The CIDI was used as an auto-questionnaire derived from the 'paper and pencil' version 1.1 from the CIDI. Although this version is not strictly validated it follows the same structure as the interviewer using the paper and pencil version. DSM-III-R criteria were operationalised using the CIDI 1.1 since, at the start of the data collection, the CIDI 2.1 (generating DSM-IV criteria) was not available in the Dutch language. The criteria for alcohol dependence for DSM-III-R are, however, very similar to DSM-IV, although DSM-IV reveals more people with alcohol abuse.[33] Patients are classified as alcohol dependent, alcohol abuser or normal, according to their results on the CIDI.[34,35] We used the DSM-III-R criteria generated by the CIDI as the reference standard.

The CAGE questionnaire is a brief screening instrument containing only four short questions: 'Have you ever felt you should cut down on your drinking?'; 'Have people annoyed you by criticising your drinking?'; 'Have you ever felt bad or guilty about your drinking?'; and 'Have you ever had a drink in the morning to get rid of a hangover?'. Each item can have either a 'yes' or 'no' response. The CAGE questions were developed from a clinical study performed in 1968 by Ewing at the North Carolina Memorial Hospital.[21] It was validated by Mayfield and his colleagues in 1970 in the setting of a psychiatric service.[36]

The AUDIT questionnaire is self-administered and includes three items on the amount and frequency of drinking, three items on alcohol dependence, and four items on common problems caused by alcohol. Each item is scored from 0 to 4, resulting in a total score of 0-40.[22] The AUDIT-PC questionnaire, a short version of AUDIT, was constructed by Piccinelli *et al*. It consisted of five AUDIT items only (items 1, 2, 4, 5, and 10) using the same scoring system.[29]

The Five-Shot questionnaire was constructed by Seppa *et al* using three CAGE items and two AUDIT items and with a different scoring system.[28] The AUDIT-C, a questionnaire constructed by Bush *et al*, includes the first three items of

the AUDIT.[31] The derivatives of the AUDIT are available on request from the authors.

For the conventional laboratory tests (GGT, MCV, ASAT, ALAT, and uric acid), serum samples were examined daily at the clinical haematology laboratory of the University Hospital Louvain using standard laboratory routines. Cut-off levels were: GGT ≥50 units/l (males) and ≥32 units/l (females); MCV ≥96 FL; ASAT ≥37 units/l; ALAT ≥40 units/l; uric acid ≥7.5 mg/dl (males) and ≥6.0 mg/dl (females). Additionally, %CDT, a relatively new marker for detecting heavy alcohol use, was tested on frozen blood samples from all patients suffering from alcohol abuse or dependence based on the DSM criteria and on a random sample of negative patients. The %CDT measures the relative amount of CDT in proportion to total transferrin. For %CDT measurements a commercial kit was used as described by the manufacturer (AXIS %CDT from AXIS Biochemicals ASA, Oslo, Norway).

Scoring of all questionnaires, including the CIDI, was fully computerised. The GP (both in giving his opinion and afterwards, in recording the answers to the CAGE questions), the research associate who entered the responses to the questionnaires on the computer, the laboratory group that performed the routine laboratory tests, and the team performing the %CDT tests were all blinded to the results of the other tests.

Statistical analysis

The diagnostic value of the CAGE, the AUDIT, and their derivatives was investigated by the calculation of sensitivity, specificity, and positive predictive values (PPVs) and negative predictive values (NPVs), and also diagnostic odds ratios with 95% confidence intervals (95% CIs). Calculations were performed using Epi-Info software.[37] In addition, receiver operating characteristic (ROC) curves were processed on the basis of all possible values of the continuous and ordered test results using MedCalc5 software (IBM version 5) based on calculations as described by Hanley and McNeil.[38] Comparison of the AUCs derived from the same population was also performed using MedCalc5 software, based on calculations for comparison of paired curves as described by Hanley and McNeil.[39] To estimate the diagnostic value and the 95% CIs of the %CDT that was performed on a random sample of patients showing no alcohol abuse or dependence, results for these patients were weighted according to the inverse of the sampling proportion. The diagnostic measurements and their 95% CIs were calculated, taking into account the design of the study. This resulted in adjusting the standard formulae of the diagnostic measurements and the CIs by deriving the probabilities and their variances of the different cells appropriately using standard probability theory.[40]

Results

Prevalence of alcohol abuse or dependence

Table 1 summarises the demographic characteristics of our study population. Complete information was available for 1992 patients. The mean age of male and female patients was 54 years and 48 years respectively. Most patients were married (males 72.5%, females 66.2%) and the educational level was somewhat higher in female patients compared with male patients. About 20% of the female patients were housewives.

