Confirming alcoholism in drivers under influence

Korzec, A.


Confirming
alcoholism
in drivers
N 247
N 247
under
influence

A. Korzec

CONFIRMING ALCOHOLISM IN DRIVERS UNDER INFLUENCE

A. Korzec

The studies described in this thesis were performed at the psychiatric department of the Sint Lucas Andreas Hospital, Amsterdam, in collaboration with the Amsterdam Institute of Addiction Research.

Sint Lucas Andreas Hospital, Bayer Diagnostics, Pharmacia & Upjohn, Axis and the Dutch Traffic Test Organization, Disqualification Division supported these studies.
Geert Franssen designed the cover; Mai Spijkers, Alex Schijffelen and Marieke Oele gave support with the publication.

Uitgeverij Prometheus, Amsterdam
Druk Wöhrmann Print Service, Zutphen

CONFIRMING ALCOHOLISM IN DRIVERS UNDER INFLUENCE

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
Prof. Mr. P.F. van der Heijden
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de Aula der Universiteit

op woensdag 16 juni 2004, te 14.00 uur

door

Aleksander Korzec
geboren te Lodz, Polen

Promotor : Prof. Dr. W. van den Brink
Co-promotor : Dr. M.A.W. Koeter

Faculteit Geneeskunde

Dankwoord

Het zuiver verwoorden van eenvoudige dankbaarheid is aanmerkelijk delicater dan het lijkt (1). Nietzsche beschrijft hoe verbluffend onhandig velen zijn in het uitdrukken van hun dankbaarheid. Soms lijkt het wel of degene op wie invloed is uitgeoefend en die daarvoor dankbaarheid verschuldigd is, zich uiteindelijk en welbeschouwd beledigd voelt. Zo iemand, die heimelijk vreest dat zijn zelfstandigheid bedreigd wordt wanneer hij referenties prijsgeeft, kan zijn dank slechts uiten in onderhuidse onhebbelijkheden (2).

Dankbaarheid bevat vaak sporen van ongemeende of overdreven lof, malicieuze wraak, de vereffening van openstaande rekeningen, het alsnog opeisen van bepaalde aanspraken en ambivalentie over eigenwaarde (3). De oorzaken hiervan zijn al lang duidelijk: mensen die een gunst bewijzen houden meer van degenen aan wie ze de gunst bewijzen, dan die begunstigden houden van hun weldoeners. Maar er bestaan volgens dezelfde Nietzsche domeinen van dankbaarheid die zuiver en evenwichtig zijn: dankbaarheid voor datgene wat je geleerd hebt, schuldvereffening tussen gelijken en dankbaarheid horend bij een afronding.

Van mijn moeder Regina heb ik geleerd, wanneer ik weer eens bij haar klagen kwam, dat het nog altijd slechter kan. Van haar zuster Katia dat je er ook iets aan kunt doen om het beter te laten gaan. Mijn vader leerde me snel op te staan als je valt. Van mijn vier kinderen, Mischa, Sanne, Sascha en Itha, heb ik veel geleerd over zuivere genegenheid. Mijn vrouw, Mecheline van der Linden, heeft mij veel over de liefde bijgebracht.

Veel heb ik opgestoken van mijn enthousiasmerende promotor, Wim van den Brink. Als belangrijkste misschien zijn visie dat de wetenschap begint met het opruimen van vooroordelen. Verder, dat de juiste manier om naar een halfvolle fles te kijken is de inhoud te benadrukken, zonder uit het oog te verliezen dat hij voor de helft leeg is; bovendien, er zijn andere flessen. Daarnaast heeft hij me geleerd dat een onderzoeker wellevend moet zijn en e-mails snel moet beantwoorden.

Van mijn co-promotor Maarten Koeter leerde ik: als je iets niet begrijpt is dat onvoldoende reden je goede humeur te verliezen.

Vaak helpt het de dingen een paar keer door te nemen en aan te horen, alvorens ze te begrijpen.

Van Carla de Bruijn, een wiskundige die in de psychiatrie verzeild is geraakt, leerde ik complexe problemen te vereenvoudigen tot onderzoekbare vraagstellingen. Van een andere wiskundige, Michiel van Lambalgen, heb ik geleerd dat problemen menigvuldig zijn; sommige zijn een vraag, andere zijn een bevel, weer andere zijn nonsens, niet alle zijn op te lossen door rekenen.

Van mijn vrienden Pieter Coen Blom, Wim Brinkman, Theo van Gogh, Frank Koerselman en van mijn broer Michel Korzec, heb ik geleerd dat onthechting, verstandige compromissen en onwankelbaar geloof in wat je doet, niet voor mij zijn weggelegd.

Van een aantal patiënten heb ik geleerd dat het onvoldoende is om interessante veronderstellingen te hebben maar dat je ze ook overtuigend moet bewijzen. Dat duurde een tijdje voor ik zover was.

Meer dan vierhonderd mensen hebben aan dit proefschrift hun medewerking gegeven. Ik heb niet onderzocht hoeveel daarvan nu spijt hebben. Ik noem hier slechts enkele groepen van al diegenen die mij geholpen hebben.

Vele patiënten, wijndrinkers en aangehouden automobilisten hebben zonder eigenbelang bloed en tijd geschonken. Directie en bestuur van het Sint Lucas Andreas Ziekenhuis hebben mij de ruimte en de middelen verschaft onderzoek te doen. Medewerkers van diverse laboratoria, met name Wim de Kievit, Torsten Arndt en Jenneke Weijers hebben me veel uitgelegd over de betekenis van getallen. Ruud Bredewoud en Prem Heera hebben me duidelijk gemaakt hoe te handelen als arts bij de afweging tussen het belang van verkeersveiligheid en het individuele belang.

Mijn collegae Frank Koerselman, Wybrand Op den Velde en Ton Vergouwen waren eerst sceptisch, en zijn daarna solidair gebleken. De assistenten in opleiding hebben me opnieuw geïnspireerd voor mijn vak.

Marianne van Weegen, Lucia Reiding, Adri Wijbenga, Ruud Smit, Mieke Commandeur, Hans Moinat en Bep de Lange hebben een belangrijke bijdrage geleverd aan het behouden van mijn goede humeur tijdens dit onderzoek. De verpleegkundigen van de afdeling psychiatrie, die hun gewicht in goud waard zijn en het zware

werk doen, hebben mij altijd geinspireerd met hun voortvarend optreden.

Tot slot een belangrijk vraagstuk: wat gebeurt met dankbaarheid na de voltooing van een bezielende en bijna verslavende bezigheid en de daarbij horende plechtigheid? Bij velen is, vreemd genoeg, de half-waarde tijd van dankbare gevoelens aanmerkelijk korter dan van bijvoorbeeld beledigde gevoelens. Er zijn religies en psychotherapeutische scholen gesticht op het idee dat het nuttig is om elke dag de aandacht te richten op dankbare gevoelens. Dat is moeilijk maar er bestaat onderzoek dat erop wijst dat mensen die daarin slagen, een betere stemming krijgen, onafhankelijk van giften en prettige gebeurtenissen (4). Een alternatief is je te richten op andere uitdagingen.

1. Wiggins S. Construction and Action in Food Evaluation. Journal of Language and Social Psychology 2001; 20:445-463
2. F. Nietzsche. De vrolijke wetenschap (100): leren huldigen. De Arbeiderspers, Amsterdam, 1976
3. McWilliams N, Lependorf S. Narcissistic pathology of everyday life: the denial of remorse and gratitude. Contemporary Psychoanalysis 1990: 26:430-451
4. McCullough ME et al. Gratitude in intermediate affective terrain: links of grateful moods to individual differences and daily emotional experience. Journal of Personality and Social Psychology 2004; 2:295-309

Contents

Chapter 1

INTRODUCTION

Medical diagnostics is classification with an aim[1]. In health care the aim is therapy and prognosis for an individual patient. This aim structures diagnostic activities: if therapy and prognosis for two different alcohol disorders are the same, the clinician feels no obligation to put a lot of effort in distinguishing the two disorders. However, the clinician can have reasons to differentiate within well-accepted diagnostic categories. There is for example a hypothetical rationale to distinguish different types of craving in alcohol dependent patients. Patients with reward craving are thought to benefit more from naltrexon while patients with relief craving are thought to benefit more from acamprosate (1,2).

In a general sense the studies in this dissertation were inspired by an interest in diagnostic reasoning, and its end product: clinical diagnosis. More precise, the aim of the studies is to examine the rationale of the clinical diagnosis of alcoholism in forensic situations.

Understanding the difference between health care diagnostics and forensic diagnostics is essential to understand some of the decisions made in the course of the studies.

In many health care situations the diagnosis of alcoholism is relatively easy. When the patient crosses the door of an outpatient alcoholism treatment center and tells about his alcohol problems there is no need for much clinical reasoning. Problems with alcohol are the core of the psychiatric diagnosis of alcoholism. The physician has only to check if the criteria for an alcohol use disorder (AUD) are met to obtain diagnosis; severity of AUD and motivation of the patient do not change diagnosis but are examined to estimate progno-

[1] Strictly speaking one can differentiate the theoretical classification of syndromes as an activity different from clinical diagnostics which is the assignment of one specific category from the classification system to one individual patient.

sis and to decide about therapy. In other diagnostic situations, such as diagnosing alcoholism in Drivers Under Influence (DUI's), obtaining a diagnosis requires more effort.

The diagnostic examination of alcoholism in DUI's, described in the third chapter of this dissertation, takes place in a legal setting. According to Dutch regulations on driving ability, a selected part of Drivers Under Influence (DUI' s) are mandatory examined by a clinician (3). Offenders are informed that they will lose their license in case of non-cooperation with the examination. Diagnostic procedures in this context are part of an administrative legal procedure to evaluate whether the subject has the right to have a driving license. Under Dutch law, it is demanded that the subject has refrained from alcohol misuse for the last 12 months. In cases where alcoholism is diagnosed, the license is withdrawn.

The clinician gives the diagnosis in a legal context, which is quite another context than the usual clinical one. Usually the clinician's goal is the patient's best interest, but in the specific legal situation, traffic safety is the point of reference. The last-mentioned does not always converge with the subject's best interest. Losing a driver's license can have great consequences for one's job and social status; many DUI's feel that their drinking habits are not severe enough to warrant a medical diagnosis. Furthermore the diagnosis, based on clinical arguments, may have to be defended in court. The focus of the diagnostic aim in forensic evaluation of DUI's is not therapeutic but prognostic. Specifically the aim is to estimate the probability of relapse DUI.

In health care the diagnostic process is dynamic. If therapy is not successful, the diagnostic process can be re-evaluated in order to check if there is an error in clinical reasoning or whether new data change the diagnosis. In forensic diagnostics clinical reasoning and arguments can not be re-evaluated after the decision that a DUI is an alcoholic. The clinician gets one chance instead of many longitudinal data on the course of the disease. Thereafter he must be sure enough of the arguments to justify his diagnostic conclusion.

So how does the legal context influence diagnostic reasoning and diagnostic decisions? An example of a case can clarify this. The clinician made an alcoholism diagnosis that had to be defended subsequently in court. The clinical considerations are given in brackets.

A 30 year old single living man, employed as accountant, has been arrested on a weekend night at 1 AM after drinking, starting at 8 PM, 6 vodka on a party at his sister's birthday and afterwards 4 beers in a bar. (Arrests in daytime and during the week are more suspect for alcoholism than nightly arrests in the weekend). The Blood Alcohol Level (BAL) was 2,0 ‰. (As each standard drink results in approximately 0,2‰ and one standard drink is eliminated in 1,5 hour, the reported intake is probably underestimated: (10 x 0,2) - 5/1,5 = 1,67‰). He has driven 10 km before being arrested (It is not easy for a non-alcoholic to drive 10 km with a BAL of 2,0 ‰)

It is his first DUI arrest. In the medical examination several months after the arrest, the subject denies present or past social, psychological or physical problems due to alcohol. He smokes but uses no drugs. (Almost all alcoholics are smokers but the converse relation does not hold). There is an elevated blood pressure: 170/105 mm Hg. (Alcoholism is a frequent cause of high blood pressure in young subjects, but the converse relation does not hold). There are no other physical signs concurrent with alcoholism. (Physical signs are uncommon in young alcoholics and are mostly seen in late stage alcoholism). He states that he did not feel intoxicated on the night of the arrest. (It is uncommon that a non-alcoholic subject does not feel intoxicated after 10 drinks; the specificity of this clinical sign is unknown). The subject says that after the arrest his drinking habits have changed from 5 alcoholic drinks each day, to one glass of wine with diner and 3 other drinks twice a week. Blood examination reveals a slightly elevated aspartate amino transferase (AST) and an elevated Gamma- glutamyltransferase (GGT) value, twice above the cut off level (AST and GGT are often elevated in alcoholism)

The clinician makes the diagnosis of alcoholism, after which the subjects' driving license is withdrawn. The subject challenges this decision in court and brings to court a medical counter-expertise. This counter-expertise states that the subject used an above average amount of alcohol but is not an alcoholic. Furthermore it states that the slightly elevated AST is without value for diagnosing alcoholism, that there is another possible reason for the elevated GGT, (the subject uses paracetamol for headache). It also states that the elevated blood pressure is essential and has no relation with the sub-

jects alcohol use, and the fact that he did not feel intoxicated after 10 drinks is not a proof of alcoholism. The expert witnesses for both parties were not able to put probability numbers to the different arguments. In the end the court decided that the clinician had enough reasons to diagnose alcoholism because of the simultaneous occurrence of different symptoms.

The question is: how valid is this decision? What are the chances that both the first clinician and the judge were wrong?

The question that the Dutch Traffic Test Organization, Disqualification Division, asks from the clinician is: What is the psychiatric diagnosis, based on clinical relevant signs and/or DSM criteria? In a subtext this is explained as: is there alcoholism in the broadest sense? Implicit in this question is the assumption that there are clinical signs that are not mentioned as DSM AUD criteria, but make an AUD diagnosis more probable.

It is important to realize that the core business of the Dutch Traffic Test Organization, Disqualification Division is not to diagnose disease but to diagnose impairments that can endanger traffic safety (3). For example, some subjects with schizophrenia with minor pathology are estimated not to be impaired for driving, but other schizophrenic patients with paranoid delusions about other car drivers are judged to be impaired. Essentially, the diagnostic evaluation is about whether somebody is unfit to safely drive a car.

Originally, the question put to the clinician, was not diagnostic but prognostic with regard to traffic safety: Is there an elevated probability of repetition of driving under influence of alcohol? The diagnosis of alcoholism was part of an overall evaluation on relapse DUI- risk, under the assumption that an addicted subject will have an enhanced probability of relapse in driving under influence because of loss of control over drinking behavior. However, physicians felt that this was not a question they were competent to answer. So in 1993, in a meeting between representatives of the Dutch Royal Medical Association, The Dutch Psychiatrist Association and The Dutch Traffic Test Organization it was agreed to change the question in: Is there a diagnosis of alcoholism? If there is an alcoholism diagnosis, but in remission, the question runs: has enough time passed to predict that the subject will not relapse in alcoholism? Because the examination takes place in a legal setting, the

Dutch Traffic Test Organization has made great efforts to standardize the clinical examination in order to enhance inter-clinician reliability.

The specific legal situation and the primary goal to enhance traffic safety lead to several conceptual, epidemiological and clinical questions:

1. How to define alcoholism?
2. What is the prevalence of alcoholism in a DUI population?
3. Which clinical arguments are used for the diagnosis of alcoholism and how valid are these arguments?
4. What is the value of the diagnostic tests used for the diagnosis of alcoholism in a DUI population?
5. Is it possible to design a diagnostic tool that, by combining probabilities of relationship between elevated biochemical markers and clinical signs, enhances the diagnostic ability to confirm whether a subject regularly uses a hazardous amount of alcohol?
6. Does such a diagnostic tool work in a real forensic situation where DUI's are examined for alcoholism?

