

TITLE PAGE

Confirming Diagnosis of Hazardous and Harmful Alcohol Use

Diagnostic accuracy of a computer assisted diagnostic system compared to conventional markers of alcoholism.

Aleksander Korzec, MD, Carla de Bruijn, MD, Michiel van Lambalgen, PhD

Key words: sensitivity, specificity, harmful alcohol use, confirmation, computer assisted diagnostic system

Word count paper : 3240

Word count abstract : 293

Authors Affiliations: Sint Lucas Andreas Hospital, Amsterdam, the Netherlands, (Korzec, psychiatrist and de Bruijn, psychiatric resident); University of Amsterdam, Amsterdam, The Netherlands (van Lambalgen, professor of logic and artificial intelligence).

Corresponding author:

A. Korzec, MD. Sint Lucas Andreas Ziekenhuis, box 9423, 1006 AE Amsterdam, the Netherlands tel. 0031 20 5108562 fax 0031 20 6126431

e-mail a.korzec@slaz.nl

Acknowledgements

All laboratory tests, except carbohydrate-deficient transferrin (CDT) were performed in the Department of Clinical Chemistry, St Lucas Andreas Hospital, Amsterdam, the Netherlands (Dr W de Kieviet). CDT was performed at *bioscientia*, Ingelheim, Germany.

ABSTRACT

Context Conventional tests for alcoholism have good screening properties for detecting hazardous and harmful alcohol use (HHAU), but they fail to confirm this diagnosis accurately and objectively.

Objective To validate a Bayesian Alcoholism Test (BAT) for confirming the diagnosis of hazardous and harmful alcohol use

Design and Subjects BAT is based on studies on the prevalence of HHAU and other diseases causing similar abnormalities, and on conditional probabilities of these disorders and associated biochemical markers and clinical signs.

BAT was compared to carbohydrate-deficient transferrin (CDT), and gamma-glutamyltransferase (GGT) in three populations: 67 treatment seeking alcoholics, 68 non-treatment seeking heavy drinkers and 47 controls. The diagnostic criterion was reported alcohol consumption.

Main Outcome Measures Test sensitivity and specificity, likelihood ratio's, receiver operating characteristic curves, comparing BAT, CDT, and GGT.

Results Comparing alcoholics and controls, sensitivity of BAT (94%) was significantly higher than CDT (63%) and GGT (73 %). The specificity of BAT (98 %) was not significantly higher than CDT (93 %) and GGT (92%). The area under the curve for BAT (0,989) was significantly higher than the area under the curve for CDT (0,909) and area under the curve for GGT (0,902).

In the group of heavy drinkers, BAT could differentiate significantly better than GGT between heavy drinkers above a harmful level, heavy drinkers with a hazardous use and heavy drinkers below a hazardous consumption. The difference with CDT was not statistically significant.

Using pooled data of all 182 subjects included in the study, the amount of drinking had a significant better correlation coefficient with BAT (0.795) than with CDT (0.657), and GGT (0.604).

Conclusions BAT has better diagnostic properties than CDT and GGT for confirming HHAU. Additional studies are necessary in other clinical populations to obtain definitive results about our diagnostic system.

INTRODUCTION

Alcoholism has severe consequences for society. Alcoholism is a heterogeneous concept that includes psychiatric diagnoses such as alcohol dependence, alcohol abuse and harmful use, as well as less severe drinking patterns, often referred to as heavy, hazardous or excessive drinking.¹ The direct and indirect costs of alcoholism are relatively constant in different countries in Europe and North America. Depending on the calculation used, these costs have been estimated to be between 1% and 2% of the gross national product (GNP).²⁻⁶ Concurrent with these estimates, research suggests a considerable prevalence of alcoholism in the general population. Epidemiological estimates about the point prevalence of excessive use of alcohol in the general population vary between 4-29% for hazardous drinking and 1-10% for harmful drinking, depending on country, the criteria for harmful and hazardous drinking and the screening instruments used.¹

Several studies indicate that, even after active screening, general practitioners identify maximally 60% of their alcoholic patients.⁷⁻⁹ The main reasons for underdiagnosis are denial on the part of patients^{10,11}, insufficient sensitivity of screening instruments in detecting patients with less severe alcoholism¹², insufficient skills of physicians, and questioning the rationale of diagnosis and intervention in young hazardous drinkers.¹³

In diagnosis, clinicians begin with different estimates of an a priori probability about the presence of a disease. According to these estimates, a diagnostic test may be used for screening, exclusion or confirmation.¹⁴ If the patient is unwilling to disclose alcoholism, or is not aware of alcohol related problems, there is no accurate diagnostic test to confirm objectively the diagnosis. There is evidence that alcoholic patients who deny or who are not aware of their condition can benefit from feed back of abnormal laboratory results^{15,16}, and also some evidence that physicians hesitate to confront patients without robust confirmatory

evidence.^{17,18} In forensic^{19,20}, insurance²¹, occupational²² and pre-operative settings^{23,24}, there is a strong need of a highly specific confirmation test of alcoholism.