Prevalence rates, according to DSM-III-R criteria as a reference standard, are given in Table 2. The total prevalence for current alcohol abuse or dependence during the past year (i.e. during the past 12 months) in this sample was 8.9% (*n* = 178/1992). This group included 132 male and 46 female patients (sex ratio = 1 female:2.8 male). During the past year 74 (7.6%) male patients met the criteria of alcohol abuse and 59 (6.1%) met the criteria of alcohol dependence. Twenty (2%) female patients were alcohol abusers and 26 (2.6%) were alcohol dependent during the past year. Mean age did not differ between patients with and without alcohol abuse or dependence. In the total population, lifetime prevalence was 14.9%. Lifetime alcohol abuse or dependence was most frequent among male patients aged 18 to 60 years (*n* = 122/623, 18%). Binge drinking occurred weekly or daily in 114 male patients (12.1%) and in 12 female patients (1.2%). Based on the information of previous patient–doctor encounters, the GPs correctly identified 33.5% of their patients with alcohol abuse or dependence during the past year. Thirty-seven per cent of the male patients and 24% of the female patients were detected with alcohol problems.

*Screening test properties in male patients (*n = *971)*

Table 3 presents the screening performance for alcohol abuse or dependence of the CAGE, the AUDIT, and the adapted versions of the AUDIT at different cut-off values. Moreover, Table 3 shows the diagnostic value for alcohol abuse or dependence of MCV, GGT, ALT, AST, uric acid and %CDT.

The sensitivity of the CAGE at a cut-point ≥1 is 62% with a specificity of 81%. Although there is a high NPV of 93%, the PPV is low (34%). At a cut-point of ≥5 for the AUDIT and AUDIT-C there is a sensitivity of 82% and 78% respectively, with a specificity of 73% and 75% respectively. It is notable that, at the recommended cut-off value of ≥8 for the AUDIT, screening properties are too low in this male population attending their GP. At its recommended cut-off point the Five-Shot questionnaire has a sensitivity of 74% and a specificity of 81%. The AUDIT-PC has a lower sensitivity (68%), but a higher specificity (84%). At the cut-off value ≥5, the PPVs were very low for the AUDIT (32%) and the AUDIT-C (32.8%) and somewhat higher for the AUDIT-PC (40%, cut-off ≥5) and the Five-Shot questionnaire (38%, cut-off ≥2.5). On the other hand, the NPVs of all these screening tests are above 90% at different cut-off values.

The likelihood ratios for a positive and negative test result are illustrated in Table 3. For a positive test result the likelihood ratio is >3 for all mentioned questionnaires. The odds ratios for all these questionnaires — except for the CAGE with a cut-off level of ≥1 — were around 10 and even higher.

The screening properties of the MCV within a male primary care population and the other conventional laboratory tests were very low. The %CDT had a very low sensitivity (18%) at a recommended cut-point of ≥6. As shown in Table

Table 1. Demographic variables.

	Male *n* = 971 (48.8%)	Female *n* = 1021 (51.2%)	All *n* = 1992 (100%)
Mean age (percentile 25/75)	54 (66/39)	48 (63/36)	51 (65/37)
Number aged over 60 years	348 (35.6)	287 (28.1)	635 (31.9)
Marital status			
Married	705 (72.5)	675 (66.2)	1380 (69.4)
Live together	74 (7.6)	79 (7.8)	153 (7.7)
Single	91 (9.4)	95 (9.3)	186 (9.4)
Divorced	83 (3.4)	55 (5.4)	88 (4.4)
Widowed	41 (4.2)	87 (8.5)	128 (6.4)
Others	26 (2.7)	27 (2.6)	63 (2.7)
Missing	0 (0)	1 (0.1)	1 (0.1)
Education			
Secondary education	519 (53.5)	473 (46.3)	992 (49.8)
Tertiary education	445 (45.8)	532 (52.3)	977 (49.2)
Missing	6 (0.7)	14 (1.4)	20 (1.0)
Work			
Employed	428 (44)	387 (38)	815 (41)
Unemployed	87 (9)	107 (10.6)	194 (9.8)
Retired	364 (37.5)	209 (20.5)	573 (28.8)
Home-maker	12 (1.2)	225 (22.1)	237 (11.8)
Student	20 (2.1)	27 (2.6)	47 (2.4)
Others	55 (5.7)	58 (5.8)	114 (5.7)
Missing	4 (0.5)	5 (0.4)	9 (0.5)

Table 2. Prevalence of alcohol abuse or dependence among primary care patients.