The studies in this dissertation concern these complex questions. The questions are discussed in this introduction. In the conclusion section some answers are given. More importantly, a method is suggested how to confirm the diagnosis of alcoholism in the context of traffic safety.

1. How to define alcoholism

Let us start with a stipulative definition of alcoholism, which distinguishes alcoholism from "social drinking".

Alcoholism refers to a heterogeneous set of disorders. Two overlapping conceptual frameworks are used to approach this set of disorders. The first approach comprises the psychiatric diagnoses alcohol dependence and alcohol abuse (Alcohol Use Disorders: AUD), and emphasizes loss of control and alcohol related social, psychological and physical consequences. The second approach comprises unhealthy drinking patterns, emphasizes their effects

on physical health, and is often referred as hazardous alcohol use (HAU) (4,5).

By using the term alcoholism, clinicians mean either an AUD diagnosis ór a HAU diagnosis, and mostly both at the same time. But it is questionable whether this last use is correct.

We started our study with a literature search about alcoholism in DUI populations and were confronted with many different conceptualizations of alcoholism. Most frequently the studies referred to unspecified populations of "heavy drinkers". It was not simple to generalize the results or translate them for populations defined with modern definitions of alcoholism.

The following eight diagnostic terms were used, often with vague definitions: alcohol dependence, alcohol abuse, hazardous use, harmful use, alcohol misuse, excessive drinking, heavy drinking, problematic drinking. Vague definitions often reflect vague ideas.

Alcohol dependence is an AUD diagnosis that is well described with relatively reliable operational and almost identical criteria in the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) and the ICD-10 Classification of Mental and Behavioral Disorders (ICD-10) (6,7).

Alcohol abuse is another well described AUD in DSM-IV, but its validity is questionable as indicated by its' low temporal stability and its' weak temporal relationship to alcohol dependence (only about 10% with alcohol abuse will become alcohol dependent). Also the fact that alcohol abuse is not associated with other forms of psychopathology and finally the fact that most people with alcohol abuse have the diagnosis because they were using alcohol while driving.

Hazardous use is a HAU diagnosis that is defined by drinking an amount of alcohol that bears a risk to health (4,5).

To complicate matters, *harmful use* is both an AUD diagnosis within ICD 10, as well as a HAU diagnosis defined by drinking an amount of alcohol that is a high risk to health (5).

The terms *alcohol misuse, excessive drinking* and *heavy drinking* have a moralizing connotation and have no informative value above the terms used in AUD and HAU diagnoses.

The term *problematic drinking* is too vague as it generally does not refer to specific alcohol related problems.

It seems unavoidable to conclude that if one wants to achieve precision, one has to define alcoholism as either an AUD diagnosis, or as a HAU diagnosis, or infer AUD diagnosis from signs of HAU, as is implicit assumed in the above mentioned question from the Dutch Traffic Test Organization. But can one infer AUD diagnosis from clinical signs indicating HAU?

There exist a vast amount of research about the relationship between ethanol intake and AUD diagnosis (8,9). It has been shown that drinking parameters like frequency of drinking, the frequency of drinking more than 5 Alcohol units/day on any one occasion and the frequency of being intoxicated increase the risk of AUD diagnosis. However, there is no research of the relationship between biochemical markers of hazardous alcohol use and AUD diagnosis. In order to examine whether it is possible to infer AUD diagnoses from HAU diagnoses we examined whether subjects with AUD diagnoses had heavier drinking patterns and more biological damage than subjects without AUD (measured in clinical and biochemical signs). As it can be assumed that the hard core alcoholics in a DUI population represent only a minority while hazardous drinkers are more frequently represented, (10) we studied this question in a population of well-functioning hazardous drinkers. This study is described in chapter two. This chapter is an abridged version of an earlier study about the discriminant validity of Alcohol Use Disorders from a different perspective (11).

Our results converged with other research regarding the low validity of the diagnosis alcohol abuse according to DSM-IV. More importantly in the context of our dissertation, we found that in our population of hazardous drinkers one could not assume that the more severe hazardous drinkers have significantly more often an AUD diagnosis.

There is another reason to question the choice of defining alcoholism as AUD in DUI populations. A specific problem in diagnosing DUI's with questions whether AUD criteria are met is the high denial rate in this population (12,13).

On first sight there is an easy solution for this problem: in order to improve traffic safety one could broaden the usual AUD definitions of alcoholism and diagnose all dubious cases as such, which seems

to be the line of reasoning of the Dutch Traffic Test Organization. But what are the consequences of such a decision?

In health care such a decision is easily defended: if one can prove that some pre-clinical alcoholism states can be treated more successfully than full blown alcoholism, a physician can defend the cost of treatment to the health insurance that pays for them. In the legal context, however, it is not as easy as that. The problem is that in the environment of administrative law the rules are different. In a legal setting one has to back up such a decision either with references to international conventions (such as ICD-10 and DSM-IV, or definitions of HAU), or with scientific arguments such as a proven relationship between specific operational diagnostic concepts and an elevated risk of relapsing in DUI behavior. Such scientific studies are indeed available (10). Alcoholics and excessive drinkers as a population are involved in significantly more collisions and driving under influence when compared to nonalcoholic drivers or the general driving population.

In conclusion one has to choose. As one cannot infer AUD from HAU, and because AUD diagnoses are dependent on the cooperation of the subject, which is questionable in DUI's, the most logical choice is to decide for HAU diagnosis as alcoholism definition in the context of traffic safety.

This conclusion, however, leads to new problems, which are discussed in the next paragraphs.

2. What is the prevalence of alcoholism in a DUI population?

The clinical value of a test is dependent of the prevalence of a specific disease in a specific population. Usually the clinician knows the test parameters, generally described as sensitivity (the probability that the test is positive if the disease is present) and specificity (the probability that the test is negative when the disease is absent). However, that is not what the clinician wants to know. For the clinician, the real value of a test is best described with the positive predictive value (PPV) of a test which is defined as the probability that the disease is present if the test is positive and the negative predictive value (NPV) which is defined that the disease is absent if the

test is negative. As PPV and NPV are dependent on the prevalence of the disease in the target population in which one uses the test one must not only know the sensitivity and specificity values but also the prevalence in the population of which the patient is a member.

With Bayes theorem one can calculate PPV and NPV from sensitivity, specificity and prevalence. Bayes theorem is a mathematical procedure how an 'a priori' probability should be revised on the basis of new information. Bayes theorem allows the exchange of the order of cause and effect. The problem addressed and solved by Bayes had to do with gambling and with inverse probabilities. It is now widely used in diagnostics: If the probability of the occurrence of an event (e.g. an elevated biochemical marker), given the presence of the cause (e.g. the disease), is known, it is possible to calculate the inverse probability that the cause (disease) was present, given the occurrence of the event (elevated biochemical marker). Translating this idea in a mathematical language: If A is the cause and B is the event, Bayes' theorem allows us to calculate the probability of *A given B, P (A|B),* if we know the probability of *B given A, P (B|A),* and the probabilities of each event alone P (A) and P (B). Bayes theorem is algebraically a simple equation.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

This theorem is applied as a method to investigate the clinical value of diagnostic tests.

In diagnostic terms Bayes' theorem concerns the value of a diagnostic test B for the presence of an illness A. Alcohol is toxic and alcoholism (A) causes a change in many bodily functions which can be used as diagnostic tests (B). The value can be formulated as the probability that the illness A is present if a diagnostic test B is positive. This probability, P (A | B), is determined by three other probabilities:

a) The inverse probability that the test is positive if the illness is present, P (B | A). This is named the sensitivity of a test and can be obtained by testing large samples of subjects with the disease.
b) The probability that the test is positive without any conditional probability, P (B). The probability that the test is positive

(meaning above a chosen cut-off value), without any conditional probability, can be obtained by testing large samples of the target-population.

c) The probability that the illness is present, P (A), without any conditional probability, also named the prevalence of the disease in the target population.

An example how prevalence influences the clinical value of a diagnostic test: The positive predictive value of an elevated GGT or AST for alcoholism, (see the example at the beginning of this introduction), changes if it is applied in populations with a different prevalence. According to Bayes' theorem, if the prevalence of alcoholism subjects in the DUI population is 10%, the probability that alcoholism is present, *if there is an elevated AST,* is approximately only 18% and *if there is an elevated GGT* only 15%. However, if the prevalence is 60%, the probability that alcoholism is present *if AST is elevated* rises to 75% and *if GGT is elevated* to 70%. (For calculation see Appendix at the end of the introduction).

Often, the only unknown value in Bayes' equation is the probability that the illness is present in the target population. This is the reason that all diagnostic research must start with the question: what is the prevalence?

Research suggests a considerable prevalence of alcoholism in DUI populations. In a review of reports on the prevalence of alcoholism in DUI populations up to 1986, Vingilis estimated the prevalence between 25 and 50%, depending on the sampling of the population and the criteria used for alcoholism (10). However, because countries differ in alcohol use patterns and because changes have occurred in alcohol use and driving under influence in many countries in the last 15 years, this estimate is not likely to be accurate anymore. In order to obtain new prevalence estimates, two different strategies can be used:

1. On the individual level by diagnosing each subject in the population with a specific diagnostic procedure. If the diagnostic procedure is not perfect one can combine different diagnostic procedures parallel or sequentially. The resulting prevalence rate is the percentage of positively diagnosed subjects in the population.

2. On the population level by assuming that a given variable is present in a certain percentage of the disease. In this case, the population based prevalence estimate of hazardous use is computed with the following formula derived from Bayes theorem: $P = [T - (1-Sp)] / (S + Sp - 1)$ (14). (P = prevalence; T = proportion of positive tests (CDT or GGT); S = sensitivity; Sp = specificity).

In the second study, chapter three, we will use both methods to obtain a prevalence-estimate of alcoholics in a DUI-population. The prevalence estimate based on individual diagnosis was obtained by applying four different diagnostic procedures to a DUI-population. The population-based prevalence estimate was obtained by using sensitivity and specificity data of either an elevated CDT or GGT value from two earlier studies: a study of alcoholics and a study of heavy drinkers.

3. **Which clinical arguments are used for the diagnosis of alcoholism and how valid are these arguments?**

When making a diagnosis of alcoholism a clinician makes a distinction between "hard signs" (high specificity) and "soft signs" (low specificity). For example both a CAGE above 2 (15) and an elevated Mean Corpuscular Volume of red blood cells (MCV) without anemia are considered as "hard signs" of alcoholism. In contrast, a slightly elevated GGT or an elevated blood pressure are considered as "soft signs" for alcoholism as many medical conditions can cause it. In clinical reasoning, these signs are interpreted as either increasing or diminishing the probability of a positive diagnosis. Clinical arguments are subjective probabilities based on clinical experience and reading research. Of some clinical signs and biochemical markers this subjective knowledge is corroborated by scientific research but for many signs and situations the scientific knowledge is insufficient or non-existent. Many clinical arguments are implicit.

One of the main goals was to develop a diagnostic confirmation test, taking into account the context of medical examination of alcoholism in a legal situation where a false diagnosis could have grave consequences. The question is: could we build a diagnostic instru-

ment based on the above mentioned clinical reasoning that could serve as a confirmation test of alcoholism?

In chapter three, in the study on the prevalence, we propose and test a diagnostic system called the restrictive diagnostic procedure (RDP), i.e. an algorithm with explicit clinical arguments and elimination of all "soft signs".

4. What is the value of the diagnostic tests used for the diagnosis of alcoholism in a DUI population?

One of the possibilities to make clinical arguments more robust is to replace one's clinical intuition by sensitivity and specificity values of the different biochemical markers and clinical signs. Unfortunately, sensitivity and specificity values are not as robust as they seem, due to the so-called spectrum effect. Studies have shown that sensitivity and specificity of markers of excessive alcohol use depend on the distribution of severe and mild cases of alcoholism in the study sample (16). A high ratio of severe/mild cases increases sensitivity, while a low ratio lowers sensitivity. On the other hand when studying specificity of alcoholism in non-alcoholic churchgoing Baptists who for the last 10 years abstained from alcohol the specificity is probably much better than in an average non alcoholic population drinking 9 alcoholic drinks a week (17).

Spectrum differences can cause a great variation of sensitivity values. For example the sensitivity values for AST for detecting HAU vary from 10-30% in a hazardous users population to 35-50% in alcoholics admitted in a detoxification center. For GGT these values are respectively 20-50% and 60-90% (16). The specificity values for AST vary less and are all above 90%, whereas for GGT the range is enormous (55-100%) (16).

In conclusion, if one wants to know the validity and usefulness of a test one has to make an empirically based assumption regarding the prevalence, sensitivity, specificity and regarding the spectrum of the target population. The findings of the prevalence from chapter three are used in chapter six. Empirical assumptions regarding the sensitivity, and specificity and spectrum are used in the development of the Bayesian Alcoholism Test (BAT) in chapter five.

Another important issue concerns the most promising diagnostic marker for HAU: Carbohydrate Deficient Transferrin (CDT). The problem here is that at the time of our studies there were many different CDT tests. One of the first commercial CDT tests, CDTect®, measures asialo-, monosialo- and disialo-isoforms of CDT but does not correct for total transferrin in plasma. In order to correct for potential differences in total plasma transferrin, the original test was replaced by another CDT test, %CDTri-TIA, that did correct for total transferrin and thereby improved the specificity of CDT. However this test also included trisialo-isoforms of CDT.

The question is which CDT test has the best diagnostic properties for detecting HAU. In the study described in chapter four we compare the diagnostic accuracy of two % CDT-tests: one including asialo-, monosialo-, disialo- and trisialo-isoforms and one without the trisialo-isoforms. Both tests correct for total transferrin in plasma.

5. Is it possible to design a diagnostic tool that, by combining probabilities of the relationships between elevated biochemical markers and clinical signs, enhances the diagnostic ability to confirm whether a subject regularly uses a hazardous amount of alcohol?

A logical step to enhance the moderate values of single diagnostic tests of hazardous alcohol use is to use different tests at the same time.

There have been many attempts to use combinations of two or more biochemical markers (17-20) or combinations of biochemical markers and clinical signs (21) to identify hazardous alcohol use. However, the proposals have found little application because of the following problems:

a. There was not sufficient improvement of sensitivity and specificity.
b. The combination of tests was too complicated to be applicable in common clinical practice.
c. The combination of tests was not applicable in legal and insurance settings because it resulted in too many false positives.

The question is whether these problems can be solved. In order to find an answer to this question, we first have to discuss the different meanings of the term 'probability' (22). In order to do so, we must distinguish between *objective* and *subjective* probability.

Objective probability (also called *frequentistic* probability) refers to a given population *S*, and can be defined as follows: the probability $P_S(A)$ of the occurrence of disease A in *S* is the number of subjects with disease A in *S*, divided by the total number of subjects in *S*. (This number is also called the relative frequency of A in *S* or the prevalence of A in S).

Subjective probability is not referring to a population, but to the knowledge *K* about a specific individual regarding the presence of an illness or a particular event. In this case $P_K(A)$ is a number between 0 and 1 (endpoints included) which expresses the physician's *degree of the belief* that a given subject has a disease A (23).

These two notions of probability are very different, even though in practice they often have to be combined. The good news about objective probability is that it is easily measured. The bad news is that it is not so clear what the numbers thus obtained mean when applied to an individual patient. Somewhat metaphorically one may say that for an individual patient (for example the 30-year old man from our example at the beginning of the introduction), the probability that this patient has disease A is the aforementioned $P_S(A)$. Here, one uses objective probabilities to guide one's subjective probabilities. Based on the fact that he is a member of the DUI population, the objective probability would be somewhere around 50 % (see chapter three). But in reality the patient belongs to many different populations, that of single 30 year old men, of first DUI arrests, of smokers, of accountants, of subjects with hypertension, *etc.*, and the probability of A may be different in each of these populations. This is one of the reasons why subjective probability is more useful for the clinician, because this is by definition relevant to single cases. Subjective probability is however beset by problems of its own.