This paper presents an expert system, Bayesian Alcoholism Test (BAT), to facilitate the confirmation of the diagnosis of a Hazardous and Harmful Alcohol Use (HHAU). A diagnostic expert system is a computer program that combines information about a disease, in this case alcoholism, in such a way that feeding in data about a particular patient (e.g. values of selected blood markers and clinical signs) yields a probability that the patient suffers from HHAU. Expert systems are useful when there is a large number of diagnostic tests and when the relationship between the disease to be diagnosed and the result of tests is of a probabilistic nature. Although the probabilistic computations involved are complex, with the advent of so-called Bayesian networks²⁵, mathematical and computational technology has now progressed so far as to make an expert system of the size we need entirely feasible. The expert system allows us to answer queries of the following type: given values obtained for some, but not necessarily all, diagnostic tests, what is the probability that a patient suffers from a particular disease? An advantage of the expert system above single diagnostic tests, is that it allows combining the results of many tests, which is common practice in diagnostics. This in contrast to the vast literature on diagnostic tests, where mostly single tests are considered.²⁶

BAT has been constructed from a literature survey, which yielded epidemiological data for about 40 % of the probabilities. The remaining probabilities were obtained by consulting experts.

The hypothesis investigated in this study is that BAT is a more accurate tool to confirm the diagnosis of HHAU than other, currently used tests, such as CDT and GGT.

Methods

Study design and study populations

The study design is a prospective cross-sectional validation study of diagnostic accuracy. The ethical committee of the St. Lucas Andreas Hospital approved the study protocol. All participants of the study gave their informed consent; the research was carried out according to the provisions of the Declaration of Helsinki of 1975, as revised in 1996. All subjects were recruited between 1998 and 2001.

We aimed to test our diagnostic system in a broad spectrum of the disease. First, we investigated whether the system was able to distinguish “clear alcoholics” from “social drinkers”. Thereafter we tested the diagnostic differentiating ability of BAT within the population of heavy drinkers (representing the spectrum in-between alcoholics and social drinkers). Three study groups were formed: controls, treatment seeking alcoholics and non treatment-seeking heavy drinkers. All subjects were male.

Non-alcoholic controls (group 1) were 79 consecutive ambulatory psychiatric patients (Sint Lucas Andreas Hospital, Amsterdam). Only patients were included in the control group (n=47) without HHAU (defined as an alcohol consumption of ≤ 280 g ethanol/week^{27,28}) in the last 4 weeks before blood sampling, and no Alcohol Use Disorder (AUD) in the last year. AUD in all groups was defined as having a disorder in accordance with the International Classification of Mental and Behavioral Disorders (ICD-10)²⁹ or in accordance with the Diagnostic Statistical Manual of Mental Disorders, (DSM-IV).³⁰

Alcoholics (group 2) were recruited from addiction treatment facilities: 73 patients consecutively admitted to a detoxification ward (Jellinek clinic, Amsterdam) and 29 consecutive patients attending an ambulatory alcoholism treatment center (Brijder stichting, Zaandam). Only patients with harmful use (defined as an alcohol intake of > 560 g

ethanol/week²⁸) in the last 4 weeks prior to examination and with an AUD diagnosis were included in the alcoholics group (n=67).

Non treatment seeking heavy drinkers (group 3) were recruited at wine-tasting conventions and by advertisements in a wine magazine, in which we informed them of the relations between alcohol and health and of the object of our study. All subjects (n=68) were included in the study group. The selection and characteristics of this group has been described previously.³¹

The data were collected by 1 psychiatrist and 6 psychiatric residents who received prior training about the instruments. For each subject all data were collected on the same day. The readers of the criterion assessments were blind to the results of the other tests and vice versa.

Instruments

1. Criterion assessments (alcohol intake and AUD diagnosis)

Since diagnostic accuracy of tests is almost always based on a correct definition of false-positives and false-negatives, we used recommended and validated diagnostic instruments for assessing AUD and alcohol intake, as criterion standard.³²

Alcohol intake was assessed using Time Line Follow Back (TLFB).³³ The TLFB is a comprehensive retrospective self-report survey that enables the collection of reliable information on drinking behavior. The amount of alcohol was documented in standardized alcoholic units (AU), a standard drink in the Netherlands containing approximately 10 grams of ethanol.

The alcohol section of the Composite International Diagnostic Interview (CIDI-2.1), section J on alcohol, was used to assess symptoms of alcohol use disorders. The CIDI is a validated and reliable, fully structured, diagnostic interview which enables making diagnoses according to ICD-10 and DSM-IV criteria.^{34,35}

2. Diagnostic system BAT

The diagnostic system was based upon a literature search, combined with clinical expertise when literature was inconsistent or not available. We searched for studies on three topics: prevalence of disorders, prevalence of clinical signs, and conditional probabilities between disorders on the one side, and associated biochemical markers and clinical signs on the other side. The investigated disorders were alcoholism and the common disorders that can cause similar clinical signs and biochemical abnormalities: liver diseases, adiposity and diabetes. The search was performed in Pubmed. We limited the search to original articles and reviews published in English between 1970 and 2002. The literature was extended by search of the reference sections of the articles obtained, and by consulting textbooks.