	Male *n* = 971 (48.8%)	Female *n* = 1021 (51.2%)	All *n* = 1992 (100%)
Abstained in past year	155 (16.6)	263 (25.8)	418 (22.1)
DSM-III-R positive			
Past year	132 (13.6)	46 (4.5)	178 (8.9)
Lifetime	219 (22.5)	78 (7.6)	297 (14.9)
Alcohol abuse			
Past year	74 (7.6)	20 (2.0)	94 (4.7)
Lifetime	127 (13.2)	41 (4.0)	168 (8.4)
Alcohol dependence			
Past year	59 (6.1)	26 (2.6)	85 (4.3)
Lifetime	92 (9.5)	37 (3.6)	129 (6.4)
Binge drinking			
Never	517 (54.7)	781 (76.6)	1298 (68.3)
Less than once monthly	237 (25.1)	142 (13.9)	379 (19.9)
Monthly	77 (8.1)	20 (2.0)	97 (5.1)
Weekly	86 (9.1)	10 (1.0)	96 (5.1)
Daily	28 (3.0)	2 (0.2)	30 (1.6)
Missing	25 (2.6)	64 (6.3)	89 (4.5)

3, screening properties for all these different laboratory tests are quite similar.

*Screening performance in female patients (*n = *1021)*

The CAGE showed a lower diagnostic performance in female patients compared with male patients, having a sensitivity of 54% at a cut-point of ≥1. The AUDIT had a sensitivity of 65% at a cut-point of ≥5, with a specificity of 92%. It was clear that the AUDIT-C (cut-off value ≥5) was not an alternative to the AUDIT among female patients in a GP population, having a sensitivity of 50% and a specificity of 93%. The sensitivity of the Five-Shot was 63% at a markedly cut-point of ≥2.5 while its specificity was markedly higher, at 95%. Because of the lower prevalence in female patients it is important that the specificity is sufficiently high to avoid too many false positives. All of these tests gave very low PPVs but, in contrast, also gave very high NPVs (above 96%). For all of the questionnaires and cut-points that were examined the odds ratios were above 10 and were greater still at higher cut-point values.

No single laboratory test was shown to be appropriate for screening. Only the %CDT was able to confirm the diagnosis at a recommended cut-point of ≥6.

Areas under the ROC curves (AUC) and difference in AUC for male and female patients

Figure 1 shows the ROC curves for male patients. The CAGE had the lowest AUC of all the questionnaires with a value of 0.74 (95% CI = 0.71–0.77) for male patients and 0.76 (95% CI = 0.73–0.79) for female patients. Among male GP patients, comparing the different AUCs from the AUDIT and the derived version of the AUDIT gave remarkable results. Only a small range existed between the AUDIT (AUC = 0.85), the AUDIT-C (AUC = 0.83), the Five-shot (AUC = 0.84) and the AUDIT of Piccinelli (AUC = 0.83). Laboratory test

Table 3. Male patients in general practice: screening properties for alcohol abuse or dependence for different tests.

Questionnaire	Sensitivity	Specificity	PPV	NPV	Likelihood ratio (+) (95% CI)	Likelihood ratio (-) (95% CI)	Odds ratio (95% CI)
CAGE							
1	62.1	81.2	34.2	93.2	3.30 (2.72–4.00)	0.47 (0.37–0.58)	7.07 (4.67–10.78)
2	47.7	92.3	49.2	91.8	6.16 (4.59–8.27)	0.57 (0.48–0.67)	10.87 (6.92–17.09)
AUDIT							
5	82.6	72.9	32.4	96.4	3.05 (2.66–3.50)	0.24 (0.16–0.35)	12.78 (7.74–21.26)
6	74.2	81.4	38.6	95.3	3.99 (3.36–4.75)	0.32 (0.24–0.42)	12.62 (8.04–19.89)
7	67.4	85.7	42.6	94.4	4.71 (3.85–5.78)	0.38 (0.30–0.49)	12.40 (8.02–19.22)
8	60.6	90.3	49.7	93.6	6.28 (4.90–8.05)	0.44 (0.35–0.54)	14.40 (9.25–22.45)
AUDIT-C							
5	78.0	74.9	32.8	95.6	3.10 (2.68–3.60)	0.29 (0.21–0.41)	10.57 (6.64–16.91)
6	66.7	84.3	40.0	94.1	4.24 (3.48–5.16)	0.40 (0.31–0.50)	10.71 (6.97–16.50)
8	48.5	94.3	57.1	92.1	8.47 (6.12–11.74)	0.55 (0.46–0.65)	15.51 (9.63–25.04)
AUDIT-PC							
5	68.2	83.9	40.0	94.4	4.24 (3.49–5.14)	0.38 (0.29–0.49)	11.17 (7.25–17.27)
6	58.3	91.5	52.0	93.3	6.89 (5.29–8.99)	0.46 (0.37–0.56)	15.14 (9.67–23.76)
7	45.5	95.7	62.5	91.8	10.59 (7.32–15.34)	0.57 (0.49–0.67)	18.59 (11.17–31.03)
8	37.9	97.5	70.4	90.9	15.13 (9.41–24.35)	0.64 (0.56–0.73)	23.75 (13.08–43.42)
Five-Shot							
1.5	93.2	50.2	22.7	97.9	1.87 (1.72–2.03)	0.14 (0.07–0.26)	13.76 (6.64–29.57)
2.0	86.4	63.6	27.2	96.7	2.38 (2.12–2.66)	0.21 (0.14–0.33)	11.09 (6.42–19.36)
2.5	74.2	80.9	38.0	95.2	3.89 (3.28–4.62)	0.32 (0.24–0.43)	12.23 (7.79–19.26)
3.0	62.1	88.3	45.6	93.7	5.32 (4.23–6.69)	0.43 (0.34–0.53)	12.40 (8.04–19.17)
Laboratory tests							
MCV	39.4	75.0	19.9	88.7	1.57 (1.24–2.00)	0.81 (0.70–0.93)	1.95 (1.30–2.90)
GGT	6.8	95.5	19.1	86.7	1.44 (0.78–2.65)	0.98 (0.93–1.02)	1.54 (0.67–3.44)
ASAT (GOT)	10.9	92.6	28.6	79.1	2.44 (1.39–4.31)	0.93 (0.87–0.99)	2.63 (1.33–5.14)
ALAT (GPT)	11.4	96.4	33.3	87.4	3.18 (1.76–5.74)	0.92 (0.86–0.98)	3.46 (1.71–6.94)
Uric acid	10.9	96.9	35.0	87.3	2.76 (1.75–4.35)	0.92 (0.87–0.98)	3.71 (1.77–7.70)
%CDT	18.2	95.6	39.0	88.0	4.09 (1.47–11.39)	0.86 (0.78–0.94)	4.78 (1.48–17.11)