First, it seldom happens that a medical expert's probability estimates are *consistent* in the following sense. If P (A) is the probability that a subject has a certain disease, e.g. alcoholism, and P (B) is the probability that this subject does not have this disease, then P

(A) + P (B) =1 (the subject either does or does not have the disease). However, due to clinical knowledge that some people have "a small amount" of disease, the medical expert's probability estimates often add up to more or less than one (e.g. the medical expert estimates the probability that the subjects suffers from the disease as p=0.20 and the probability that the subject does not have the disease as p=0.90). If the subjective probabilities do not satisfy this basic law of probability, the computations yield meaningless results.

Second, it has been shown that subjective estimation of probability, when one has to keep track of many different probabilities (such as membership of many different populations), is often counterintuitive and therefore results in wrong estimates (24).

Third, it is not clear how subjective probabilities relate to objective probabilities, when known. Clearly an uninformed subjective estimate can differ vastly from the true relative frequency in a population. Weather-forecasters give generally reasonable estimates; physicians, alas, do not, which brings us to the next topic: expert systems.

An expert system is a computer program that codifies existing general knowledge about a domain, (in our study alcoholism), in such a way that feeding in data about a particular patient (e.g. values of selected blood markers) may yield a valid diagnostic probability for the patient to suffer from alcoholism. Expert systems are useful when there are a large number of diagnostic tests for a given disease, and when the relationship between the disease to be diagnosed and the result of tests for the disease is of a probabilistic nature.

Although the probabilistic computations involved are complex, mathematical and computational technology has now progressed so far as to make an expert system of the size necessary for the description of alcoholism entirely feasible. This progress has resulted in so-called Bayesian networks with graphical structures in which the nodes represent diseases, syndromes, patho-physiological entities, symptoms, and diagnostic tests, and where an arrow going from node n to node m indicates that the probability of m (causally) depends on the probability of n (See chapter 5, figure 1). The objective probabilities involved can be obtained from epidemiological data or can be elicited from experts. To avoid misunderstanding, the

issue here is not objective or subjective information but objective and subjective probabilities. It is an important assumption that the probabilistic information in these conditional probability tables can always be combined consistently, i.e. sum of to 100% and not more.

The two kinds of information, graphical and probabilistic, are sufficient to answer queries of the following type: Given values obtained for some, but not necessarily all, diagnostic tests, what is the probability that a given patient suffers from a particular disease? It is important here that the results of many tests may be combined; this is in contrast to the vast literature of diagnostic tests, where mostly single tests are considered (25). The predictions made by a Bayesian network depend entirely on the graphical structure and the associated conditional probability tables. This is still about objective probabilities. However, sometimes these probabilities are based on objective/empirical information and sometimes on subjective estimates made by experts.

The question now is, can such a Bayesian Alcoholism Test be developed and if so, how does it work with populations of patients of which the diagnosis is known?

The most important contribution of this dissertation is the development of a Bayesian expert system for diagnosing HAU. In chapter five, the development of this Bayesian Alcoholism Test (BAT) is described and compared to single diagnostic tests in a population of known alcoholics, heavy drinkers and non-alcoholic controls.

6. Does a Bayesian network diagnostic tool work in a real forensic situation where DUI's are examined for alcoholism?

Feinstein suggests that a diagnostic tool must be validated like the validation of the therapeutic value of drugs in different phases (26).

He writes: *"In phase I, the test would be compared for cases of substantially diseased people and for healthy controls. If good discrimination is shown in Phase I, the test can advance to Phase II, in which the spectrum of comparison is extended. The test would be now challenged with different types of diseased cases and controls, covering a suitably wide spectrum of disease and health. If discriminating remains good, the challenge spectrum would be en-*

larged in Phase III so that the selected cases and controls encompass the clinical, co-morbid and pathological issues... If the test passes the challenges of Phase III, the architecture of Phase IV can become prospective rather than case-control. The results of the marker would be noted, reported and analyzed for a large consecutive series of clinically suitable patients. If definitive standard results are not known for many of these patients, their data would be analyzed separately, using alternative standards".

In our first study on BAT, described in chapter five, phase I and II of the validation process is conducted. Phase III remains to be tested because no data could be collected among internal medicine patients of our hospital. Since our primary goal was to enhance the diagnostic process in the context of DUI's, we conducted a phase IV study and compared BAT to conventional methods used for diagnosing alcoholism in DUI's. Because no gold standard exists for alcoholism, we used alternative (clinical) standards to validate BAT.

Medical diagnosis is classification with an aim. The diagnostic tool BAT is developed and tested in this study with the aim to diminish the number of false positives and false negatives and to diminish the variability between the subjective clinical estimates of different clinicians in a forensic psychiatric setting.

Summarizing: In chapter two we show that AUD diagnoses cannot be inferred from HAU diagnosis. In chapter three we find a prevalence estimation of alcoholism (AUD and/or HAU) in the DUI population of about 50%. In chapter four we show that the %CDT test without trisialo-isoforms is superior to identify dependent alcoholics. In chapter five we develop and validate BAT, a confirmation test for HAU. In chapter six, BAT is validated in a population of DUI's. In chapter seven we summarize the results of this dissertation, discuss its' clinical implications and provide suggestions for future research.

References

1. Verheul R, Brink W van den, Geerlings P. A three-pathway psychobiological model of craving for alcohol. Alcohol and Alcoholism 1999; 34:197-222.
2. de Bruijn C, Korzec A, Van den Brink W: Behandeling van alcoholisme: maatwerk. In: Bakker A, Korzec A: De gecompliceerde kliniek 2003. Benecke N.I. Amsterdam 2002.
3. Dutch regulations on driving ability: Regeling Maatregelen rijvaardigheid en geschiktheid. Regeling van de Minister van Verkeer en Waterstaat, houdende nadere regels met betrekking tot maatregelen rijvaardigheid en geschiktheid, bedoeld in Hoofdstuk VI, paragraaf 9, van de Wegenverkeerswet 1994. RV 216650.
4. Saunders JB, Lee NK. Hazardous alcohol use: its delineation as a subthreshold disorder, and approaches to its diagnosis and management. Comprehensive Psychiatry 2000; 41:95-103.
5. 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. Alcoholism: Clinical and Experimental Research 2002; 26:332-339.
6. Diagnostic and Statistical Manual of Mental Disorders, 4th edn. American Psychiatric Association Press, Washington, DC 1994.
7. ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines. Geneva: World Health Organization, 1992.
8. Grant BF, Harford TC The relationship between ethanol intake and DSM III alcohol use disorders. Journal of Substance Abuse 1988-89; 1:231-252.
9. Dawson DA Drinking patterns among individuals with and without DSM-IV alcohol use disorders. Journal of Studies on Alcohol 2000; 61:111-120.
10. Vingilis, E. (1989) Are drinking-drivers alcoholics? A review of the literature. In High alcohol consumers and traffic: Proceedings of the International Workshop: High Alcohol Consumers and Traffic, Nov 28-30, pp.165-182. Inrets 1989.
11. 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. European Addiction Research 2003; 9:182-187.
12. Mischke, HD and Venneri RL. Reliability and validity of the MAST, Mortimer Fuilkins Questionnaire and CAGE in DWI Assessment. Journal of Studies on Alcohol 1987; 48:492-501.
13. Korzec A, Bär M, Koeter MWJ, de Kieviet W. Diagnosing alcoholism in high risk drinking drivers: comparing different diagnostic proce-

dures with estimated prevalence of hazardous alcohol use. Alcohol and Alcoholism 2001; 36: 594-602
14. Poole KW, Flynn PM, Rao AV, Cooley, PC. Estimating the prevalence of drug use from self-reports in a cohort for which biological data are available for a subsample. American Journal of Epidemiology 1996; 144: 413-420.
15. Acronym for 4 questions about alcohol problems.
Mayfield D, McLeod G, Hall P. (1974) The CAGE questionnaire: validation of a new alcoholism screening instrument. American Journal of Psychiatry 1974; 131: 1121-1123.
16. Conigrave, K. M., Saunders, J. B. and Whitfield, J. B. () Diagnostic tests for alcohol consumption. Alcohol and Alcoholism 1995; 30:13-26.
17. Bean P, Harasymiw J, Peterson CM, Javors M. Innovative Technologies for the Diagnosis of alcohol abuse and monitoring abstinence. Alcoholism: Clinical and Experimental Research 2001; 25:309-316.
18. Huseby NE, Nillssen O, Kanitz RD. Evaluation of two biological markers combined as a parameter of alcohol dependency. Alcohol and Alcoholism 1997; 32: 731-737.
19. Sillanaukee P, Massot N, Jousilahti P, Vartiaine E, Poikolainen K, Plsson U, Alho H. Enhanced clinical utility of Gamma -CDT in a general population. Alcoholism: Clinical and Experimental Research 2000; 8: 1202-1206.
20. Hartz A, Guse CI, Kajdscy-Balla A. Identification of heavy drinkers using a combination of laboratory tests. Journal of Clinical Epidemiology 1997; 50: 1357-1368
21. Skinner HA, Holt S, Sheu WJ, Israel Y. Clinical versus laboratory detection of alcohol abuse: the alcohol clinical index. British Medical Journal 1986; 292:1703-1708.
22. V. Barnett, Comparative Statistical Inference, Wiley 1999.
23. Kraus P, Clark D. Representing uncertain knowledge; an artificial Intelligence approach. Intellect Books London 1993
24. Kahneman D, Slovic P, Tversky A. (Eds) Judgement under uncertainty: Heuristics and biases. Cambridge University Press, Cambridge 1991.
25. 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.
26. Feinstein AR. Clinical Epidemiology. Philadelphia: WB Saunders. 1985.

Appendix

Bayes theorem can be translated in diagnostic terms of positive predictive value, sensitivity, specificity and prevalence. Positive predictive value means the probability that the illness is present if the test is positive, sensitivity is the ratio of positive test results/ all test results if the illness is present, specificity is the ratio of negative test results / all test results if the illness is absent and prevalence is the ratio of illness present/ all subjects in the sample.
The formula used is:

1) Positive predictive value = sensitivity x prevalence / sensitivity x prevalence + (1-specificity) x (1- prevalence)
2) If one uses the average sensitivity values found in studies for hazardous drinking of AST as 20% and GGT 35% and specificity values for AST as 90% and for GGT as 77,5% (14) this signifies:
3) When prevalence is 10% then using AST the equation results in:
   Positive predictive value AST = 0,2 x 0,1/0,2 x 0,1 + 0,1 x 0,9 = 0,02/0,09 = 0,18

   using GGT the equation results in:
   Positive predictive value GGT = 0,35 x 0,1/0,35 x 0,1 + 0,225 x 0,9 = 0,035/0,203 = 0,15
4) When prevalence is 60% then using AST the equation results in:
   Positive predictive value AST = 0,2 x 0,6/0,2 x 0,6 + 0,1 x 0,4 = 0,12/0,16 = 0,75

   using GGT the equation results in:
   Positive predictive value GGT = 0,35 x 0,6/0,35 x 0,6 + 0,225 x 0,4 = 0,21/0,3 = 0,70

For easy calculation see the online clinical calculator at:
http://www.intmed.mcw.edu/clincalc/bayes.html

Chapter 2*

IS THERE A RELATION BETWEEN BIOCHEMICAL MARKERS OF HAZARDOUS DRINKING AND DSM-IV ALCOHOL USE DISORDERS?

A study in a population of well-functioning men with hazardous alcohol use.

Abstract

The purpose of this study was to examine the relation between biochemical markers of hazardous alcohol use and the presence of Alcohol Use Disorder (AUD) diagnoses within a population of well-functioning male heavy drinkers.
A group of 57 subjects with a consumption of at least 28 alcoholic units (AU)/week was recruited from wine-tasting clubs. Within this group, a comparison was made between those individuals who met the criteria of AUD and those who did not. We compared biochemical markers and drinking habits of both groups. No significant differences were found between the individuals with AUD and those without AUD, or between individuals with alcohol dependence and those without AUD, except for their drinking pattern. These findings raise doubt of the possibility to infer AUD-diagnoses from clinical and biochemical signs of hazardous use in heavy wine drinkers.

This chapter is an abridged version of an earlier study about the discriminant validity of AUD disorders (1) from a different perspective: 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. European Addiction Research 2003; 9:182-187

Acknowledgements
We thank Gareth Davies, M.D., and Jan van Lissum, Master Taster, for essential help in the recruiting of subjects.

INTRODUCTION

According to Dutch regulations on driving ability, a selected proportion of Drivers Under Influence (DUI' s) are mandatory examined by a clinician. Offenders are informed that they will lose their license in case of non-cooperation with the examination. Diagnostic procedures in this context are part of an administrative legal procedure to evaluate whether the subject has the right to have a driving license. Under Dutch law, it is demanded that the subject has refrained from alcohol misuse for the last 12 months. In cases where alcoholism is diagnosed, the license is withdrawn.

The question that the Dutch Traffic Test Organization, Disqualification Division, asks from the clinician is: What is the psychiatric diagnosis, based on clinical relevant signs and/or criteria of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)? In a subtext this is explained as: is there alcoholism in the broadest sense? Implicit in this question is the idea that clinical signs, like elevated biochemical markers that are indicative of hazardous alcohol use, make a clinical DSM-IV diagnosis of dependence or abuse more probable, when criteria of DSM-IV cannot be assessed reliably. A specific problem in diagnosing DUI's with questions whether AUD criteria are met is the high denial in this population. DSM-IV Alcohol Use Disorders comprise the psychiatric diagnoses alcohol dependence and alcohol abuse. These diagnoses emphasize loss of control and alcohol related social, psychological and physical consequences.

This raises questions about the relation between DSM-IV Alcohol Use Disorders (AUD), and elevated biochemical markers that are indicative of hazardous alcohol use.

Besides approaching alcoholism as a mental disease, as defined in DSM-IV and ICD-10, one can approach alcoholism as all unhealthy drinking patterns, defined as hazardous alcohol use (HAU), emphasizing their effect on physical health (2,3). A great advantage of HAU diagnosis for diagnosing alcoholism in DUI's is that it can be diagnosed with the aid of different biochemical markers, thus bypassing the problem of denial.

A vast amount of research exists on the relationship between hazardous alcohol intake and AUD diagnosis. Although it could be ar-

gued that excessive consumption is implicit in the criteria for AUD there is no fixed volume of ethanol intake that is necessary or sufficient for a classification of alcohol dependence (4). Rather than quantity items, drink preference and drinking patterns, such as frequency of intoxication or daily use, are correlated with AUD (4-10). It is unknown whether biochemical markers that indicate hazardous alcohol use can predict the existence of AUD disorders.

In this study, we investigated a group of non-treatment seeking well-functioning wine drinking men. We chose this group because clinical experience tells us that the majority of hazardous drinking DUI's are non-treatment seeking and are relatively well functioning. The purpose of the study was to explore whether in populations of DUI's (that partly consists of heavy drinkers) positive biomarkers of heavy drinking can differentiate between subjects with and without AUD. Therefore we were interested whether in a group of heavy drinkers who had no apparent reason to deny their alcohol problems, biomarkers would predict AUD diagnosis.

The main question was therefore whether one can infer AUD diagnosis from clinical or biochemical signs, that are used to detect HAU. We expected the group with AUD, especially those with dependence, to have higher outcomes indicative of HAU. Our second question was: If a difference is found, is it sufficiently large to be used for diagnostic purposes?

SUBJECTS AND METHODS

Study design

The study design was observational, non-intervention case-control. In order to get a homogenous population, we included only wine-drinking males. The inclusion criterion was a minimal average consumption of 28 alcoholic units per week, which is described in recent literature as hazardous alcohol use (2,3). The subjects were recruited at wine-tasting conventions and by means of advertisements in a wine magazine.

Subjects

We recruited 68 men, of which 57 met the inclusion criteria. The remaining 11 subjects had consumed less than 28 alcoholic units per week over the last 90 days. A psychiatric resident and psychiatrist examined all subjects in the period of July 1998 until March 2001.