Since our study groups included only men, we included mostly data of studies on men in Western Europe or in the United States and of studies performed in the general male population. If such data were not available, we used studies with mixed male and female populations, primary care populations, clinical populations etc. A total of 68 studies was included. Two reviewers appraised all articles for methodological content and results. If, on certain prevalence or condition, no studies were found, we used estimates obtained by consulting experts. The interested reader is welcome to contact us for more information regarding the literature search outcomes. See also the homepage of the paper at:

<http://www.science.uva.nl/staff/~Michiell>, for a summary table of data that were used as input into BAT.

The data mentioned above were used to create a Bayesian network, a graphical structure the nodes of which represent diseases, symptoms and biochemical tests, and where an arrow going from disease to symptom or biochemical test, indicates that the symptom or test is dependent on the disease (FIGURE 1). Apart from their graphical structure, the Bayesian network works with conditional probability tables that give the conditional probability

distribution of a disease causing different symptoms and biochemical abnormalities. The two kinds of information, graphical and probabilistic, are combined and result in probabilities that a patient is suffering from different diseases. BAT combined the results of the components listed below and showed a probability for each subject to suffer from HHAU, as well for diabetes and for liver disease.

3. BAT components

a. Clinical signs

As an indication of level of response to alcohol (LRA), subjects were asked how many units were required to become aware of an effect. Furthermore, the average number of cigarettes smoked per day was documented. Each participant was also asked the four CAGE questions (acronym based on its four questions: **C**ut down drinking, **A**nnoyed by criticism about drinking, **G**uilty feelings about drinking, and **E**arly drink [drinking in the morning]).³⁶

Physical consequences of alcohol use were assessed by a standardized physical examination, including inspection of the skin (spider) and palpation of the liver, and by biochemical tests.

b. Biochemical tests

Venous blood samples were taken for determination of mean cell volume (MCV), carbohydrate-deficient transferrin (CDT), Gamma-glutamyltransferase (GGT), aspartate aminotransferase (AST), alanine aminotransferase (ALT) and alkaline phosphatase (AP).

These biochemical markers (with the exception of AP) are known as indicators of enzymatic induction or cellular damage due to alcohol, and are predictive of adverse health outcomes.³⁷

For details concerning the analytical procedures see our report on diagnosing alcoholism in drinking drivers.¹⁹ In the present study, we used another CDT test: (ChronAlcoI.D. (Sangui Biotech Inc., USA). This test has been validated analytically and clinically.^{38,39} For CDT, ALT, AST, GGT, AP and MCV the upper reference limits were: CDT: 3.0, ALT: 45, AST: 40, GGT: 65, AP 135 U/l and MCV 97fL. The rationale for choosing CDT cutoff at 3.0 is

based on the recommendation in a study of a comparable CDT test⁴⁰. MCV cutoff was based on several studies comparing social drinkers and alcoholics.⁴¹⁻⁴³ The other cutoffs were recommended by the laboratory where the tests were performed. All biochemical tests were performed in the laboratory of the Sint Lucas Andreas Hospital, except CDT which was performed at bioscientia, Ingelheim, Germany.

4. Other measurements

The subjects were asked if they had diabetes or used (anti-diabetic) medication. Hepatitis risk was screened with questions on earlier hepatitis, intravenous drug use, or blood transfusion before 1985. BMI was measured by weight and height measurement.

Data analysis

The statistics software HUGIN 5.7 was used for building the Bayesian network (Hugin Expert A/S, Aalborg, Denmark).

The statistics software Confidence Interval Analysis (CIA), version 2.05, was used for calculating diagnostic sensitivity, specificity and likelihood ratios. Confidence intervals for sensitivity, and specificity were computed using Wilson's method.⁴⁴ Confidence intervals of the

likelihood ratios was computed using the score method.⁴⁴ Receiver operating characteristic (ROC) analysis was performed with Statistics Package for Social Sciences (SPSS for Windows, 11.0, 2000). Area under the curve (AUC) was used as a measure of overall test accuracy. Differences of AUC between tests were examined according to the method by Hanley and McNeil.⁴⁵

Differentiation in the group of heavy drinkers was computed with entropy.⁴⁶ Entropy is a measure of information in distribution. Increase of entropy is associated with a decrease of available information and increase of uncertainty. Confidence intervals were calculated with the method suggested by Esteban and Morales.⁴⁷

The Spearman test with confidence intervals was performed with CIA to assess difference in correlations between alcohol intake and results of BAT, CDT and GGT in the combined populations of alcoholics, heavy users and controls.

RESULTS

Sample characteristics of the three selected groups are shown in table 1. The data were collected between 1998 and 2001. The mean age of the heavy drinkers group (49.3) was significantly higher than that of the alcoholics (43.6).

Cutoff level BAT

In order to make a comparison with other diagnostic tests of excessive alcohol use possible, we calculated the cutoff level which gave the best accuracy (total of true positives and true negatives divided by all subjects). The cutoff range with the best accuracy (95.6%) was between 41% and 50%. Because we aimed to design a confirmation test we chose 50% as cutoff level.

Test Sensitivity, specificity and likelihood ratio

In the group of 67 treatment seeking alcoholics the sensitivity of BAT was significantly better than CDT and GGT (table 2).