Table 4. Female patients in general practice: diagnostic properties for alcohol abuse or dependence for different tests.

Questionnaire	Sensitivity	Specificity	PPV	NPV	Likelihood ratio (+) (95% CI)	Likelihood ratio (-) (95% CI)	Odds ratio (95% CI)
CAGE							
1	54.3	92.1	24.5	97.7	6.88 (4.89–9.68)	0.50 (0.36–0.68)	13.88 (7.07–27.33)
2	37	96.8	35.4	97	11.62 (6.96–19.40)	0.65 (0.52–0.81)	17.85 (8.34–38.16)
AUDIT							
5	65.2	91.9	27.5	98.2	8.05 (5.97–10.85)	0.35 (0.25–0.56)	21.27 (10.58–43.12)
6	58.7	95.9	40.3	98.0	14.31 (9.70–21.10)	0.43 (0.31–0.61)	33.22 (16.13–68.84)
7	56.5	97.6	53.1	97.9	23.96 (14.87–38.60)	0.45 (0.32–0.62)	53.81 (24.71–118.30)
8	50.0	98.7	63.9	97.7	37.50 (20.33–69.18)	0.51 (0.38–0.68)	74.00 (31.01–179.46)
AUDIT-C							
5	50.0	93.2	25.8	97.5	7.39 (5.10–10.71)	0.54 (0.40–0.72)	13.77 (6.97–27.24)
6	39.1	97.3	40.9	97.1	14.67 (8.70–24.76)	0.63 (0.50–0.79)	23.46 (10.82–50.92)
7	28.3	99.0	56.5	96.7	27.55 (12.76–59.48)	0.72 (0.60–0.87)	38.02 (14.21–102.89)
8	21.7	99.6	71.4	96.4	52.99 (17.27–162.60)	0.79 (0.67–0.92)	67.43 (18.06–273.08)
AUDIT-PC							
5	56.4	95.7	38.2	97.9	13.12 (8.89–19.37)	0.45 (0.33–0.63)	28.88 (14.13–59.33)
6	41.3	98.8	61.3	97.3	33.56 (17.36–64.90)	0.59 (0.47–0.76)	56.47 (23.08–140.17)
7	30.4	99.1	60.9	96.8	32.97 (15.06–72.17)	0.70 (0.58–0.85)	46.96 (17.32–129.60)
8	19.6	99.5	64.3	96.3	38.15 (13.32–109.30)	0.81 (0.70–0.93)	47.19 (13.43–173.63)
Five-Shot							
1.5	80.4	73.4	12.5	98.8	3.03 (2.54–3.61)	0.27 (0.15–0.48)	11.37 (5.15–25.96)
2.0	67.4	87.4	20.1	98.3	5.34 (4.12–6.93)	0.37 (0.25–0.57)	14.32 (7.17–28.89)
2.5	63.0	94.7	35.8	98.2	11.82 (8.37–16.69)	0.39 (0.27–0.57)	30.28 (14.84–62.29)
3.0	37.0	97.3	39.5	97.0	13.86 (8.12–23.66)	0.65 (0.52–0.81)	21.40 (9.80–46.69)
Laboratory test							
MCV	41.3	79.3	8.6	96.6	2.00 (1.39–2.89)	0.52 (0.37–0.74)	2.71 (1.41–5.20)
GGT	6.5	91.8	3.6	95.4	0.79 (0.26–2.42)	1.02 (0.94–1.10)	0.78 (0.19–2.72)
ASAT (GOT)	6.5	97.9	13.0	95.7	3.18 (0.98–10.32)	0.95 (0.88–1.03)	3.33 (0.75–12.61)
ALAT (GPT)	0.0	98.6	0.0	95.4	Not computable		
Uric acid	6.5	96.4	7.9	95.6	1.82 (0.58–5.69)	0.97 (0.90–1.05)	1.87 (0.44–6.77)
%CDT	15.2	95.5	14.0	96.0	3.38 (1.13–10.10)	0.89 (0.78–1.01)	3.81 (0.99–15.01)