Main outcome measures

Within this group, a comparison was made between those individuals who met the criteria of an AUD-diagnosis according to either DSM-IV or ICD-10 and those who did not (11,12). In addition, the subgroup of individuals with dependence was compared to those without an AUD-diagnosis according to DSM-IV or ICD-10.

The alcohol section of the CIDI-2 (section J) was used to assess symptoms of alcohol use disorders. The CIDI is a validated and reliable, fully structured diagnostic interview, which enables for making diagnoses according to ICD-10 and DSM-IV- criteria. Several studies have found that the diagnostic concordance between the CIDI and other diagnostic instruments (AUDADIS, SCAN) was excellent for dependence, but somewhat lower for abuse and harmful use (13-17). Subsequently DSM-IV and ICD-10 diagnoses were made using the CIDI computer algorithm.

Alcohol intake and patterns of alcohol use over the last three months were assessed using Timeline Followback (TLFB-90). This is a retrospective self-report survey, which enables the collection of reliable information on drinking behaviour (18,19). The amount of alcohol was documented in alcoholic units (AU), a standard drink in the Netherlands containing approximately 10 grams of ethanol.

As an indication of level of response to alcohol, subjects were asked how many units were required to produce an alcohol effect. Furthermore, the number of cigarettes smoked was documented.

Biochemical markers, including mean corpuscular volume of erythrocytes (MCV), aspartate aminotransferase (AST), alanine aminotransferase (ALT), gamma-glutamyltransferase (GGT) and carbohydrate-deficient transferrin (CDT), were assessed as indicators of cellular damage due to alcohol or its metabolites and as possible predictors of future harm (20). For details concerning the analytical procedures, see our report on diagnosing alcoholism in drinking drivers (21). In the present study, we used another CDT

test: ChronAlcoI.D. (Sangui Biotech Inc., U.S.A.). This test has been validated analytically and clinically (22,23). The upper reference limits in the current study at 37°C were GGT> 65U/l, MCV> 100fl, CDT> 2.7%.

Statistics

Main group comparisons were performed by using χ^2-test or Fisher's exact test for categorical variables, and, because of the relatively small number of subjects, the Mann-Whitney U-test was used for continuous variables. We used a two-tailed significance level of 5%. For these analyses, Statistics Package for Social Sciences was used (SPSS for Windows, 10.1, 2000).

RESULTS

AUD-diagnosis

Of the 57 participants, 30 met the criteria of an AUD-diagnosis according to DSM-IV or ICD-10. Out of the 30 participants with an AUD-diagnosis, 13 met the criteria of alcohol abuse or harmful use; 17 met the criteria of alcohol dependence. The remaining 27 subjects did not meet the criteria of an AUD-diagnosis. Out of the 30 participants with AUD, two participants had an AUD-diagnosis based on only one symptom (recurrent alcohol use in situations in which it is physically hazardous, e.g., drunken driving). The other 28 all met at least three criteria of either abuse or dependence. Out of the 27 participants with no AUD-diagnosis, 11 subjects did meet one criterion of dependence (concerning loss of control), two subjects met two criteria (one met two criteria concerning loss of control and one had loss of control and withdrawal symptoms) (see table 1). Those individuals are sometimes described as diagnostic orphans (24,25).

We made a comparison between the participants with an AUD-diagnosis and those without. We also made a comparison between the subgroup of participants with dependence and the participants without an AUD-diagnosis.

Table 1.
AUD diagnoses and number of diagnostic criteria (of either abuse or dependence) met (n=57)

	AUD - (n=27)	AUD + (n=30)
Abuse / harmful use	0	13
Dependence	0	17
No criterion +	14	0
One criterion +	11	2
Two criteria +	2	0
Three or more criteria +	0	28[1]

[1]Of these 28 subjects, 17 met three or more dependence criteria. They met the number of criteria needed for the dependence diagnosis. The other 11 subjects only met one or two dependence criteria and thus did not meet the number needed for the dependence diagnosis. However, they also met one or more of the abuse criteria.
So, when taking these criteria together, they met three or more criteria.

Demographic variables
There were no significant differences between the groups concerning demographic variables (see table 2).

Drinking behaviour and smoking status
There was no significant difference in the total amount of alcohol consumed per week between the groups, nor when corrected for body weight. The percentage of days in which drinking occurred did not differ significantly either. However, the participants with an AUD diagnosis did engage more often in binge drinking (drinking 10 AU/day or more) than those without AUD. The subgroup of participants with a dependence diagnosis had more drinks on an average drinking day (abstinent days not included) than those without AUD (see table 3).
The number of alcoholic units needed to notice a first effect - level of response (LRA- an indication of tolerance) was not significantly different between the groups (see table 3).

The smoking status between the groups as well as the quantity of cigarettes smoked did also not differ significantly between the groups. However, there was a trend showing that the subjects with AUD, especially those with dependence, smoked more cigarettes per day (see table 3). From epidemiological studies, it can be concluded that nicotine dependence rates increase sharply up to half a packet of cigarettes per day (26).

Therefore, we also made a comparison between the groups of subjects smoking more than ten cigarettes a day. This difference was not statistically significant either.

Table 2.
Demographic variables of 57 wine drinkers with and without AUD

	AUD- (n=27)	AUD+ (n=30)	Df	Chi square	Mann Whitney U	p
Age (mean ± SD)	51 ± 11	48 ± 10			370	0,58
Living with partner (%)	89	77	1	1,5		0,23
Education			2	1,4		0,50
Primary (%)	4	10				
High (%)	33	40				
University(%)	63	50				
Employment			3	5,4		0,14
Full time (%)	70	73				
Part time (%)	0	10				
Unemployed (%)	0	3				
Pension/retired (%)	30	14				

No differences significant at p<0,05

Clinical signs and biochemical markers

There were no significant differences between the groups regarding abnormalities at physical and laboratory examination (see table 4). The mean values of the biochemical markers did not differ significantly between the groups either.

Subjects with AUD: mean CDT 3,4; subjects without AUD: mean CDT 2,9.

Subjects with AUD: mean GGT 57 U/l; subjects without AUD: mean GGT 72 U/l. Subjects with AUD: mean ALAT 54 U/l; subjects without AUD mean ALAT 40 U/l . Subjects with AUD: mean ASAT 37 U/l; subjects without AUD mean ASAT 31 U/l. Subjects with AUD: mean MCV 92,4 fl; subjects without AUD: mean MCV 91,7 fl.

Table 3.
Drinking and smoking behaviour over the past 90 days

	AUD - (n=27)	AUD + (n=30)	Mann Whitney U	P
AU/week[1]	50,1 (27,3)	53,1 (18,4)	319	0,17
AU/drinking day[2]	7,3 (4,0)	8,1 (3,0)	288	0,06
% of days on which drinking occurred[3]	98,2 (3,5)	94,6 (9,3)	352	0,33
% of days with binge-drinking[4]	16,2 (34,0)	26,9 (33,0)	283	0,04*
Level of Response[5]	4,3 (3,0)	3,2 (2,1)	349	0,36
Cigarettes/day[6]	1,7 (5,5)	5,6 (11,6)	315	0,06
	AUD- (n-27)	Dependence (n=17)	Mann Whitney U	P
AU/week[1]	50,1 (27,3)	54,2 (17,5)	164	0,11
AU/drinking day[2]	7,3 (4,0)	8,4 (3,2)	147	0,04*
% of days on which drinking occurred[3]	98,2 (3,5)	94,1 (10,6)	209	0,56
% of days with binge-drinking[4]	16,2 (34,0)	23,5 (30,0)	170	0,13
Level of Response[5]	4,3 (3,0)	2,9 (2,1)	180	0,22
Cigarettes/day[6]	1,7 (5,5)	6,6 (13,4)	180	0,10

( ) standard deviation

[1] mean number of alcoholic units (± 10 g alcohol) per week over the last 90 days

[2] mean number of alcoholic units per drinking day (abstinent days not included) over the last 90 days

[3] mean percentage of the last 90 days on which the subjects drank at least one alcoholic unit

[4] mean percentage of the last 90 days on which the subjects drank ten alcoholic units or more

[5] mean number of alcoholic units needed to notice a first effect

[6] mean number of cigarettes smoked per day over the last 90 days

*Significant difference in the Mann-Whitney U-test ($p<0,05$)

Table 4.
Abnormalities in physical examination and biochemical markers

	AUD - (n=27)	AUD + (n=30)	Df	Chi square	p
BMI > 25 (%)	56	37	1	2,04	0,15
RR > 160/95 (%)	26	23	1	0,05	0,82
Facial erythema (%)	17	7	1	1,00	0,32
Liver enlargement (%)	4	3	1	0,01	0,94
CDT > 2.7 %† (%)	44	52	1	0,32	0,57
GGT > 65 U/l (%)	26	27	1	0,00	0,95
MCV > 100 fl (%)	0	3	1	0,01	0,94
	AUD - (n=27)	Dependence (n=17)			
BMI > 25 (%)	56	29	1	2,88	0,09
RR > 160/95 (%)	26	24	1	0,03	0,86
Facial erythema (%)	17	6	1	0,83	0,36
Liver enlargement (%)	4	6	1	0,11	0,74
CDT > 2.7 %† (%)	44	59	1	0,89	0,35
GGT > 65 U/l (%)	26	29	1	0,06	0,80
MCV > 100 fl (%)	0	6	1	1,63	0,20

† Three missing values. No differences in chi-square test at $p<0,05$

DISCUSSION

To our surprise, in our population of wine-drinking men, we found almost no differences in drinking amount, biochemical markers and clinical signs between those with, and those without an AUD-diagnosis.

A number of limitations of the current study deserve comment. To begin with, the relatively small sample size must be considered. We did not have the power to detect small differences. Therefore, our results, especially those concerning the subgroup with dependence, must be interpreted with caution. The results regarding the subgroup with dependence should rather be seen as a support for the results of the total group with AUD, than as an independent result. The fact that our measurements nearly all point in the same direction, can be regarded as a corroboration of our results. In addition, when using a more lenient significance level of 10%, hardly any more differences are found (see table 3 and 4).

Secondly, there might be selection. Evidently, wine drinkers are not representative of heavy drinkers in general. Wine drinkers seem to be better educated and of a higher socio-economic class than other heavy drinkers (27). DUI's are a very specific subgroup of heavy drinkers as well. Therefore, the generalizability of our results to DUI's may be limited. In western countries, drinking beer or preference for beer is more likely to be associated with high-risk behaviours, such as heavy and excessive drinking, drive after drinking and other alcohol-related problems, than are other types of beverage (4-8). It is therefore possible that in a DUI population of heavy drinkers there would be a statistical difference.

Also, there is a chance that, due to our sampling methods, we selected a special group of wine drinkers. Denial is traditionally considered a cardinal feature of alcoholism (28). Therefore, it is possible that in the total population of well-functioning wine-drinkers there are more alcoholics, and that we only selected the less severe part of the alcoholism spectrum because the more severe alcoholics were unwilling to participate in the study. This may have reduced the probability to show discriminant validity.

Thirdly, it is possible that the results represent the low discriminant validity of CIDI rather than of AUD diagnoses. However, this seems unlikely as CIDI has been well validated, especially for dependence. Furthermore, the clinical relevance of symptoms was checked.

In hindsight, the results of this study are not remarkable from a biological perspective. The two groups, with and without AUD diagnosis, have the same drinking behaviour, except the percentage of days of binge drinking. Binge drinking is defined as more than 10 AU/day. Compared to the average drinking behaviour of the subjects in our study population (>7 AU/day), this seems, biologically, not a large extra impact.

Lack of biological knowledge and an enormous variety of responses to alcohol mar the medical scientific debate about the definition of alcoholism. It is influenced by the choice of the type of definition, cultural attitudes about the use of alcohol, considerable individual differences of stimulant, disinhibitative, sedative effects of alcohol, and social or physical damage of excessive alcohol use. It is known that some subjects respond by elevation of CDT when

using hazardous amount of alcohol, some with GGT elevation and some with elevation of both or none.

Whatever the reasons for this phenomenon are, the reasons are unlikely to be found in the current AUD criteria. In our population it is impossible to infer AUD diagnoses from biomarkers indicative of HAU.

CONCLUSION

Within this population of well-functioning wine-drinking men, individuals with AUD hardly differ from those without AUD in terms of biochemical markers. Even individuals with dependence can hardly be distinguished from those without AUD. Taking into consideration the methodological limitations of our study, we must question the possibility to infer current AUD-diagnoses by means of clinical signs and biochemical markers.

References

1. de Bruijn H, Korzec A, Arndt T, van den Brink W. The validity of alcohol use dis- order in well-functioning men with hazardous alcohol use. European Addiction Research 2003; 9:182-187.
2. 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.
3. 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. Alcoholism: Clinical and Experimental Research 2002; 26:332-339.
4. Dawson DA: Drinking patterns among individuals with and without DSM-IV alcohol use disorders. Journal of Studies on Alcohol 2000; 61:111-129.
5. Muthen BO, Muthen LK. The development of heavy drinking and alcohol-related problems from ages 18 to 37 in a U.S. national sample. Journal of Studies on Alcohol 2000; 61:290-300.
6. Lemmens P, Tan ES, Knibbe RA Measuring quantity and frequency of drinking in a general population survey: a comparison of five indices. Journal of Studies on Alcohol 1992; 53:476-486.
7. Jensen MK, Andersen AT, Sorensen TI, Becker U, Thorsen T, Gronbeack M. Alcoholic beverage preference and risk of becoming a heavy drinker. Epidemiology 2002; 13: 127–132.
8. Chung W. Types of alcoholic beverage and high-risk drinking: how risky is beer drinking in Korea? Alcohol and Alcoholism 2004; 39:39-42.
9. York L. Progression of alcohol consumption across the drinking career in alcoholics and social drinkers. Journal of Studies on Alcohol 1994; 56:328-336.
10. Grant BF. The relationship between ethanol intake and DSM-III-R Alcohol dependence: results of a national Survey. Journal of Substance Abuse 1993; 5:257-267.
11. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, ed 4, American Psychiatric Association Press, Washington, DC, 1994.
12. World Health Organisation. The ICD-10 Classification of Mental and Behavioural Disorders. World Health Organisation, Geneva, 1992.

13. Cottler LB, Grant BF, Blaine J, Mavres V, Pull C, Hasin D: Concordance of DSM-IV alcohol and drug use disorder criteria and diagnoses as measured by AUDADIS-ADR, CIDI and SCAN. Drug and Alcohol Dependence 1997; 47:195-205.

14. Üstün B, Compton W, Mager D, Babor T, Baiyewu O, Chatterji S: WHO study on the reliability and validity of the alcohol and drug use disorder instruments: overview of methods and results. Drug and Alcohol Dependence 1997; 47:161-169.

15. Pull CB, Saunders JB, Mavreas V, Cottler LB, Grant BF, Hasin DS, Blaine J, Mager D, Üstün B: Concordance between ICD-10 alcohol and drug use disorder criteria and diagnoses as measured by the AUDADIS-ADR, CIDI and SCAN: results of a cross-national study. Drug and Alcohol Dependence 1997; 47:207-216.

16. Cottler LB, Robins LN, Grant BF, Blaine J, Towle LH, Wittcehn HU, Satorius N: The CIDI-core substance abuse and dependence questions: cross-cultural and nosological issues. The WHO/ADAMHA field trial. British Journal of Psychiatry 1991; 159:653-658.

17. Compton WM, Cottler LB, Dorsey KB, Spitznagel EL, Mager DE: Comparing assessments of DSM-IV substance dependence disorders using CIDI-SAM and SCAN. Drug and Alcohol Dependence 1996; 41:179-187.

18. Sobell LC, Maisto SA, Sobell MB, Cooper AM: Reliability of alcohol abusers' self-reports of drinking behavior. Behavioral Research and Therapy 1979;17:157-160.

19. Allen JP, Fertig RZ, Litten P, Sillanaukee P, Anton RF Proposed recommendations for research on biochemical markers for problematic drinking. Alcoholism: Clinical and Experimental Research 1997; 21:244-247.