BAT also yields a probability of the presence of the diagnosis. 76.1% of the 67 alcoholics scored a probability of >95% in BAT of having hazardous or harmful use.

The specificity of BAT was not significantly higher than those of CDT and of GGT.

The positive and negative likelihood ratio's of BAT were superior to that of CDT and GGT (table 2).

False positives

There was only one subject of the 47 controls scoring positive with BAT, against three subjects showing an elevated CDT and four subjects with an elevated GGT. To the best of our knowledge, the subjects were not consuming alcohol above a hazardous level. The reasons for the subject scoring false positive on the BAT were abnormal results of ALT (176 U/l, normal range 5-50) AST (112 U/l, normal range 10-45) and GGT (163 U/l, normal range 10-65). This subject was also the only control subject scoring with BAT above 50% probability of having

hepatitis. The subject had no prior history of (intravenous) drug use or blood transfusion. Blood examination for antibodies for hepatitis B and C (anti-HCV and HbsAg) was negative.

False negatives

Of the 67 alcoholics, BAT did not recognize four subjects, against 24 subjects having normal CDT values and 18 subjects having normal GGT values.

Compared with the rest of the alcoholics, the subgroup of four subjects, not identified with BAT, was not different from those correctly identified. One subject was much younger than the average of the alcoholic population (29 years); another subject had a relatively low alcohol use, just above harmful use level (570g alcohol/week).

ROC curves

Comparison of the ROC curves (populations of alcoholics and controls, n=114) showed that BAT was superior to that of CDT and GGT (FIGURE 2). The area under the curve for BAT (0,989 [95% CI, 0.976-1.000]) was significantly higher ($p < 0.005$) than the area under the curve for CDT (0,909 [95% CI, 0.851-0.967]) and area under the curve for GGT (0,902 [95% CI, 0.848-0.957]).

Using receiver operating characteristic curves, 100% specificity was achieved, with a corresponding sensitivity of the BAT of 92 %, sensitivity of CDT of 28% and sensitivity of GGT of 49%.

Differentiating power of BAT in heavy drinking

Subsequently we investigated the differentiating ability of BAT, CDT and GGT to distinguish between three levels of drinking: harmful drinking (>560 g/week), hazardous drinking (280-560g/week) and non-hazardous use(<280 g/week).As can be seen in table 3 from the entropy values, BAT gives more information than GGT. The difference with CDT was not significant. Of the 19 heavy drinkers with harmful use, BAT identified 63 %, CDT identified 53 % and GGT identified 32 % of the subjects. The difference between BAT and CDT was not

significant (95% CI of the difference $-0,152 - 0,342$). The difference between BAT and GGT was significant (95% CI the difference $0,072 - 0,502$).

Correlation alcohol intake and test results

Using pooled data of all 182 subjects, BAT had a significant better correlation coefficient with the alcohol intake, 0.795 (95% CI 0.735-0.843) than CDT, 0.657 (95% CI 0.564-0.734) and than GGT 0.604 (95% CI 0.503-0.689).

COMMENT

The diagnosis of heavy drinking is difficult when dealing with subjects that deny excessive alcohol use or alcohol related problems. In case of suspicion, the available diagnostic tests are too insensitive and unspecific to be able to support the diagnosis in legal and health care settings. This study used a combination of clinical signs and biochemical tests and compared its diagnostic properties with available markers of excessive alcohol use.

The diagnostic system that we evaluated in this study has several advantages above the usual diagnostic tests for excessive alcohol use. First, our results indicate that, in the population of alcoholics, this test has better diagnostic properties than the regular tests. A second advantage is that BAT produces a probability that a subject is suffering from HHAU. A third advantage is that it also produces a probability that the clinical and biochemical abnormalities are caused by another disease. The fourth advantage above other suggestions for using combinations of biochemical tests for HHAU⁴⁸, is that BAT can be easily accommodated with a node of the expected prevalence of the disease, without changing cutoff values of the used tests.

However, our study has several limitations that deserve attention.

Firstly, our study results are applicable for men only. The conditional probability tables for women, especially for CDT, GGT, MCV and smoking, are different. Secondly, the external validity of our study must be considered. There might be a selection bias, causing BAT to perform better or CDT and GGT to perform worse. However, the sensitivity and specificity values of the usual alcohol use markers, CDT and GGT, were similar to those found in different other studies.^{49,50} Thirdly, the majority of the conditional probabilities used in designing BAT are based on estimates. Many of the studies we used had methodological shortcomings or produced inconclusive data. However, the prevalences that are not fed in BAT but were produced by BAT (such as prevalence of abnormal biochemical and clinical signs in the general population) are within the range indicated in the literature.

Our results failed to indicate that BAT is significantly better than CDT in identifying harmful drinkers in a population of heavy drinkers. There are some possible reasons for this. The first reason to consider is the relatively small sample of heavy drinkers. A second reason might be that wine drinkers are not representative of heavy drinkers in general. They smoke less than the general population and certainly less than usual heavy drinkers and they might have less drinking problems, as indicated by cage, than usual populations of heavy drinkers³⁰. Cage and smoking status influence the BAT score.