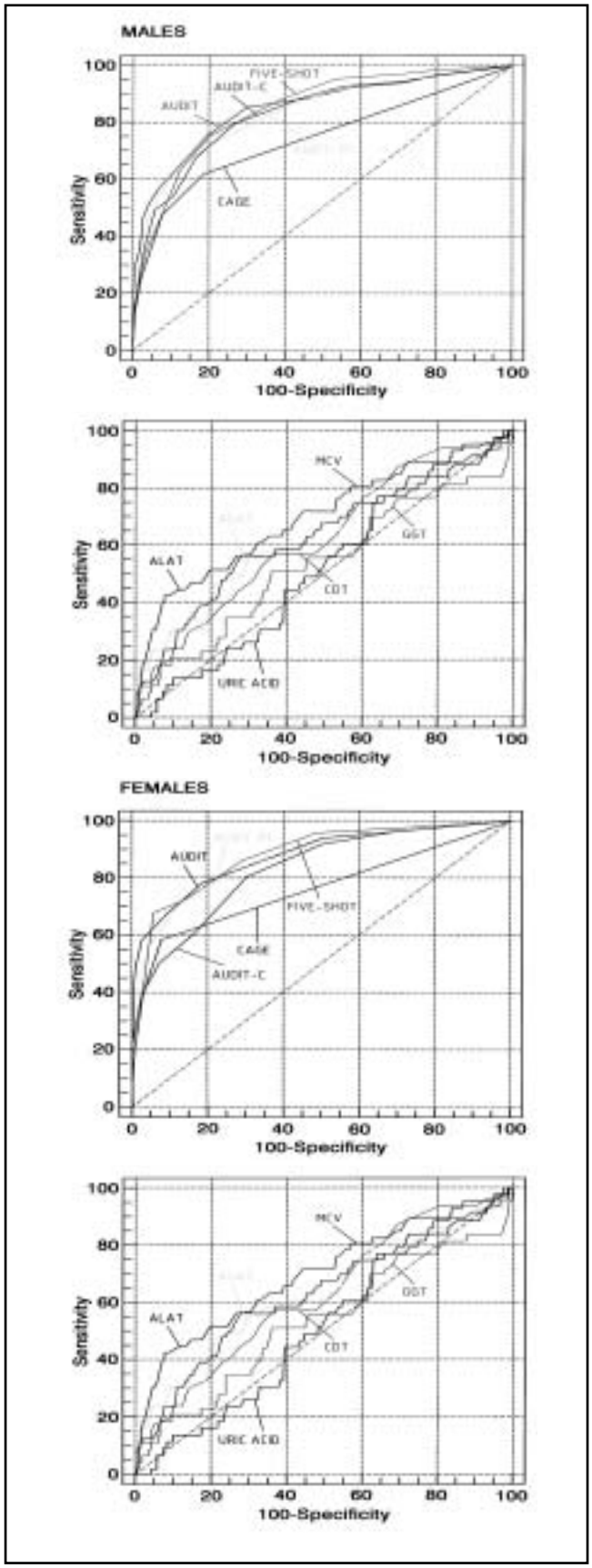

Figure 1. ROC curves for male and female patients in general practice.

results gave AUCs from 0.57 (GOT) to 0.66 (%CDT). However, the differences between the conventional laboratory tests and the %CDT were small and statistically not significant. Figure 1 illustrates the small differences between conventional laboratory tests and the %CDT. Figure 1 also depicts the ROC curves for the questionnaires with one outlier: the CAGE.

Figure 2 presents the ROC curves for female patients. The Five-Shot has the best results among a female population with an AUC of 0.88 (95% CI = 0.86–0.90). The CAGE performed better among female patients than among male patients (AUC = 0.76; 95% CI = 0.73–0.79), but performed less well than the AUDIT (0.87, 95% CI = 0.85–0.89). The optimal cut-points were lower for female compared with male primary care attenders. Within this population the AUDIT-C performs less well than expected, having an AUC of 0.82 (95% CI = 0.80–0.85).