20. Conigrave KM, Saunders JB, Reznik RB, Whitfield JB: Prediction of alcohol-related harm by laboratory test results. Clinical Chemistry 1993; 39:2266-2270.

21. Korzec A, Bär M, Koeter MWJ, Kievit de W: Diagnosing alcoholism in high-risk drinking drivers: comparing different diagnostic procedures with estimated prevalence of hazardous alcohol use. Alcohol and Alcoholism 2001; 36:594-602.

22. Korzec A., Arndt T, Bär M, Koeter MWJ: Trisialo-Fe_2-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.

23. 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. Clinical Chemistry 2000; 46:483-492.

24. Hasin D, Paykin A: Dependence symptoms but no diagnosis: diagnostic "orphans" in a community sample. Drug and Alcohol Dependence 1998; 50:19-26.

25. Pollock NK, Martin CS: Diagnostic orphans: adolescents with alcohol symptoms who do not qualify for DSM-IV abuse or dependence diagnoses. American Journal of Psychiatry 1999;156:897-901.

26. Kandel DB, Chen K: Extent of smoking and nicotine dependence in the united states: 1991-1993. Nicotine and tobacco research 2000; 2:263-274.

27. Mortensen EL, Jensen HH, Sanders SA, Reinisch JM: Better psychological functioning and higher social status may largely explain the apparent health benefits of wine: a study of wine and beer drinking in young Danish adults. Archives of Internal Medicine 2001; 161:13-27.

28. Tarter RE, Alterman AI, Edwards KL: Alcohol denial: a biopsychological interpretation. Journal of Studies on Alcohol 1984; 45:214-218.

Chapter 3*

DIAGNOSING ALCOHOLISM IN HIGH RISK DRINKING DRIVERS:
comparing different diagnostic procedures with estimated prevalence of hazardous alcohol use.

Abstract

In several European countries drivers under influence (DUI), suspected of an alcohol use disorder (AUD, 'alcoholism') are referred for diagnostic examination. The accuracy of diagnostic procedures used in diagnosing AUD in the DUI population is unknown. The aim of this study was to compare three prevalence estimates of AUD based on a structured clinical interview (SCID), a restrictive diagnostic procedure (RDP) and usual clinical diagnostic procedure (CDP), with a prevalence estimate based on sensitivity and specificity data of biological markers of excessive use of alcohol in non-judicial samples. The latter unbiased estimate provides an external yardstick against which the biased patient-based prevalence estimates in this special sample can be evaluated.The unbiased estimate derived from sensitivity and specificity data resulted in a prevalence estimate of excessive use of alcohol between 74 % and 82 %, which is much higher than the three diagnostic procedures. SCID identifies maximally 5% of alcoholics found with the unbiased estimate. RDP identified ≥ 31% of the unbiased estimate, while CDP identified ≥ 60% of the unbiased estimate. The high chance of false

* Previously published: Alex Korzec, Marij Bär, Maarten W.J. Koeter, Wim de Kieviet. Diagnosing alcoholism in high risk drinking drivers: comparing different diagnostic procedures with estimated prevalence of hazardous alcohol use. Alcohol and Alcoholism 2001; 36:594-602.

positive diagnosis, however, makes CDP unacceptable in the legal context of AUD diagnosis in DUI populations.

INTRODUCTION

Alcohol Use Disorder (AUD, "alcoholism") increases the risk of involvement in road traffic accidents (1-4). The collision rates of alcoholics are twice as much as the collision rates of non-alcoholic drivers (5). These findings have engendered traffic law regulations in several European countries. The regulations stipulate that drivers under influence (DUI), suspected of being alcoholics, need to submit to a medical examination, in order to refute or substantiate that suspicion. (6). According to Dutch regulations on driving ability (7) different groups of DUI' s are examined (see subject and methods section). Offenders are informed that they will lose the license in case of non-cooperation with the examination.

Diagnostic procedures in this context are part of an administrative legal procedure to evaluate whether the subject has the right to have a driving license. Under Dutch law, it is demanded that the subject has refrained from alcohol misuse for the last 12 months. In cases where alcoholism is diagnosed, the license is withdrawn.

The legal context causes two problems in identifying alcoholics. The first problem is the understandable low validity of self-reporting of alcohol problems in DUI subjects (8). Secondly, in many instances a diagnosis of alcoholism has to be defended in legal procedures. Diagnoses, based on clinical judgement and data with an unreliable correlation with alcoholism, are increasingly challenged in court with questions about the chance of false positive diagnosis.

The accuracy of diagnostic procedures used in diagnosing alcoholism in DUI's is unknown and has, to the best of our knowledge, never been investigated before.

Research suggests a considerable prevalence of AUD in DUI populations. In a review of prevalence reports up to 1986, Vingilis estimated the prevalence between 25 and 50%, depending on the sampling of the population and the criteria used for alcoholism (9). More recent studies, using DSM III criteria and biochemical tests show the same prevalence range (Table 1). However, there is reason for scepticism about the validity of these prevalence values.

Table 1
Summary of investigations on prevalence of alcoholism in DUI populations using DSM III and biochemical tests

AUTHOR(S)	No	SELECTION
Scoles et al. 1986	500	DUI - population prior to trial
Miller et al. 1986	461	Convicted DUI's; 25% involved in traffic accident
Iffland et al., 1995	534	Arrested DUI with BAC>0,8‰
Dunbar et al. 1985	58	Male and female DUI's older than 30 years and involved in accident
Dunbar et al. 1985	140	Male and female DUI's older than 30 years
Papoz et al. 1986	3427	Male accident victims presented at emergency ward
Papoz et al. 1986	3427	Male accident victims presented at emergency ward of Hospital
Pikkarainen et al. 1989	176	Apprehension at roadblock DUI's with BAC>0.5 ‰
Pikkarainen et al. 1989	183	Apprehension not at road block, on suspicion for alcohol use, DUI's and BAC>0.5 ‰
Pikkarainen et al. 1989	176	Apprehension road block. DUI's with BAC>0.5 ‰
Michiels 1996	877	Male and female DUI"s
Ruud et al. 1993	150	Males convicted for DUI
Gjerde et al. 1986	269	Male DUI's. 61% younger than 30 years
Gjerde et al. 1987	50	
Jaster et al. 1993	110	Male DUI's
Lutz et al. 1992	219	Male and female DUI's
Morgan et al. 1996	93	Male DUI's with BAC ≥ 200 mg/dl, or repeated conviction in last 10 years, or failure to provide specimen for analysis

DSM-III (American Psychiatric Association, 1980); CDT, carbohydrate-deficient transferrin; GGT, gamma-glutamyltransferase; MCV, mean corpuscular value; BAC, blood-alcohol concentration.

CRITERION ALCOHOL USE DISORDER	PREVA-LENCE %	CORRECTION FOR TEST PA-RAMETERS
DSM III alcoholism classification	27.4%	
DSM III alcoholism classification	54%	
CDTect > 20 U/l	54,5%	No
GGT > 50 U/l	48%	No
GGT > 50 U/l	24.3%	No
GGT > 40 U/l	30%	No
Combination of abnormal values of GGT and MCV corresponding to values of 90% of control population that used 80g pure alcohol daily	27%	Yes
GGT >50 U/l (Method not described)	25%	No
GGT >50 U/l (Method not described)	29%	No
Combination of DUI recidivism and GGT> 50 U/l (Method not described)	20.5%	No
GGT > 56 U/l in males	32%	Yes
GGT > 42 U/l in females (Method not described).	16%	
Combination of elevated GGT, recidivism DUI and BAC > 2.0 ‰	14%	
CDT>74mg/l	35%	No
GGT>50 U/l	23%	
GGT > 50 U/l in males	21%	
CDT >74 mg/l	60%	No
Combination of two abnormal values: GGT>58mmol/L; MCV> 96fl; CDT index >15%.	34%	Yes
GGT>28 U/l in males		No
GGT> 18 U/l in women	20.5%	
CDT > 20	28 %	No
GGT abnormal. (Method not described)	21.5 %	

The number of cases with elevated biochemical markers cannot be equalised with true cases of alcoholism, as was done in most of the reviewed studies.

Firstly because elevated biochemical markers are not in a strict sense markers of alcoholism but of hazardous use of alcohol. More importantly, research indicates that the sensitivity of biochemical markers drops dramatically in young alcoholics, and also in drinkers with less severe alcoholism (10-14). As DUI populations consist for a not negligible part of young drivers, and severe alcoholics represent only a small minority, the reported prevalence values can be considered as conservative (9,15).

However, the number of cases with elevated biochemical markers can be used to obtain a better estimate of prevalence, if one takes into account sensitivity and specificity data of biochemical markers of alcoholism in non-judicial samples. With a formula derived from 'Bayes theorem' one can calculate the prevalence of hazardous use in a population by incorporating test results with knowledge of the sensitivity and specificity. This population-based method can be used as external criterion for the accuracy of different diagnostic procedures.

In evaluating diagnostic procedures one must consider the differences between diagnosing alcoholism in health care settings and in legal settings. In health care the main diagnostic aim is to enhance health. Therefore it is important to identify all alcoholic patients. In order to minimise the risk of missed diagnoses a high sensitivity of diagnostic procedures and tests is important. Usual clinical diagnostic procedure (CDP) in health care depend on clinical judgement which incorporates all available historical, clinical and laboratory data.

In the legal setting of medical examination in a DUI population the aim is not to enhance health but to enhance traffic safety. Because diagnosis may be challenged in court, diagnosis is restricted to sure cases. In order to minimise the risk of false positive diagnoses a high specificity of the diagnostic procedure is important. Therefore more restrictive diagnostic procedures (RDP) are used. Ideally, legal diagnostics must rely on objective, reliable and specific data, such as recent history of alcohol problems, physical signs of alcoholism or specific biochemical tests of hazardous alcohol use. In

legal settings high specificity of diagnostic tests is more important than high sensitivity, because incorrect diagnoses have unacceptable legal consequences.

Understanding the legal dilemma is essential in choosing from the different diagnostic procedures. The dilemma is to find a balance between two opposite aims.

On the one side, the requirement to enhance traffic safety (for the public); each missed diagnosis endangers traffic safety. On the other side, the requirement to protect the rights of the individual; each incorrect diagnosis has unacceptable consequences (for the individual) as incorrectly diagnosed DUI's may, for example, lose their job after being disqualified to drive.

In this study three prevalence estimations, obtained with different diagnostic procedures, are compared with each other and with an unbiased prevalence estimate based on sensitivity and specificity data of biological markers of alcoholism in non-judicial samples. The central question in this study is: How do different clinical diagnostic procedures perform in the detection of AUD, compared to prevalence of hazardous use in the population-based method?

SUBJECTS AND METHODS

Subjects

The population under study consisted of 241 consecutive male DUI's who were referred for medical examination between September 1996 and May 1998 after driving under the influence of alcohol. Of these 29 were excluded because of incomplete clinical or chemical data, leaving a study population of 212.

In accordance with Dutch traffic regulations the following groups were included for referral and examination:

1. DUI's with at least one arrest with a Blood Alcohol Concentration (BAC) ≥2.1‰ or three DUI's arrests with any BAC above 0.5 ‰ within 5 years, or refusal to cooperate with breath analysis (examination group). This group is referred by the Dutch Traffic Test Organization, Disqualification Division, who pays for the medical examination. Some basic information of the characteristics of drivers processed under these regulations were obtained from the Dutch

Traffic Test Organization, which supplied data on all DUI's examined in the Netherlands in 1997.
2. DUI's who apply for re-granting of the driving license after previous DUI, medical examination and loss of permanent driving license for 12 months because of diagnosis of alcoholism (re-examination group). In this group almost all individuals are self-referred, applying for re-licensing, and have to pay for the examination.

Standardized clinical data collection
All DUI's were examined and diagnosed by the same physician. The examination was recorded in a standardized clinical report, which was part of a legal procedure on behalf of the Dutch traffic test organization. The clinical report of each subject consisted of extensive history taking, instruments to assess AUD, namely Structured Clinical Interview (SCID) and the CAGE-questionnaire, physical examination, biochemical measurements and a conclusive clinical judgement as to whether it was probable the subject had AUD in the last 3 or 12 months. History taking was focused on clinical signs of alcoholism and on possible and probable non-alcoholic causes for elevated biochemical markers. The latter included questions about current and past illness, specifically diabetes, liver diseases, blood transfusions and intravenous drug use (because of the possibility of hepatitis C which can affect carbohydrate-deficient transferrin (CDT) (27), anaemia, and drugs that could affect biochemical markers (anti-epileptics, folate antagonists, anti- AIDS medication, fenothiazines, some diuretics and thyrostatics).

Alcoholism or AUD is defined as either alcohol abuse or alcohol dependence according to DSM- IV .

Biochemical measurements:
Venous blood samples for determination of hemoglobin (Hb), Hematocrit (Ht), Red blood cell count (E), Mean cell volume (MCV), carbohydrate-deficient transferrin (CDT), Gamma glutamyltransferase (GGT), aspartate aminotransferase (AST), and alanine aminotransferase (ALT) were taken. Serum samples for CDT were frozen within 4 hours after collection and stored at - 20°C until use. CDT was analyzed in duplicate, using a commercial kit, CDTect, of

Pharmacia and Upjohn. Measurement of serum GGT, ALT, AST was executed within 4 hours with VITROS (Ortho Clinical Diagnostics) at 37°C and re-coded for the value at 30°C. Hb, Ht, E, MCV were kept at room temperature and analyzed within 4 hours with Technicon H2 analyzer, Bayer. The reference limit of CDTect was ≥ 20 U/l, GGT ≥ 40 U/l, ALT ≥ 34 U/l, AST ≥ 33 U/l, MCV ≥ 100 fl.

Diagnostic procedures

For reasons of comparability with the population based method, as the diagnostic window of biochemical markers does not exceed 3 months, only current AUD diagnosis (within the last 3 months) is used in the different diagnostic procedures.

Data from clinical reports of every subject were processed in three diagnostic procedures: SCID, RDP and CDP. The three diagnostic procedures are not independent; SCID is incorporated into the RDP and both SCID and RDP are incorporated into CDP. The diagnostic procedures are described below in detail. Essentially SCID identifies alcoholics that are willing to report alcohol problems, RDP identifies those positive with SCID and with elevated biochemical markers that can be seen as proof of current hazardous drinking, while CDP identifies those positive on SCID and RDP and subjects with more "soft signs" of alcoholism. All resulting diagnoses refer to AUD in the 3 months prior to examination.

1. Diagnostic procedure 1: SCID. Recent alcohol problems were assessed with the Structured Clinical Interview for DSM IV Axis I disorders, clinician version, module E (alcohol use disorders) over the last 3 months [SCID-CV, (28)]. The SCID-CV is a semistructured interview for making the major DSM - IV diagnoses and is based on the Diagnostic and Statistical Manual of Mental Disorders 4 th. Edition (29). It was designed for use in clinical settings as a way of ensuring standardized assessments. A study, using earlier versions of SCID, report test-retest kappa's for current diagnoses of AUD of 0.75 (30).

AUD diagnosis was made if the subject scored positively on one of the SCID criteria of alcohol abuse or scored positively on three criteria of alcohol dependence, in the 3 months prior to the interview.