The clinical applicability of BAT is confirmation of HHAU in different settings. All data used to feed in the BAT system are results of standard clinical examination of subjects suspected to have alcohol problems. Feeding in the data and producing the BAT results takes less than 5 minutes per subject.

The present study includes only the first two phases of development of diagnostic tests.¹⁴ It should be further validated in other clinical settings (phase III), such as a population with liver diseases and eventually in a prospective consecutive series of clinically suitable patients (phase IV). Further research addressing these questions is necessary to obtain definitive results about our diagnostic system.

References

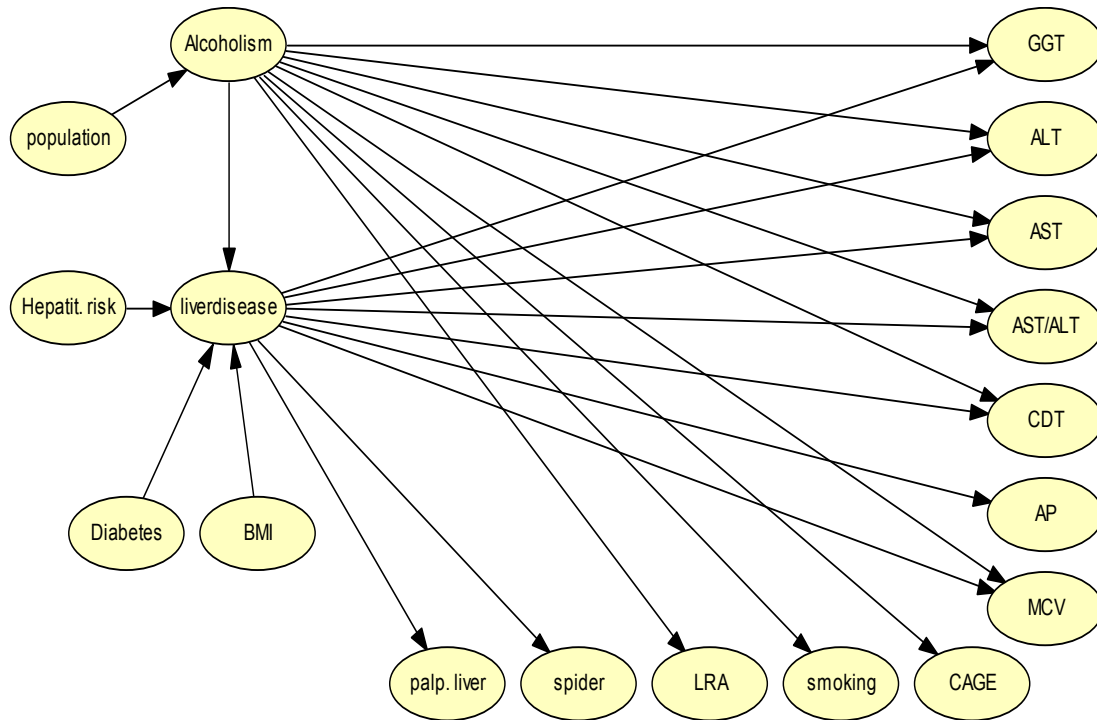
1. Carrington Reid M, Fiellin DA, O'Connor PG. Hazardous and harmful alcohol consumption in primary care. *Arch Intern Med* 1999;159:1681-1689
2. Reynaud M, Gaudun-Colombel AF, Le Pen C. Two methods of estimating costs linked to alcoholism in France. *Alcohol Alcohol* 2001;36:89-95
3. Brecht JG, Poldrugo F, Schädlich P. Alcoholism: The cost of illness in the Federal Republic of Germany. *Pharmaco-Economics* 1996;10:484-493.
4. Heien DM, Pittman DJ. The external costs of alcohol abuse. *J Stud Alcohol* 1993;54:302-307.
5. Rice DP, Kelman S, Miller LS. The economic cost of alcohol abuse. *Alcohol Health and Research World* 1991;15; 4-8.
6. Single E, Robson L, Xie X, Rehm J. The economic costs of alcohol, tobacco, and illicit drugs in Canada, 1992. *Addiction* 1998; 93: 991-998.
7. Cornell M. Detection of problem drinkers in general practice. Dissertation. 1994 Rijksuniversiteit van Limburg. Maastricht.
8. McQuade WH, Levy SM, Yanek LR, Davis SW, Liepman MR. Detecting symptoms of alcohol abuse in primary care settings. *Arch Am Med.* 2000;9:814-821.
9. Fiellin DA, Carrington Reid M, O'Connor PG. Screening for alcohol problems in primary care. *Arch Intern Med* 2000;160:1977-1989
10. Samet JH, O'Connor PG. Alcohol abusers in primary care: readiness to change behavior. *Am J Med* 1998;105:302-306
11. Ness DE, Ende J. Denial in the medical interview; recognition and management *JAMA* 1994;272:1777-1780
12. Rumpf HJ, Hapke U, Meyer C, John U. Screening for alcohol use disorder and at risk drinking in the general population: psychometric performance of three questionnaires. *Alcohol Alcohol* 2002;37:261-268
13. Beich A, Gannik D, Malteru K. Screening and brief intervention for excessive alcohol use: qualitative interview study of the experiences of general practitioners. *BMJ* 2002;325:870-872.
14. Feinstein A.R. Clinical Epidemiology. Philadelphia: WB Saunders. 1985.
15. Kristenson H Methods of intervention to modify drinking patterns in heavy drinkers. *Recent Dev Alcohol* 1987;5:403-423.
16. Wick HD. Alcohol problems in general practice – the value of biological markers. *Schweiz Rundsch Med Prax* 1999;88:1711-1714.