There was no significant difference between the performance of either GGT and MCV compared with the %CDT in this female patient population.

Discussion

Screening for alcohol abuse or dependence remains a difficult issue within primary care. Despite many questionnaires and efforts to introduce them into the busy practice of a family physician, there is still a need to continuously develop and select the most powerful and user-friendly questionnaire in general practice. However, a powerful test in a high prevalence specialised setting may not be appropriate in a low prevalence setting, such as in primary care.[41] We therefore designed this study in general practice to assess the screening properties of several short questionnaires, some conventional laboratory tests, and a relatively new marker for higher alcohol intake, the %CDT.

The prevalence of alcohol abuse or dependence in the previous year (8.9%) in our study population is similar to other studies[1,3-6,42] and the sex ratio was the same as in the study by Hill (1 female:2.8 male). Marital status, educational level or work circumstances did not provide us with diagnostic clues or increased prior odds.

With a sensitivity of 62% for males and 54% for females, the CAGE was proven to be an insufficient screening instrument to detect alcohol abuse or dependence among primary care patients. Among male patients attending their GP, the AUDIT seems to be a powerful screening tool with a high sensitivity (83%) and a reasonable specificity but at a lower cut-off (≥5) than usually recommended. At the usual cut-point of 8 we found a sensitivity of 60% that is comparable with other studies in similar populations.[43,44] Evidence that the AUDIT seems equally appropriate for males and females, until now, was based on only a few studies and the WHO test development samples.[27] Our results support these findings. The AUDIT-C — using only the first three questions of the AUDIT as suggested by Bush[31], at a cut-off point ≥5 — results in nearly the same diagnostic parameters for a male patient population. Given the fact that this questionnaire is shorter and easier to use, we confirm the results of Bush and find it an efficient screening instrument in male primary care patients. Furthermore, the AUC of the AUDIT-C of our study for male GP patients fits very well with their

results among patients from three Veteran Affairs general medical clinics.

To detect alcohol abuse or dependence among female patients is much more problematic. Even an often-studied questionnaire, such as the AUDIT, with a cut-point of 65 only detects 65% of the current DSM-III-R-positive female patients. At this low cut-off point, a large group of female patients is falsely labelled with a drinking problem and results in a positive predictive value of 27% only. For this population the Five-Shot questionnaire has the best screening properties. It is also shorter compared with the AUDIT and therefore easy to use as a screening instrument for the busy GP.[28]

Once again, we confirm that conventional laboratory tests are of no use for detecting alcohol abuse or dependence in a primary care setting.[16] Even the new alcohol marker, %CDT, cannot be used as a screening instrument in this population. Nevertheless, this being the case, %CDT seems to be the best alcohol marker available, although the difference between the MCV and %CDT is small and our study confirms the results of Meerkerk *et al*.[17]

As pointed out by Hill *et al*,[6] one of the major disadvantages of most diagnostic studies in this field is that most of the previous studies relied on screening instruments alone to determine prevalence rates. Furthermore, no differentiation was made between abuse and dependence and between current and lifetime prevalence rates. We made sure to include these specifications in our report.

With a population of 1992 patients, this study is probably one of the largest on the diagnosis of alcohol abuse or dependence, using DSM criteria as the reference standard within a general practice population. As a result, we were able to calculate accurate diagnostic parameters with relatively small 95% confidence intervals.

To perform a reliable diagnostic study we recognised three criteria. The first was to include consecutive patients (or a random sample of them) from a well-defined population. The second was the need for blinding between the test evaluators and the reference standard evaluators, as well as between multiple test evaluators, if applicable, as it is the case in this study. The third was that verification bias (the relationship between the likelihood of receiving the gold standard and the results of the evaluated test) should be avoided. Our study design complied completely with the first two requirements. We also fulfilled the third criterion as far as questionnaires and routine laboratory tests are concerned. With respect to the %CDT characteristics we also complied with this criterion, after weighing the results of DSM-negative patients for the inverse of the sample proportions of this fully randomised sample.

As a reference standard we used DSM-III-R criteria, as no Dutch language version of the CIDI 2.0 (generating DSM-IV criteria) was available at the start of the data collection. DSM-III-R criteria were operationalised using the CIDI questionnaire which was fully validated for this purpose.[35] However, we used the CIDI as an auto-questionnaire which, though not validated as such, had exactly the same wording as the interview of the 'paper and pencil' version. It also follows the same structure as the interviewer using the validated paper and pencil version and was used in exactly the same way. We therefore have no reason to believe that this might result in any relevant bias with respect to our study results.

For both questionnaires and for the reference standard we have only reported past-year results and prevalence rates of alcohol abuse or dependence; this is because we believe that only an active alcohol problem is worthwhile being detected by the GP. Our study population also contains a large sample of female patients and we can therefore present results for this group in which problem drinking generally is less studied.