2. Diagnostic procedure 2: Restrictive diagnostic procedure (RDP). We devised a restrictive diagnostic procedure for detection of alco-

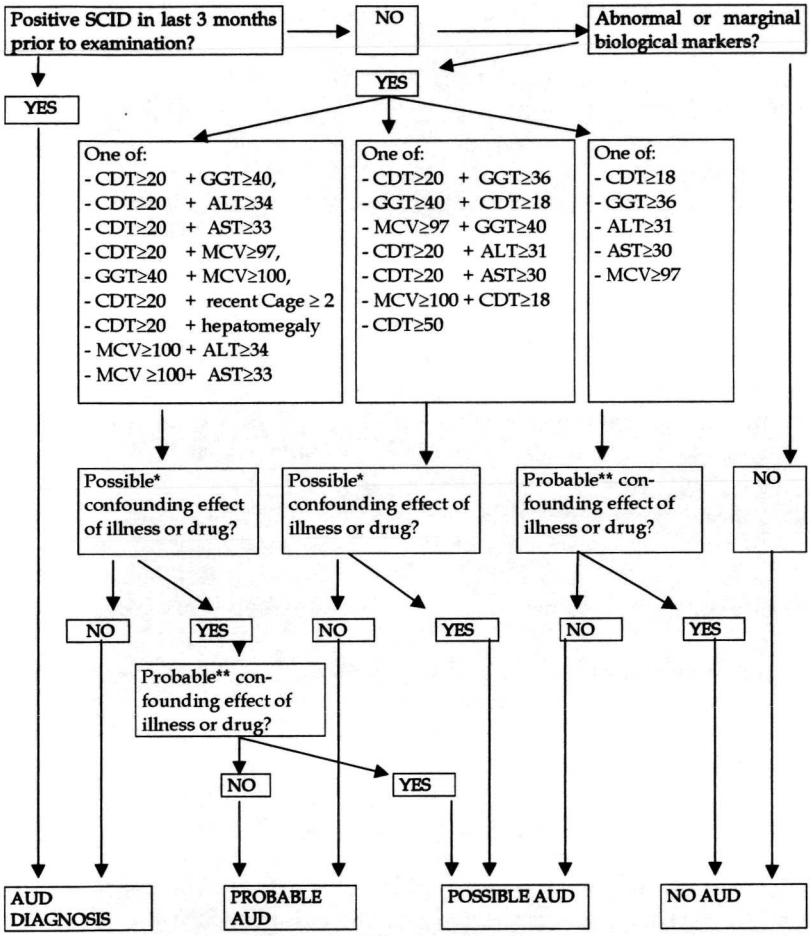
holism with the aim to maximize reliability and specificity of diagnosis. From the standardized recorded history only data from SCID-interview, the 4 CAGE questions (31) and data from history and medication were used to check for possible non-alcohol causes for raised tests. From physical examination only liver palpation was used. All biochemical measurements were used. The restrictive diagnostic procedure is described in figure 1.

AUD diagnosis was only made if SCID was positive, or if a simultaneous combination of elevated biochemical tests, or simultaneous combination of biochemical tests and clinical signs was present. In case of possible non-alcoholic causes for positive biochemical and clinical signs, diagnosis was not made. When two or more of the enzymes ALT, AST and GGT were simultaneously elevated, no AUD diagnosis was made. In the presence of indication of liver illness, such as liver enlargement, highly elevated ALT and AST, or highly elevated GGT, no AUD diagnosis was made. In the case of a moderately elevated GGT, ALT or AST there is a difficulty whether to interpret a simultaneous MCV elevation as an independent test in diagnosing AUD, as this elevation may be possibly caused by the same liver illness.

In order to diminish the small chance (in this population) of incorrectly diagnosing subjects with non-alcoholic liver disease, without increasing the much greater chance (in this population) of missing diagnosis in non-abstinent subjects with alcoholic liver disease, we used a higher cut off value of MCV in the combinations of elevated MCV, with elevated ALT, AST or GGT, as an extra safeguard against incorrect diagnosis.

The algorithm for RDP was made before analysis. Two items (no effect of alcohol $\geq$ 4 AU and blood pressure $\geq$ 160/95 were deleted post hoc, as these items didn't provide additional differentiating value.

Figure 1.
Flow chart Restrictive Diagnostic Procedure. 3 alternative criteria of increasing strictness for testing excessive use of alcohol



*Possible confounding effect when the chance that abnormal biochemical or clinical signs are caused by non-alcoholic illness or drug is estimated to be > 5%.

**Probable confounding effect when the chance that abnormal biochemical or clinical signs are caused by non-alcoholic illness or drug is estimated to be > 50%. (These criteria derive from the albeit sparse published data on the rates of elevated marker tests in various non-alcoholic conditions, and the authors' clinical estimates). SCID, structured clinical interview, CDT, carbohydrate-deficient transferrin; GGT, gamma-glutamyltransferase; MCV, mean corpuscular value; ALT, alanine aminotransferase, AST, aspartate aminotransferase; AUD, alcohol use disorder.

The choice of the different combinations of elevated biochemical measurements in RDP was motivated by our aim to achieve a specificity of approximately 95%. In non clinical settings, specificity of GGT is reported between 80- 90 %, of ALT above 80%, of AST above 90%, of elevated MCV, in men, above 90% and of elevated CDT above 80% (14,32,33). As hazardous use of alcohol elevates CDT, MCV and the enzymes GGT, ALT and AST through partly independent biological pathways, these markers can be considered as partly independent tests (34).

Demanding that a pair of diagnostic tests are simultaneously elevated before making a positive diagnosis of alcoholism, maximizes specificity and minimizes false positive labeling of innocent patients, but pays the price of lots of missed diagnoses (35). If CDT, GGT and MCV are independent tests, simultaneous elevation of CDT and GGT, or CDT and MCV, or MCV and GGT will result in specificity values between 96 and 99%. This assumption is partly confirmed in a study that measured specificity of simultaneously elevated CDT and GGT in non-alcoholic controls (36).

3.Diagnostic procedure 3: Clinical diagnostic procedure (CDP). In this diagnostic procedure a diagnosis was reached through clinical judgement after evaluation of all available data, according to usual clinical practice. Besides biochemical measurements, historical data, clinical signs, and instruments to assess alcohol problems were used.

Histories included time and circumstances of arrest. A police report of BAC, data of earlier DUI and reports of earlier medical examinations after DUI were available. No information from GP's was asked.

Recent alcohol intake was assessed by means of a structured interview. This included questions about the exact amount of alcohol units (AU) in the week prior to the examination, an estimate of the average AU per week during the last year, and questions about changes in quantity and frequency of drinking in the last year.

Hazardous drinking is defined as the level of persistent alcohol consumption being likely to result in adverse health effects: >280 g ethanol/week (37). As 1 AU is defined as a standard drink of approximately 10 g alcohol, hazardous use signifies an average of more than 28 AU weekly.

Lifetime and current alcohol problems were assessed using the CAGE questions, SCID, and questions about any past treatment for alcohol problems. A subject was considered to have had a life time alcohol problem if CAGE ≥ 2, or if a subject was ever treated for alcohol problems, or received a SCID diagnosis in the 12 months prior to the interview.

As in RDP, history included also questions about different diseases and drugs, to control for possible confounders in regard to non-alcoholic causes for elevated biochemical markers or liver enlargement.

Physical examination included breath smelling of alcohol during examination, (but no alcohol breath test), blood pressure, liver palpation and observation of skin abnormalities indicative for liver dysfunction and neurological dysfunction indicative of polyneuropathy or withdrawal symptoms.

Diagnosis of recent AUD in CDP procedure was based on clinical reasoning. All data and clinical signs were assessed as either diminishing the chance of recent AUD, increasing the chance of recent AUD, or confirming AUD diagnosis. A positive diagnosis of recent AUD was made if the above described SCID and RDP procedures resulted in an AUD diagnosis or if several AUD chance -increasing - data were present without the presence of confounding effects of illness or drugs.

Population based prevalence estimate of hazardous use.

Studies have shown that sensitivity and specificity of markers of hazardous alcohol use depend on the distribution of severe and mild cases of alcoholism in the studied cohort. A high ratio of severe/mild cases heightens sensitivity, while a low ratio lowers sensitivity.

Because we assumed that our population consists of a high risk population of hazardous users, alcoholics and social drinkers without AUD, we used sensitivity and specificity values found in studies with two high risk populations Sillanaukee et al., 1993, Huseby et al., 1997b (38,39) . Sillanaukee et al. compared hazardous drinkers with some signs of AUD, to social drinkers and found a sensitivity of 57% and a specificity of 79%. Huseby et al. compared alcohol dependent patients to non-dependant patients from a population of

men admitted to a surgical ward and found a sensitivity of 55% at a specificity of 85%. Sensitivity and specificity values refer to the relation of AUD and elevated CDT or GGT.

The estimated prevalence of AUD was computed with the following formula:

$$P = [T - (1-Sp)] / (S + Sp - 1) \quad (40)$$

where: P = prevalence; T = proportion of elevated tests (CDT or GGT) = (true positives + false positives)/ all tests. S = sensitivity = number of true-positives/(number of true-positives + number of false-negatives). Sp = specificity = number of true-negatives/(number of true-negatives + number of false-positives). Below: PPV = Positive Predictive Value = number of true positives/ (number of true-positives + number of false-positives). NPV = Negative Predictive Value = number of true negatives/ (number of true negatives + number of false negatives).

Statistical analysis

SPSS was used for computation of frequencies. Comparison of groups was performed with T- Test. Comparison of multiple groups was conducted with ANOVA.

RESULTS

Sample characteristics

The sample characteristics of the examination group were not significantly different from all DUI's examined in the Netherlands in 1997 when compared for age, average BAC and mean number of DUI arrests. The mean age of our cohort was 42.1. 31% of the DUI's were younger than 35 years (Table 2).

Table 2:
Sample characteristics of 212 DUI subjects and of all "first examined DUI's" in 1997 in the Netherlands

Examined population	First examination n=93	Re-examination n=119	Sign	First Examination in the Netherlands in 1997 n=2045	Sign
Mean Age	40,17 (11.8)	44 (11.2)	0.025*	40.29 (10.3)	0.92
Mean BAC	1.98 ‰ (0.56) ‰	1.90 ‰ (0.61) ‰	0.428	2.12 ‰ (0.58) ‰	0.064
Mean number of DUI arrests in last 5 years	1.91 (1.23)	1.41 (1.20)	0.003*	1.93 (1.19)	0.459
Mean time between last DUI and medical examination in months	6.2 (3.2)	47 (35)	0.000*		
Reported AU/week	10.4 (12.9)	5.5 (8.0)	0.002*		

* <0.05 independent sample T-test ( ) SD

The re-examination group, that consisted of subjects who applied for re-granting the drivers license, reported much less alcohol use (5,5 AU/week) than the examination group (10, 4 AU/week). In comparison, the average self reported alcohol intake in the Dutch male population is 21 AU per week (41). Only 7 subjects, (7.5%), from the examination group, and 4 subjects, (3.4%), from the re-examination group reported to drink more than 28 AU average per week in the 3 months prior to the interview.

107 out of 212 DUI's reported life time alcohol problems according to our definition (if CAGE ≥ 2, or if a subject was ever treated for

alcohol problems, or received a SCID diagnosis in the 12 months prior to the interview). In agreement with expectation the re-examination group reported more lifetime alcohol problems (61.3%) than the examination group (36.6%). The last percentage is probably the result of underreporting. According to an epidemiological study performed in 1996, the one-month, 12 month and life-time prevalence of AUD in the Dutch male population was respectively 8.5%, 13.4% and 28.3% (42).

Prevalence according to different diagnostic procedures

SCID. Applying SCID over the last three months as diagnostic procedure identified 7 DUI's with AUD in the Examination group and only one in the Re-examination group. According to SCID the estimated prevalence of AUD over the last three months in the examination group is 7.5% and 0.8% in the re-examination group (Table 3). Applying SCID over the last 12 months identified 21 DUI's with AUD in the examination group (22.6%) and 4 in the re-examination group (3.4%).

RDP. The restrictive diagnostic procedure resulted in an AUD diagnosis in 32 DUI's from the Examination group (prevalence according to RDP 34.4%), and 18 from the re-examination group (prevalence of AUD in the re-examination group 15.1%).

CDP. The clinical diagnostic procedure using all data resulted in an AUD diagnosis in 54 DUI's in the Examination group (prevalence according to CDP 58.1%), and 43 DUI's in the re-examination group (prevalence according to CDP 36.1%).

Population based prevalence computation.

The total amount of subjects with elevated CDTect or GGT was 101 (51 from the examination group and 50 from the re-examination group). The proportion of DUI's with elevated biochemical markers is 101/212. Using the sensitivity and specificity values found by Sillanaukee et al. results in an estimated prevalence of AUD for all DUI's in our study of 74%. Using the sensitivity and specificity values found by Huseby et al. result in an estimated prevalence of 82 %.

Table 3:
Estimate of Prevalence of AUD according to different diagnostic procedures compared to population based prevalence estimate of excessive use of alcohol

Diagnostic procedure	SCID 3 months		Restrictive Diagnostic Procedure				Clinical Diagnostic procedure		Population based method
DUI population n=212	+ n=8	- n=204	AUD n=50	Probable AUD n=21	Possible AUD n=59	No AUD n=82	AUD n=97	No AUD n=115	Elevated CDT or GGT n=101 (51 examination group; 50 re-examination group)
Prevalence AUD	3.8%		23.6 %				45.8%		82% or 74%
Life time alcohol problems n=107	7	100	31 62.0%	17 81.0%	25 42.4%	34** 41.5	60	47*	
Mean AU/week 7,7 (10,7)	25.6 (29.6)	7.1* (8.7)	15.5 (15.4)	8.6 (8.0)	6.7 (8.2)	3.5** (6.1)	11,9 (13)	4,1 (6,4)	
Examination group N=93	7 7.5%	86* 92.4%	32 34,4%	11 11.3%	19 20.4%	31 33.3%	54 58,1%	39 41.9%	
Re-examination group n= 119	1 0.8%	118* 99.2%	18 15.1%	10 8.4%	40 33.6%	51** 42.9%	43 36.1%	76* 63.8%	

* p <0.05 independent sample T-test ( ) SD

** p <0.05 one-way ANOVA

DISCUSSION

The three diagnostic procedures are not independent. SCID is incorporated in RDP and both SCID and RDP are incorporated in CDP. Not surprisingly, additional data result in higher AUD

prevalence values: 3.8% with SCID only, 23.5% with the restrictive diagnostic procedure and 45.8% with clinical judgement. How to explain the great difference between prevalence found with diagnostic procedures and the prevalence found with the population based method?

On one side, one has to reckon with the possibility that the low sensitivity of biochemical markers, used in the population based method, inflates the estimated prevalence of AUD beyond results of earlier research and beyond face validity. Another possible explanation is that the estimated prevalence found with the population-based method, between 82% and 74%, can be considered as maximal prevalence only. As 'hazardous drinking' encompass a larger group than the group with AUD, the criterion can be only used as maximal reference level.

On the other hand, one has to consider the possibility that the diagnostic tools to detect alcoholism in DUI's result in considerable under diagnosing.

SCID identifies maximally 5 % of all AUD found with the unbiased estimate. This performance was not unanticipated; SCID identifies only those alcoholics that are aware of their problems and are willing to be open about it. For obvious reasons most DUI's will not be open about their alcohol consumption (which was reported as 3 times lower than average in the Dutch population) or about their alcohol problems (which was reported as just a little lower than in the Dutch population).

RDP identifies 6 times as many as SCID procedure only, and at least 28 % and maximally 31% of the unbiased AUD estimate. This is a significant gain compared to SCID. At the same time it is evident that the sensitivity of the Restrictive Diagnostic Procedure is low. This result is also according to expectation. One can assume that RDP will result in under-diagnosis because physical signs of alcoholism are late symptoms of alcoholic disease, because approximately 5-20 % of alcohol dependent patients and 40 - 60 % of alcohol abusers show no elevations of biochemical tests (12,38,43) and because 31 % of our population consisted of subjects younger than 35 years. In young subjects biochemical markers have a low sensitivity for detection of alcoholism. Another reason for under diagno-

sis of RDP is that any possible non-alcoholic cause had also automatically led to exclusion of the diagnosis.

CDP identifies at least 56% of the unbiased AUD estimate and up to 60% of the examination group. Even if one thinks that the prevalence of AUD found with CDP is rather high, one has to consider that the prevalence values in this study refer to prevalence of AUD found several months after the DUI arrest. It seems reasonable to assume that prevalence of AUD at the time of arrest would be much higher.