17. Cleary PD, Miller M, Bush BT, Warburg MM, Delbanco TL, Aronson MD. Prevalence and recognition of alcohol abuse in a primary care population. *Am J Med* 1988; 85:466-471.
18. Conigliaro J, Lofgren RP, Hanusa BH. Screening for problem drinking. Impact on physician behavior and patients drinking habits. *J Gen Intern Med* 1998;13:251-256.
19. Korzec A, Bär M, Koeter MWJ, de Kieviet W. Diagnosing alcoholism in high risk drinking drivers: comparing different diagnostic procedures with estimated prevalence of hazardous alcohol use. *Alcohol Alcohol* 2001; 36: 594-602
20. Malcolm R, Anton RF, Conrad SE, Sutherland S. Carbohydrate-deficient transferrin and alcohol use in medical Examiner cases. *Alcohol* 1999;17:7-11.
21. Bean P, Daniel P. Carbohydrate-deficient transferrin: Current facts and future projections for the insurance industry. *On the risk* 1996;12:43-48.
22. Curry SJ, Ludman E, Grothaus L, Gilmore T, Donovan D. At-risk drinking among patients in an occupational medicine clinic. *Alcohol Alcohol* 2002;37(3):289-94.
23. Kitchens JM. Does this patient have an alcohol problem? *JAMA* 1994;272:1782-1786.
24. Spies C, Tonnesen H, Andreasson S, Conigrave K. Perioperative morbidity and mortality in chronic alcoholic patients. *Alcohol Clin Exp Res* 2001; 25:164S-170S.
25. Lauritzen SL, Spiegelhalter DJ. Local computations with probabilities on graphical structures and their application to expert system. *Proc. Royal Statistic Society* 1988; B50:157-224.
26. Moons KGM, van Es GA, Michel BC, Büller HR, Dik J, Habbema F, Grobbee DE. Redundancy of the single test evaluation. *Epidemiology* 1999;10: 276-281
27. Saunders JB, Lee NK. Hazardous alcohol use: its delineation as a sub-threshold disorder, and approaches to its diagnosis and management. *Comprehensive Psychiatry* 2000;41:95-103
28. Conigrave KM, Degenhardt LJ, Whitfield JB, Saunders JB, Helander A, Tabakoff B. CDT, GGT, and AST as markers of alcohol use: The WHO/ISBRA collaborative project. *Alcohol Clin Exp Res* 2002;26:332-339
29. ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines. Geneva: World Health Organization, 1992.
30. Diagnostic and Statistical Manual of Mental Disorders, (1994) 4th edn. American Psychiatric Association Press, Washington, DC.
31. de Bruijn H, Korzec A, Arndt T, van den Brink W. The validity of alcohol use disorder in well-functioning men with hazardous alcohol use. *Eur Add Res*. In press
32. Allen JP, Fertig JB, Litten RZ, Sillanaukee P, Anton RF. Proposed recommendations for research on biochemical markers for problematic drinking. *Alcohol Clin Exp Res* 1997;21:244-247.