We can recommend the use of the Five-Shot in a male as well as in a female GP population. While the four questions of the CAGE did not perform well, a combination of the last three questions and the first two AUDIT questions (i.e. the Five-Shot questionnaire) seems to have good diagnostic test characteristics according to DSM-III-R criteria for alcohol abuse or dependence.

At a cut-point of ≥ 2.5 the Five-Shot detects more than twice as many patients with alcohol abuse or dependence as the GP, based on his previous encounters. Nevertheless, with a specificity of 81% (male general practice patients) and 95% (female general practice patients) we must keep in mind that two out of three patients who were screened positive with the Five-Shot will not have an alcohol problem (PPV = 36% for females, 38% for males). On the other hand it should be remembered that, apart from its diagnostic performance, a routine administration of a brief screening questionnaire could serve the GP well as a basis for discussing alcohol problems. A positive screening test should also lead to further investigation of drinking and related problems, so that patients can be offered brief interventions or referrals as appropriate.

Acknowledgments

This research was partially supported by a grant from Merck-Belgolabo. Orange Medical generously provided free agents and micro-anion exchange columns for the determination of the AXIS %CDT. The authors are greatly thankful to the participating GPs for recruiting the patients; to Lieve Heyvaert and Peter Vermaelen for their dedicated work in visiting GPs, gathering their questionnaire responses, and entering them into the computer; Professor Blanckaert for his fast and accurate supervision of the laboratory tests; all the other laboratories for collecting blood samples; and Marina Devis for providing the finishing touch to the manuscript.

References

1. Cleary PD. Prevalence and recognition of alcohol abuse in a primary care population. *Am J Med* 1998; **85:** 466-471.
2. Phelps GL, Johnson NP. Bright light in dark places: physician recognition of alcoholism. *JSC Med Assoc* 1990; **86(1):** 17-18.
3. Magruder-Habib K, Durand AM, Frey KA. Alcohol abuse and alcoholism in primary health care settings. *J Fam Pract* 1991; **32:** 407-413.
4. Buchsbaum DG, Buchanan RG, Lawton M, *et al*. Alcohol consumption patterns in a primary care population. *Alcohol Alcohol* 1991; **26:** 215-220.
5. Leckman AL, Umland BE, Blay M. Prevalence of alcoholism in a family practice center. *J Fam Pract* 1984; **18:** 967-970.
6. Hill A, Rumpf H-J, Hapke U, *et al*. Prevalence of alcohol abuse and dependence in general practice. *Alcohol Clin Exp Res* 1998; **22:** 935-940.
7. Burge SK, Schneider FD. Alcohol-related problems: recognition