The above mentioned prevalence is dependant on the administrative selection of DUI's for examination, which variegates in different countries. The issue here is to provide the clinician, working within a legal situation, with a method to calculate PPV and NPV for different diagnostic procedures. The diagnostic gain of CDP above RDP has significant legal disadvantages that can be illustrated by the consequences for Positive Predictive Value (PPV) of this procedure. If we would use this procedure in a population with 40% prevalence of AUD, under the optimistic assumption that CDP has a specificity of 80% and sensitivity between 60% and 95 %, the positive predictive value of CDP will vary between 66 and 75%. This may be quite acceptable in health care settings, but is evidently not acceptable in legal settings. The high chance of false positive diagnosis makes CDP unacceptable in the legal context of AUD diagnosis in DUI populations. Until better markers are available we advise physicians who participate in diagnosing AUD in DUI populations to use RDP enhanced with secondary data like circumstances of arrest.

It remains to be researched if RDP (enhanced or not) has a high enough PPV and an acceptable NPV. However, it is too optimistic to hope that such research will be able to replace clinical reasoning completely (44). As different sub-groups of DUI's have different a priori prevalence (table 3), and test parameters of biochemical markers are dependent on age and gender, different norms must be used in diagnostic procedures. Even if precise knowledge of the positive predictive values of different diagnostic procedures in different groups becomes available, one has still to answer a social, as well as the legal question: How sure one has to be of diagnosis in diagnosing alcoholism in DUI populations?

References

1 Waller JA, Turkel HW. Alcoholism and traffic death. New England Journal of Medicine 1966; 275:532-536.

2 Dunbar JA, Ogston, SA Ritchie, A, Devgun MS, Hagart J, Martin BT. Are problem drinkers dangerous drivers? British Medical Journal 1985; 290:827-829.

3. Papoz L, Weill J, l'Hoste J, Chich Y, Got C, Goehrs Y. Biological markers of alcohol intake among 4796 subjects injured in accidents. British Medical Journal 1986; 292:1234–1237.

4. Östrom M, Eriksson A. Single-vehicle crashes and alcohol: a retrospective study of passenger car fatalities in Northern Sweden. Accident Analysis and Prevention 1993; 25:171-176.

5. Vingilis E. Drinking drivers and alcoholics: Are they from the same population?. In Research advances in alcohol and drug problems, Smart, R.G., et al. eds, pp. 299-342. Plenum Press, New York 1983.

6. Nickel WR. Regranting of driving licenses- report on the working group on regranting of licenses. International council on alcohol drugs and traffic safety. Dec 1995.

7. Dutch regulations on driving ability: Regeling Maatregelen rijvaardigheid en geschiktheid. Regeling van de Minister van Verkeer en Waterstaat, houdende nadere regels met betrekking tot maatregelen rijvaardigheid en geschiktheid, bedoeld in Hoofdstuk VI, paragraaf 9, van de Wegenverkeerswet 1994. RV 216650.

8. Mischke HD and Venneri RL. Reliability and validity of the MAST, Mortimer Fuilkins Questionnaire and CAGE in DWI Assessment. Journal of Studies on Alcohol 1987; 48:492-501.

9. Vingilis E. Are drinking-drivers alcoholics? A review of the literature. In High alcohol consumers and traffic: Proceedings of the International Workshop: High Alcohol Consumers and Traffic, Nov 28-30, pp.165-182. Inrets 1989

10. Nystrom M, Peraslo J, Salaspuro M. Carbohydrate-deficient transferrin (CDT) in serum as a possible indicator of heavy drinking in young students . Alcoholism: Clinical and Experimental Research 1992; 16:93-97.

11. Allen JP, Litten RZ, Anton RF, Cross RM. Carbohydrate-deficient transferrin as a measure of immoderate drinking: Remaining issues. Alcoholism: Clinical and Experimental Research 1994; 18:799- 812.

12. Litten RZ, Allen JP, Ferrite JB. Gamma glutamyl transpeptidase and carbohydrate deficient transferrin: alternative measures of excessive alcohol consumption. Alcoholism: Clinical and Experimental Research 1995; 19:1541- 1546.

13. Huseby NE, Nillssen O, Erfurth A, Wetterling T, Kanitz RD. Carbohydrate-deficient transferrin and alcohol dependency: variation in response to alcohol intake among different groups of patients. Alcoholism: Clinical and Experimental Research 1997; 21:201-205.
14. Salaspuro M. Carbohydrate-deficient transferrin as compared to other markers of alcoholism: a systematic review. Alcohol 1999; 19(3):261-271. Elsevier Science Inc.
15. Hasin D, Paykin A, Endicott J, and Grant B. The validity of DSM-IV alcohol abuse: drunk drivers versus all others. Journal of Studies on Alcohol 2000; 60:746-755.
16. Scoles, EA, Fine EW, Steer RA. DUI offenders presenting with positive blood alcohol levels at pre-sentencing evaluation. Journal of Studies on Alcohol 1986; 47:500-502.
17. Miller BA, Whitney R, Washousky R. Alcoholism diagnoses for convicted drinking drivers referred for alcoholism evaluation. Alcoholism: Clinical and Experimental Research 1986; 10:651-656.
18. Iffland R, Grassnack F. Epidemiologische untersuchung zum CDT und andere indikatoren für alcoholprobleme in blut alcoholauffalliger Deutsche PKW fahrer. Blutalcohol 1995; 32:27-41.
19. Pikkarainen J, Pentilla A. Frequency of heavy and problem drinkers among Finish drivers. In High alcohol consumers and traffic: Proceedings of the international Workshop; High Alcohol Consumers and Traffic, 1988 Nov 28-30 pp 37-48. Inrets 1989.
20. Michiels W, La Harpe, R. Les ivresses dans le trafic a Genève: Abus occasionels ou alcoolisme. Sozial- und Praventivmedizin 1996; 41:28-35.
21. Ruud J, Gjerde H, Morland J. Alcohol consumption among convicted drunken drivers in Norway. Journal of traffic medicine 1993; 21:121-125.
22. Gjerde H, Sakshaus J, Morland J. Heavy drinking among Norwegian male drunken drivers: a study of gamma glutamyltransferase. Alcoholism: Clinical and Experimental Research 1986; 10:209-212.
23. Gjerde H, Moerland J. Concentrations of carbohydrate-deficient transferrin in dialyzed plasma from drunken drivers. Alcohol and Alcoholism 1987; 22:271-276.
24. Jaster R, and Wegener R. Diagnosis of chronic alcohol abuse in intoxicated drivers: CDT in combination with other parameters. Blutalcohol 1993; 30:257-265.
25. Lutz FU, Bausach S. Leberenzymwerte (Gamma-GT, GOT, GPT) bei trunkenheitstätern zum tatzeitpunkt. Blutalcohol 1992; 29:211-215.
26. Morgan MY, Major HE. The use of serum carbohydrate deficient transferrin in the assessment of 'high risk offenders' in Great Britain. Alcohol and Alcoholism 1996; 31:625-628.

27. Perret R, Froehlich F, Levanchy D, Henry H, Bachman C, Pécoud A, Bianchi L, Gonvers JJ. Is Carbohydrate deficient transferrin a specific marker for alcohol abuse? A study in patients with chronic viral hepatitis Alcoholism: Clinical and Experimental Research 1997; 21:1337-1342.
28. First MB, Spitzer RL, Gibbons M, Williams JBW. Structured Clinical Interview for DSM IV Axis I Disorders (SCID I) - Clinical Version. American Psychiatric Press, Washington 1997.
29. American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders, 4th edn. American Psychiatric Association Press, Washington, DC. 1994.
30. Williams JB, Gibbon M, First MB, Spitzer RL, Davies M, Borus J, Howes MJ, Kane J, Pope H G Jr, Rounsaville B., et al. () The structured clinical interview for DSM-III-R (SCID) multisite test retest reliability. Archives of General Psychiatry 1992; 49 (8):630-636.
31. Mayfield D, McLeod G, Hall P. The CAGE questionnaire: validation of a new alcoholism screening instrument. American Journal of Psychiatry 1974; 131:1121-1123.
32. Conigrave KM, Saunders JB, Whitfield JB. Diagnostic tests for alcohol consumption. Alcohol and Alcoholism 1995; 30:13-26.
33. Sillanaukee P, Seppa K, Koivula T. Acetaldehyde-modified hemoglobin as a marker of alcohol consumption: comparison of two new methods. Journal of Laboratory and Clinical Medicine 1992; 120:42-47.
34. Bean P, Daniel P. Carbohydrate-deficient transferrin; current facts and future projections for the insurance industry. On the risk 1996; 12:43-48.
35. Sackett DL, Haynes RB, Guyatt GH, Tugwall P. Clinical Epidemiology. Little Brown and Company, Boston 1991.
36. Pelt van J. Carbohydrate deficient transferrin: a new biochemical marker for chronic excessive alcohol consumption. Nederlands Tijdschrift voor Geneeskunde 1997; 141:773-777.
37. Saunders JB, Lee NK. Hazardous alcohol use: its delineation as a subthreshold disorder, and approaches to its diagnosis and management. Comprehensive Psychiatry 2000; 41:95-103.
38. Sillanaukee P, Seppa K, Koivula T. CDT by anion-exchange chromatography by RIA as a marker of heavy drinking among men. Alcoholism: Clinical and Experimental Research 1993; 17:230-233.
39. Huseby NE, Nillssen O, Kanitz RD. Evaluation of two biological markers combined as a parameter of alcohol dependency. Alcohol and Alcoholism 1997; 32:731-737.
40. Poole KW, Flynn PM, Rao AV, Cooley PC. Estimating the prevalence of drug use from self-reports in a cohort for which biological data are

available for a subsample. American Journal of Epidemiology 1996; 144:413-420.
41. Zwart de WM. Ontwikkelingen in het gebruik van alcohol en drugs. In Verslaving, Mensink C, Spruit IP eds. Bohn, Stafleu en van Loghum, Houten 1998.
42. Bijl RV, Zessen van G, Ravelli A. Psychiatric morbidity among adults in the Netherlands: the NEMESIS study. II. Prevalence of psychiatric disorders. Nederlandse Tijdschrift voor Geneeskunde 1998; 141:2453-2460.
43. Hillman A, Sykes RAD, McConnell AA. Limitation in the use of gamma glutamyltransferase estimations in alcohol-dependant subjects. Alcohol and Alcoholism 1998; 33:626-630.
44. Gilg T. Einsatzmöglichkeiten von CDT in der rechts- und verkehrmedizin. In: Klinische alcoholismusdiagnostik, Soyka, M. ed. Steinkopf, Darmstadt 1999.

Chapter 4*

TRISIALO-FE_2-TRANSFERRIN DOES NOT IMPROVE THE DIAGNOSTIC ACCURACY OF CARBOHYDRATE-DEFICIENT TRANSFERRIN AS A MARKER OF CHRONIC EXCESSIVE ALCOHOL INTAKE

Abstract

We studied the diagnostic efficiency of two commercial tests for analysis of carbohydrate-deficient transferrin (CDT) as a marker of chronic alcohol abuse in alcoholics, %CDTri-TIA (including about 50% trisialo-Fe_2-transferrin in CDT) and ChronAlcoI.D. (excluding this transferrin isoform from CDT). TLFB (Timeline-Followback) and Composite International Diagnostic Interview (CIDI) 2.1-alcohol section, which are valid, reliable and fully structured diagnostic interviews, were used as gold standard for assessment of frequency and amount of alcohol intake. %CDTri-TIA showed a distinctly reduced diagnostic sensitivity (52.8% %CDTri-TIA, 71.7% ChronAlcoI.D., $p = 0.00$) and accuracy (66.2% %CDTri-TIA, 77.9% ChronAlcoI.D., $p = 0.01$). Diagnostic specificity was statistically not different between the tests (95.8% %CDTri-TIA, ChronAlcoI.D. 91.7%, $p = 0.30$). Inclusion of trisialo-Fe_2-transferrin in CDT does not improve its diagnostic efficiency.

* Previously published: A. Korzec, T. Arndt, M. Bär, M. W. J. Koeter. Trisialo-Fe_2-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

Acknowledgements
We thank Britta Semmelroggen and Pauli Zuijdhoek for excellent technical assistance, and Lloyd Allen Jones for stylistic emendations. Six kits of %CDTri-TIA were a gift from AXIS.

INTRODUCTION

Carbohydrate-deficient transferrin (CDT) is widely used for laboratory diagnosis of chronic alcohol abuse. A review on CDT was published recently (1). There is still controversy as to the diagnostic benefit from including trisialo-Fe_2-transferrin in CDT and/or using CDT concentrations or CDT/transferrin (CDT/Tf) ratios (1-5). We investigated these issues by comparing diagnostic sensitivity, specificity and accuracy of two commercially available CDT tests: %CDTri-TIA (Axis, Norway) and ChronAlcoI.D. (Sangui Biotech Inc., U.S.A.).

MATERIALS AND METHODS

The study was in accordance with the Declaration of Helsinki of 1975, as revised in 1996 and approved by the ethical committee of the St. Lucas Andreas Hospital. The guidelines for studies of the diagnostic accuracy of diagnostic tests [6] were observed: spectrum bias was avoided by assessing consecutive patients, reviewer bias by blinding case history on alcoholism and alcohol intake to laboratory results and vice versa, verification bias by applying the criterion standards to all subjects. Each test was performed without knowledge of the CDT results obtained by the other.

Patients and Assessment of Alcohol Intake

All subjects were male. Elevated ("hazardous") drinking was defined as the level of persistent alcohol consumption being likely to results in adverse health effects: >280g ethanol/week (7, 8).

57 Controls were recruited from consecutive ambulatory psychiatric patients. 24 patients with an alcohol consumption of ≤280 g ethanol/week in each of the last 4 weeks before blood sampling and who had no Alcohol Use Disorder (AUD, "alcoholism") diagnosis were included in the control group (mean and median of ethanol consumption and age were 47 and 21 g/week and 46.5 and 45.5 years). The remaining 33 patients had an AUD diagnosis in the last year or had been drinking >280 g ethanol/week in the last 4 weeks and were excluded from the study.

101 Alcoholics were recruited from treatment facilities: 72 patients consecutively admitted to a detoxification ward and 29 consecutive patients attending an ambulatory alcoholism treatment centre. Alcoholism in this study group was defined as having an AUD diagnosis in accordance with ICD-10 (International Classification of Mental or Behavioural Disorders) or DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, 4th edition) (9,10). 53 patients with an alcohol intake of >280 g ethanol/week in each of the last four weeks and with an AUD diagnosis were included in the alcoholics group (mean and median of ethanol consumption and age were 1326 and 1113 g/week and 42.3 and 43.0 years). The remaining 48 patients were excluded from the study due to cessation of drinking in the last 4 weeks.

Widely accepted, reliable and validated diagnostic instruments were used as criterion standards in assessing alcohol intake and alcoholism (11). Alcohol intake was assessed from TLFB (12), a comprehensive retrospective self-report survey that allows the collection of information up to 12 months before the interview date. Alcohol use disorder (AUD) was assessed by means of the Composite International Diagnostic Interview (CIDI) 2.1-alcohol section, which is a valid, reliable and fully structured diagnostic interview and enables diagnosis to be computer-generated according to ICD-10 and DSM-IV criteria (9,10,13,14).

Blood samples

Blood was collected into evacuated sterile gel-tubes (Becton-Dickinson, vacutainer). Serum was obtained by centrifugation at 2600g, 5°C for 10 min. Serum aliquots were stored at -20°C. Samples were thawed only once for assay. To check if the delay between CDT analysis by %CDTri-TIA (summer 1999) and ChronAlcoI.D. (winter 1999) affected the CDT results, the %CDTri-TIA assay was repeated on a subset of 20 samples at the time of ChronAlcoI.D.: T-test for paired samples showed no significant differences between the "summer" and "winter" CDT values (mean CDT/Tf ratio 7.2% ± 7.2% (summer) and 7.0% ± 6.2% (winter), mean summer-winter difference 0.19% ± 1.18% ($p = 0.553$ two tailed), correlation of test-retest 0.997 ($p = 0.000$)). Passing and Bablok correlation (15) yielded no significant difference from zero for the intercept and from 1 for

the slope, proving the CDT concentrations to be stable (statistically non-different) between summer and winter 1999. Unaltered CDT concentrations after freezing serum samples for several months were also reported in (16-18).