- 33 Sobell LC, Sobell MB. Timeline Followback: a technique for assessing self-reported alcohol consumption. In: Litten RZ, Allen JP, eds. *Measuring alcohol consumption: psychosocial and biological methods*. New Jersey: Humana Press 1996:41-72.
34. Cottler LB, Grant BF, Blaine J, Mavres V, Pull C, Hasin D, et al. Concordance of DSM IV alcohol and drug use disorder criteria and diagnoses as measured by AUDADIS-ADR, CIDI and SCAN. *Drug Alcohol Dep* 1997;47:195-205.
35. Üstün B, Compton W, Mager D, Babor T, Baiyewu O, Chatterji S, et al. WHO study on the reliability and validity of the alcohol and drug use disorder instruments: overview of methods and results. *Drug Alcohol Dep* 1997;47:161-9.
36. Mayfield D, Mcleod G, Hall P. The CAGE questionnaire: validation of a new alcoholism screening instrument. *Am J Psychiatry* 1974;131:1121-1123.
37. Conigrave KM, Saunders JB, Reznik RB, Whitfield JB: Prediction of alcohol-related harm by laboratory test results. *Clin Chem* 1993;39:2266-2270.
38. Korzec A, Arndt T, Bär M, Koeter MWJ: Trisialo-Fe₂-transferrin does not improve the diagnostic accuracy of carbohydrate-deficient transferrin as a marker of chronic excessive alcohol intake. *Journal of Laboratory Medicine* 2001;25:407-410.
39. Hackler R, Arndt T, Helwig-Rolig A, Kropf J, Steinmetz A, Schaffer JR: Investigation by isoelectric focusing of the initial carbohydrate-deficient transferrin (CDT) and non-CDT transferrin isoform fraction step involved in determination of CDT by the ChronAlcol.D. assay. *Clin Chem* 2000;46:483-492.
40. Helander A, Fors M, Zakrisson B. Study of Axis-shield new %CDT immunoassay for quantification of carbohydrate-deficient transferrin (CDT) in serum. *Alcohol Alcohol* 2001;36:406-412.
41. Chalmers DM, Chanarin I, Macdermott S, Levi AJ. Sex related differences in the hematological effects of excessive alcohol consumption. *J Clin Path* 1980;33:3-7.
42. Sillanaukee P, Seppa K, Lof K, Koivula T. CDT by anion-exchange chromatography followed by RIA as a marker of heavy drinking among men. *Alcohol Clin Exp Res* 1993;17: 230-233.
- 43 Seppa K, Sillanaukee P, Koivula T. Abnormalities of hematological parameters in heavy drinkers and alcoholics. *Alcohol Clin Exp Res* 1992;16:117-121.
44. Newcombe GR, Altman DG. Proportions and their differences. In: Altman DG, Machin D, Bryant TN, Gardner MJ (eds). *Statistics with confidence*. BMJ Books sec edition 2001.
45. Hanley JA, McNeil BJ. A method of comparing the areas under the receiver operating characteristic curves derived from the same cases. *Radiology* 1983;148:839-843.
46. Paris JB. *The uncertain reasoner's companion; a mathematical perspective*. Cambridge University Press. Cambridge 1994.

47. Esteban MD, Morales D. A summary on entropy statistics. *Kybernetica* 1995;31:337-346
48. Bean P, Harasymiw J, Peterson CM, Javors M. Innovative Technologies for the Diagnosis of alcohol abuse and monitoring abstinence. *Alcohol Clin Exp Res* 2001;25:309-316.
49. Conigrave KM, Saunders JB, and Whitfield JB. Diagnostic tests for alcohol consumption. *Alcohol Alcohol* 1995;30:13-26.
50. Salaspuro M. Carbohydrate-deficient transferrin as compared to other markers of alcoholism: a systematic review. *Alcohol* 1999;19:261-271.

Figure 1 Network for the Bayesian Alcoholism Test. The a priori probabilities for diseases and states (left) are combined with the biochemical (right) and clinical findings (under). An arrow going from disease to symptom or biochemical test, indicates that the symptom or test is dependent on the disease or state.



Example of a conditional table for the node ALT i.e. probabilities of values of alanine aminotransferase depending on presence of alcohol use and of liver disease.

No HHAU	No liverdisease	Fatty liver	Hepatitis	Livercirrhosis
ALT not elevated	0,975	0,8	0,5	0,6
50 U/l <ALT <100 U/l	0,025	0,15	0,3	0,3
ALT>100 U/l	0	0,05	0,2	0,1
HAZARDOUS USE				
ALT not elevated	0.9	0.8	0.4	0.5
50 U/l <ALT <100 U/l	0.05	0.15	0.35	0.25
ALT>100 U/l	0.05	0.05	0.25	0.25
HARMFUL USE				
ALT not elevated	0.7	0.5	0.3	0.4
50 U/l <ALT <100 U/l	0.2	0.35	0.4	0.3
ALT>100 U/l	0.1	0.15	0.3	0.3

Table 1. Sample characteristics of 47 controls, 68 non treatment seeking heavy drinkers and 67 treatment seeking alcoholics.

	Controls (n=47)	Heavy users (n=68)	Alcoholics (n=67)
Age	45,3 ±12,7 (24-76)	49.3 ±10,2 (29-80)	43.6 ± 7 (28-58)
Alcohol units/week,	4 ± 6,5 (0-24)	47 ± 22,2 (17-160)	134 ±75 (56-492)
Percentage AUD diagnosis in last year	0%	41,2%	100%
Percentage abstinent in last year	27,3 %	0 %	0 %

Values are mean ± SD and (range)

Age difference between controls and heavy drinkers and controls and harmful users n.s.

Age difference between heavy drinkers and alcoholics, 5.7 significantly different (95% CI 2.7-8.7)

Table 2. Sensitivities, specificities and likelihood ratios of BAT, CDT and GGT for diagnosing harmful and hazardous alcohol use, comparing treatment seeking alcoholics and controls

Diagnostic parameters	BAT (n=114)	CDT(n=110)	GGT(n=114)	Difference and 95% CI for the difference BAT- CDT	Difference and 95% CI for the difference BAT- GGT
Sensitivity	94 (86 – 98)	63.1 (50.9 – 73.8)	73.1 (61.5 – 82.3)	30.8†* (19.0 – 42.5)	20.9* (8.9 – 32.9)
Specificity	97.9 (89 – 100)	93.3 (82.1 – 97.7)	91.5 (80.1 – 96.8)	4.4† (-6 – 15.9)	6.4 (- 2.9 – 17.6)
Likelihood ratio +	44.2 (8.5-249.9)	9.5 (3.5 – 27.9)	8.6 (3.6 – 22.0)	4.7#* (4.5 – 4.8)	5.1#* (5.0-5.3)
Likelihood ratio -	0.06 (0.02-0.14)	0.40 (0.30 – 0.54)	0.29 (0.19-0.43)	0.15#* (0.14-0.17)	0.21#* (0.19- 0.23)