and intervention. *Am Fam Physician* 1999; **59:** 361-370, 372.
8. Allen J, Litten RZ, Lee A. What you need to know: detecting alcohol problems in general medical practice. *Singapore Med J* 1998 **39:** 38-41.
9. Fleming MF, Barry KL, Manucel CB, *et al*. Brief physician advice for problem drinkers. A randomised controlled trial in community-based primary care practices. *JAMA* 1997; **227:** 1039-1045.
10. Anderson P. Effectiveness of general practice interventions for patients with harmful alcohol consumption. *Br J Gen Pract* 1993; **43:** 386-389.
11. Romelsjo A, Andersson L, Barner H, *et al*. A randomized study of secondary prevention of early stage problem drinkers in primary health care. *Br J Addict* 1989; **84:** 1319-1327.
12. Senft RA, Polen MR, Freeborn DK, *et al*. Brief intervention in a primary care setting for hazardous drinkers. *Am J Prev Med* 1997, **13:** 464-470
13. Deehan A, Marshall EJ, Strang J. Tackling alcohol misuse: opportunities and obstacles in primary care. *Br J Gen Pract* 1998 **48:** 1779-1782.
14. McAvoy BR. Training general practitioners. *Alcohol Alcohol* 1997; **32:** 627-628.
15. Cowan PF. An intervention to improve the assessment of alcoholism by practising physicians. *Fam Pract Res J* 1994; **14:** 41-49.
16. Hoeksema HL, De Bock GH. The value of laboratory tests for the screening and recognition of alcohol abuse in primary care patients. *J Fam Pract* 1993; **37:** 268-276.
17. Meerkerk G-J, Njoo KH, Bongers IMB, *et al*. Comparing the diagnostic accuracy of carbohydrate-deficient transferrin, gamma-glutamyl transferase, and mean cell volume in a general practice population. *Alcohol Clin Exp Res* 1999; **23:** 1052-1059.
18. Yamauchi M, Hirakawa J, Maezawa Y, *et al*. Serum level of carbohydrate deficient transferrin as a marker of alcoholic liver disease. *Alcoholism Clin Exp Res* 1993; **28:** 3-8.
19. Lesh OM, Walter H, Freitag H, *et al*. Carbohydrate-deficient transferrin as a screening marker for drinking in a general hospital population. *Alcoholism Clin Exp Res* 1996; **31:** 249-256.
20. Maisto SA, Connors GJ, Allen JP. Contrasting self-report screens for alcohol problems: a review. *Alcohol Clin Exp Res* 1995; **19:** 1510-1516.
21. Ewing JA. Detecting alcoholism, the CAGE questionnaire. *JAMA* 1984; **252:** 1905-1907.
22. Babor TF, Ramon de la Fuente J, Saunders J, *et al. AUDIT The Alcohol Use Disorders Identification Test: guidelines for use in primary health care.* Geneva: Substance Abuse Department, World Health Organization, WHO/PSA/92.4, 1992.
23. Aertgeerts B, Buntinx F, Ansoms A, *et al*. Is there a difference between CAGE interviews and the written CAGE questionnaires? *Alcohol Clin Exp Res* 2000; **24(5):** 733-736.
24. Perdrix A, Decrey H, Pecoud A, *et al*. Detection of alcoholism in the medical office: applicability of the CAGE questionnaire by the practising physician. *Schweiz Med Wochenschr* 1995; **125:** 1772-1778.
25. Lawner K, Doot M, Gausas J, *et al*. Implementation of CAGE alcohol screening in a primary care practice. *Fam Med* 1997; **29:** 332-335.
26. Lairson DR, Harrist R, Martin DW, *et al*. Screening for patients with alcohol problems: severity of patients identified by the CAGE. *J Drug Educ* 1992; **22:** 337-352.
27. Allen JP, Litten RZ, Fertig JB, *et al*. Review of research on the Alcohol Use Disorders Identification Test (AUDIT). *Alcohol Clin Exp Res* 1997; **21:** 613-619
28. Seppa K, Lepisto J, Sillanaukee P. Five-Shot Questionnaire on heavy drinking. *Alcohol Clin Exp Res* 1998; **22**: 1788-1791.
29. Piccinelli M, Tessari E, Bartolomasi M, *et al*. Efficacy of the alcohol use disorder identification test as a screening tool for hazardous alcohol intake and related disorders in primary care. A validity study. *BMJ* 1997; **314:** 420-427.
30. Seppa K, Makela R, Sillanaukee P. Effectiveness of the Alcohol Use Disorders Identification Test in occupational health screenings. *Alcohol Clin Exp Res* 1995; **19:** 999-1003.
31. Bush K, Kivlahan DR, McDonnel MB, *et al*. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. *Arch Intern Med* 1998; **16:** 1789-1795.
32. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental disorders*. 3rd revised edition. Washington DC: American Psychiatric Press, 1987.
33. Hasin D, Paykin A. Alcohol dependence and abuse diagnosis: concurrent validity in a nationally representative sample. *Alcohol Clin Exp Res* **23:** 1999; 144-150.
34. Robins LN. The Composite International Diagnostic Interview. An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Arch Gen Psychiatry* 1988; **45:** 1069-1077.
35. Wittchen HU, Robins LN, Cottler LB, *et al*. Cross-cultural feasibility, reliability and sources of variance of the Composite International Diagnostic Interview (CIDI). The Multicentre WHO/ADAMHA Field Trials. *Br J Psychiatry* 1991; **159:** 645-658.
36. Mayfield DG. The CAGE-questionnaire; validation of a new alcoholism screening instrument. *Am J Psychiatry* 1974; **131:** 1121-1123.
37. Dean AG. *Epi Info, version 6: a word processing, database, and statistics program for epidemiology on microcomputers*. Atlanta, GA: Centers for Disease Control and Prevention, 1994.
38. Hanley JA, McNeil BJ. The meaning and use of the area under a receiving operating characteristic (ROC) curve. *Radiology* 1982; **143:** 29-36.
39. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristics curves derived from the same cases. *Radiology* 1983; **148:** 839-843.
40. Agresti A. *Categorical data analysis.* New York: Wiley, 1990.
41. Knottnerus JA, Leffers P. The influence of referral patterns on the characteristics of diagnostic tests. *J Clin Epidemiol* 1992; **45(10):** 1143-1154.
42. Cornel M, Knibbe RA, Van Zutphen WM, Drop MJ. *Problem drinking in a general practice population*. [Thesis.] Maastricht, 1993.
43. Schmidt A, Barry KL, Fleming MF. Detection of problem drinkers: the Alcohol Use Disorder Identification Test (AUDIT). *Southern Med J* 1995; **88:** 52-59.
44. Isaacson JH, Butler R, Zacharek M, *et al*. Screening with the Alcohol Use Disorder Identification Test (AUDIT) in an inner-city population. *J Gen Intern Med* 1994; **9**: 550-553.