* significant difference at p level < 0.05

Ratio of compared Likelihood Ratios with 95% Confidence interval

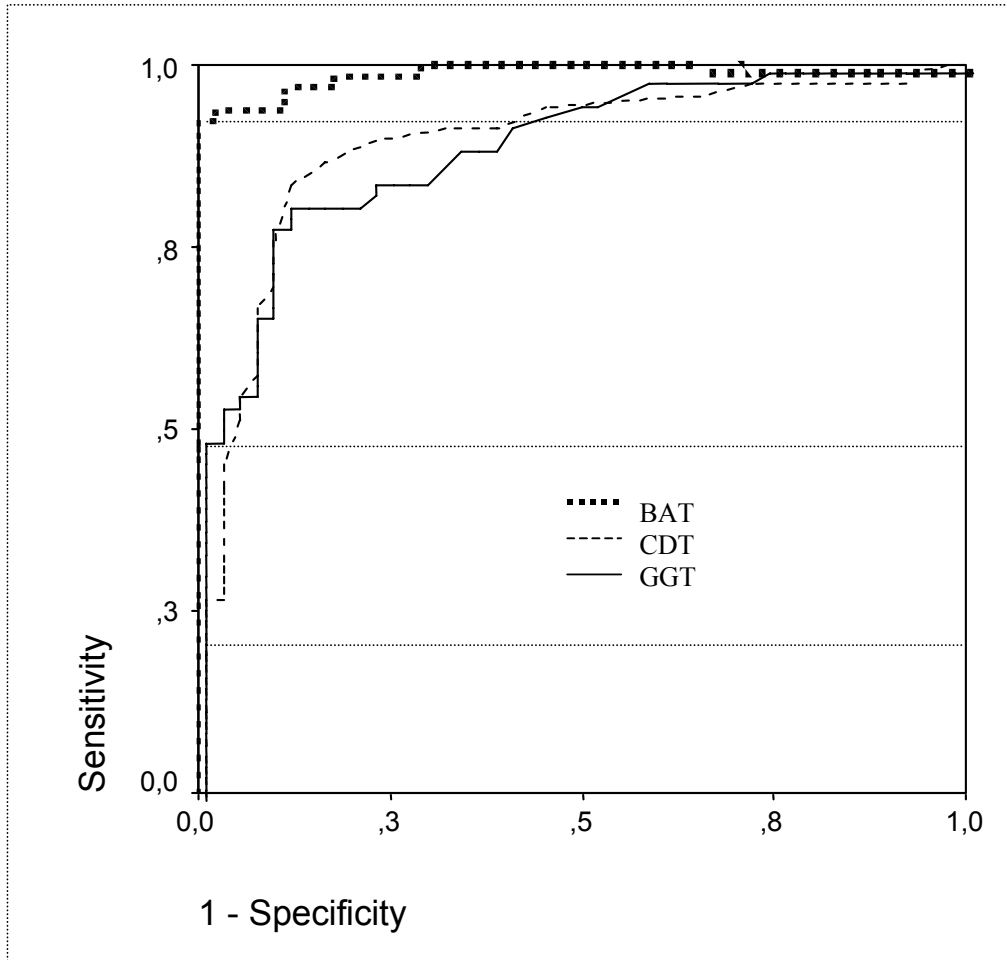
† The difference is calculated without missing values (n=110)

Table 3. Distribution of positive test results of BAT, CDT and GGT over the three subgroups of 68 heavy drinkers: a. Non-hazardous use (<280g/week), b. Hazardous use (280-560g alcohol/week), c. Harmful use (>560 g alcohol/week)

Heavy users	BAT+	CDT+	GGT+
(n= 68)*	(n=23) 100%	(n=23) 100%	(n=17) 100%
a.<280g/week (n=11)	(n=0) 0%	(n=1) 4.3%	(n=2) 11.8%
b.280-560g/week (n= 38)	(n= 11) 47.8%	(n=12) 52.2%	(n=9) 52.9%
c. >560g/week (n=19)	(n= 12) 52.2%	(n=10) 43.5%	(n=6) 35.3%
Entropy	0.6921	0.8367	0.9566
(95% CI)	(0.6723 - 0.7130)	(0.6334-1)	(0.7321-1)

*3 missing values of CDT

Figure 2 Receiver operating characteristic curves, comparing 46 controls and 64 alcoholics. Criterium is harmful use of alcohol (>560 g alcohol/week)



	BAT	CDT	GGT
Area under the ROC curve (95% CI)	0,989 (0,975-1,000)	0,909* (0,852-0,966)	0,902* (0,847-0,957)

Area under the roc curves are given in percentages with confidence intervals. Comparison was done with 110 subjects because of 4 missing values of CDT

* p<0,005