%CDTri-TIA- and ChronAlcoI.D.™-Assays

%CDTri-TIA Assay was provided by AXIS Biochemicals ASA (Oslo, Norway), distributed by Orange Medical, The Netherlands and performed in Amsterdam. The test includes about 50% of trisialo-Fe_2-Tf in CDT, and reports CDT/Tf ratios. ChronAlcoI.D. Assay was provided by Sangui BioTech, Inc. (Santa Ana, U.S.A.), distributed by Biodiagnostics (Kiel, Germany) and performed in Ingelheim. The test excludes trisialo-Fe_2-Tf from CDT and reports CDT concentrations and CDT/Tf ratios. Both tests are based on anion-exchange chromatography for fractionation of CDT isoforms and non-CDT isoforms, followed by nephelometric (Array nephelometer, Beckman/Array Flexisoft program by Beckman Coulter, Mijdrecht, The Netherlands, for the %CDTri-TIA) or turbidimetric (Dynatec MR 5000 reader/Dynex Revelation 3.2 software by Dynex Technologies, Denkendorf, Germany, for the ChronAlcoI.D.) quantification of CDT. Quality control was done by internal (delivered with the test kits and analyzed in each series) and external quality control material (DGKC, Bonn; GTFCH, Heidelberg; Instand, Düsseldorf). The CV's for the low and high controls in the appropriate quality-control periods were <12.0% and <5.3% (%CDTri-TIA) and <7.5% and <7.9% (ChronAlcoI.D). Analytic specificity and precision of the ChronAlcoI.D. were assessed previously (19)]. For both tests, borderlines indicating elevated alcohol consumption have been suggested: 5-6% CDT for the %CDTri-TIA (test instructions) and 2.5-2.7% CDT or 100-110 mg CDT/L for the ChronAlcoI.D. (20). Taking into account the social consequences of false-positives regarding chronic alcohol abuse, we used the upper limits of these borderlines as decision criteria (cut-offs): 6% CDT for the %CDTri-TIA and 2.7% CDT or 110 mg CDT/L for the ChronAlcoI.D. (Table 1).

Statistics

Diagnostic sensitivity, specificity, accuracy, ROC curves, confidence intervals (CI) and inter-assay variation coefficients were computed

with the statistics software SPSS base 10.0 for Windows NT (SPSS Inc., Chicago, U.S.A.). Differences in the criteria of diagnostic efficiency between %CDTri-TIA and ChronAlcoI.D. were checked for significance by the McNemar test for paired samples. Confidence intervals were calculated with a formula given in (21). P-values <0.05 indicate significant differences.

RESULTS AND DISCUSSION

The parameters of diagnostic efficiency obtained at cut-offs of 6% CDT for the %CDTri-TIA and 2.7% for the ChronAlcoI.D. assay are summarised in Table 1. Compared with %CDTri-TIA, ChronAlcoI.D. showed significantly higher diagnostic sensitivities and accuracies. There were no significant differences in the diagnostic specificities between %CDTri-TIA and ChronAlcoI.D. (Table 1).

Table 1.
Parameters of diagnostic efficiency of %CDTri-TIA and ChronAlcoI.D. for 53 patients (alcohol intake of >280g ethanol/day and alcoholism diagnosis) and 24 controls (alcohol intake ≤280 g ethanol/day).

	ChronAlcoI.D.	%CDTri-TIA			
	cut-off 2.7%	cut-off 6.0%	Difference	95% CI[a]	p[b]
Diagnostic sensitivity	71.7%	52.8%	18.9%	7.2% - 30.6%	0.00
Diagnostic specificity	91.7%	95.8%	-4.1%	-12.0% - 3.8%	0.30
Diagnostic accuracy	77.9%	66.2%	11.7%	2.9% - 20.5%	0.01

[a] Confidence interval for difference between the two tests

[b] p values based on McNemar test without continuity correction (21).

Using cut-offs of 5% CDT (instead of 6% CDT, Table 1) for the %CDTri-TIA and 2.5% (instead of 2.7%, Table 1) for the ChronAlcoI.D. improved the diagnostic sensitivities (from 52.8% to 69.8% for %CDTri-TIA, from 71.7% to 81.1% for ChronAlcoI.D.) and accuracies (from 66.2% to 75.3% for %CDTri-TIA, from 77.9% to 84.4% for

ChronAlcoI.D.) for both tests, but diminished the diagnostic specificity of the %CDTri-TIA assay (from 95.8% to 87.5%). The diagnostic specificity of the ChronAlcoI.D. was unaffected (91.7% at the low and the high cut-off).

Compared with CDT/Tf ratios (ChronAlcoI.D.), absolute CDT concentrations obtained by the same assay (ChronAlcoI.D.) showed a significantly reduced sensitivity (45.3% for absolute vs 71.7% for relative CDT concentrations; 95% CI -40.0% - - 13.4%, p=0.00) and accuracy (61.0% for absolute vs 77.9% for relative CDT concentrations; 95% CI -26.6% - -7.1%, p=0.00).

Our findings are in accordance with an earlier study (2), comparing the %CDT-TIA (identical with %CDTri-TIA, including about 50% of trisialo-Fe_2-Tf, measuring CDT/Tf ratios) and the CDTect (excluding trisialo-Fe_2-Tf, measuring absolute CDT concentrations). Compared with CDTect, %CDT-TIA showed an overall reduced diagnostic accuracy for detecting alcohol abuse in men, this being mainly due to a diminished diagnostic sensitivity (2). For ChronAlcoI.D., absolute CDT concentrations (as used by the CDTect) showed an overall weaker diagnostic accuracy when compared with the corresponding CDT/Tf ratios (see above). Thus, the findings in (2) cannot solely be due to the different units used by the two tests (% of total Tf by the %CDT-TIA and U/L by the CDTect). CDTect (22) and the ChronAlcoI.D. (19) show a similar analytic specificity. The fact that both tests exclude trisialo-Fe_2-Tf from CDT makes the greatest difference in comparison with the %CDTri-TIA. Thus, it is more likely that the diminished diagnostic accuracy of so-called "trisialo-tests" (%CDTri-TIA or %CDT-TIA[1]) is due to the inclusion of trisialo-Fe_2-Tf in CDT. This conclusion is supported by findings by others (3,5,23): No increase of trisialo-Fe_2-Tf concentration after chronic alcohol consumption, but significant increases for asialo-Fe_2-Tf (by 219% of its normal serum concentration), monosialo-Fe_2-Tf (28% increase) and disialo-Fe_2-Tf (148% increase) were described in (23). Increased concentrations of asialo- and disialo-Fe_2-Tf in serum samples with pathological CDT/Tf ratio and almost identical trisialo-Fe_2-Tf concentrations in serum samples with normal and pathological CDT/Tf ratio were reported in (3). Classi-

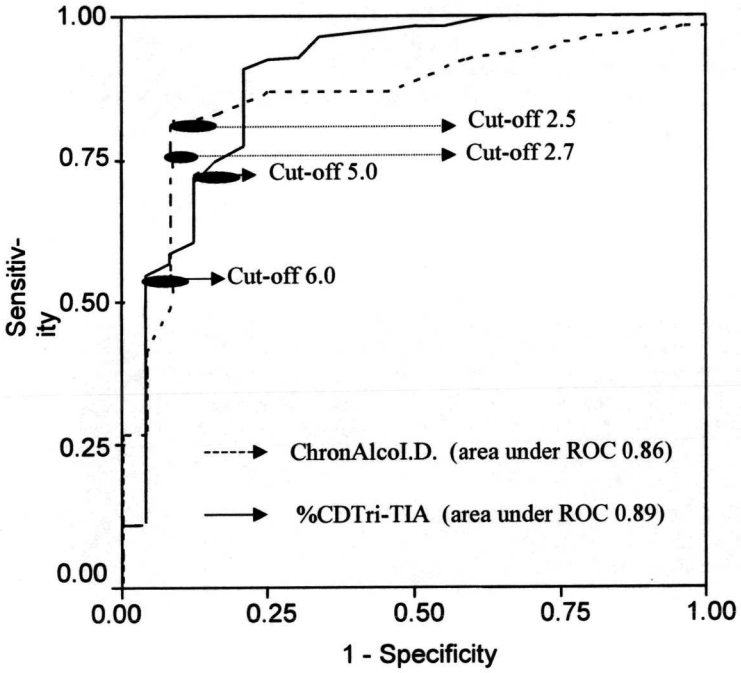
[1] Unfortunately, the new product by Axis, excluding trisialo-Fe_2-Tf from CDT and using the common CDT definition has the same name, %CDT TIA.

fying relative CDT concentrations obtained by ChronAlcoI.D., %CDT TIA (including 50% of trisialo-Fe_2-Tf into CDT) and HPLC as either normal or elevated, Lipkowski et al. found 22% discrepancies between %CDT TIA and HPLC, but only 9% between ChronAlcoI.D. and HPLC (5). The authors strongly recommend not to include trisialo-Fe_2-Tf into CDT.

The significant differences in diagnostic sensitivity and diagnostic accuracy between %CDTri-TIA and ChronAlcoI.D. were not matched by analogous differences in the corresponding areas under the ROC curve (AUC, see Fig. 1). This discrepancy is most probably caused by an intersection of both curves outside the clinically important part: In the cut-off area where in normal clinical practice the diagnostic measurements or decisions are made (at the recommended cut-offs), %CDTri-TIA performs worse than ChronAlcoI.D. In the cut-off area where diagnostic decisions will never be made (because of the corresponding unacceptable low diagnostic specificities), %CDTri-TIA performs better than ChronAlcoI.D. and thus gains AUC. However, ROC analysis seems less suitable for comparing the tests under study because this method assumes that the choice of cut-off is made only from data plotted, without information from previous published work suggesting what is the best cut-off value (24).

This is in contrast to the concept of our study (use of non-arbitrary, recommended and widely accepted cut-offs). Comparisons of sensitivities of diagnostic tests are usually made on the same level of specificity or vice versa. In our study, this approach would mean comparison of the test performance for one test at the recommended (optimal and widely used) and for the other at a non-recommended (non-optimal and never used) cut-off. Therefore, we have assessed the diagnostic efficiency of both tests at their recommended cut-offs. The relatively small number of controls may limit the significance of our study. However, it does not explain the significant differences in diagnostic accuracy between %CDTri-TIA and ChronAlcoI.D.

Figure. 1.
ROC plot for two commercial CDT tests for laboratory diagnosis of chronic excessive alcohol intake for 53 alcoholics and 24 healthy controls: %CDTri-TIA (including about 50% of trisialo-Fe_2-transferrin in CDT), ChronAlcoI.D. (excluding this transferrin isoform).



Conclusion

If trisialo-Fe_2-Tf increases at all after chronic alcohol abuse, its proportional change might be less than that for the common CDT isoforms. If this is true, the comparably less affected but large amounts of trisialo-Fe_2-Tf might mask the alcohol-induced increases in the CDT isoforms and thus lower the diagnostic sensitivity of CDT. As a consequence, the production of so-called "trisialo-tests" by Axis has been terminated recently.

References

1. Arndt T. Carbohydrate-deficient transferrin as a marker of chronic alcohol abuse: a critical review of preanalysis, analysis, and interpretation. Clinical Chemistry 2001; 47:13-27.
2. Viitala K, Lähdesmäki K, Niemela O. Comparison of the Axis %CDT TIA and the CDTect method as laboratory tests of alcohol abuse. Clinical Chemistry 1998; 44:1209-15
3. Dibbelt L. Does trisialo-transferrin provide valuable information for the laboratory diagnosis of chronically increased alcohol consumption by determination of carbohydrate-deficient transferrin? Clinical Chemistry 2000; 46:1203-5.
4. Heggli D-E, Aurebekk A, Granum B, Westby C, Løvli T, Sundrehagen E. Should tri-sialo-transferrins be included when calculating carbohydrate-deficient transferrin for diagnosing elevated alcohol intake? Alcohol and Alcoholism 1996; 31:381-4.
5. Lipkowski M, Dibbelt L, Seyfarth M. Is there an analytical or diagnostic advantage from including trisialo transferrin into the fraction of carbohydrate-deficient transferrin? Lessons from a comparison of two commercial turbidimetric immunoassays with the carbohydrate-deficient transferrin determination by high-performance liquid chromatography. Clinical Biochemistry 2000; 33:635-41.
6. Bruns DE, Huth EJ, Magid E, Young DS. Toward a checklist for reporting of studies of diagnostic accuracy of medical tests. Clinical Chemistry 2000; 46:893-5.
7. Carrington Reid M, Fiellin DA, O'Connor PG. Hazardous and harmful alcohol consumption in primary care. Archives of Internal Medicine 1998; 159:1681-1689.
8. Piccinelli M, Tessari E, Bortolomasi M, Piasere O, Semenzin M, Garzotto N, Tansella M. Efficacy of the alcohol use disorders identification test as a screening tool for hazardous alcohol intake and related disorders in primary care: a validity study. British Medical Journal 1997; 314: 420-4.
9. ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines. Geneva: World Health Organization, 1992.
10. American Psychiatric Association: Diagnostic and statistical manual of mental disorders, 4th ed. Washington: American Psychiatric Association, 1994
11. Allen JP, Fertig JB, Litten RZ, Sillanaukee P, Anton RF. Proposed recommendations for research on biochemical markers for problematic drinking. Alcoholism: Clinical and Experimental Research 1997; 21:244-247.

12. 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.
13. 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 and Alcohol Dependence 1997; 47:161-9.
14. 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 and Alcohol Dependence 1997; 47:195-205.
15. Passing H, Bablok W. A new biometrical procedure for testing the equality of measurements from different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry. Journal of Clinical Chemistry and Clinical Biochemistry 1983; 21:709-20.
16. Stibler H, Borg S, Joustra M. Micro anion exchange chromatography of "carbohydrate-deficient transferrin" (CDT) in serum in relation to alcohol consumption (Swedish patent 8400587-5). Alcoholism: Clinical and Experimental Research 1986;10:535-44.
17. Mårtensson O, Brandt R. Stability of CDT in vivo and in vitro. Alcohol and Alcoholism 1997; 32:384.
18. Arndt T. Kohlenhydrat-defizientes Transferrin (CDT): die derzeit spezifischste Kenngröße chronischen Alkoholmißbrauchs. Journal of Laboratory Medicine 1999; 23:392-406.
19. Hackler R, Arndt T, Helig-Rolig A, Kropf J, Steinmetz A, Schaefer JR. Investigation by isoelectric focusing of the initial carbohydrate-deficient transferrin (CDT) and non-CDT transferrin isoform fractionation step involved in determination of CDT by the ChronAlcoI.D. assay. Clinical Chemistry 2000; 46:483-492.
20. Arndt T, Behnken L, Martens B, Hackler R. Evaluation of the cut-off for serum carbohydrate-deficient transferrin as a marker of chronic alcohol abuse determined by the ChronAlco I.D.™-Assay. Journal of Laboratory Medicine 1999; 23: 507-10.
21. Armitage P, Berry G, eds. Statistical methods in medical research, Oxford: Blackwell Science Ltd, 1994:126.
22. Arndt T, Hackler R, Kleine TO, Gressner AM. Validation by isoelectric focusing of the anion-exchange isotransferrin fractionation step involved in determination of carbohydrate-deficient transferrin by the CDTect assay. Clinical Chemistry 1998; 44:27-34.

23. Mårtensson O, Härlin A, Brandt R, Seppä K, Sillanaukee P. Transferrin isoform distribution: gender and alcohol consumption. Alcoholism: Clinical and Experimental Research 1997; 21:1710-5.
24. Kazmierczak SC. Statistical techniques for evaluating the diagnostic utility of laboratory tests. Clinical chemistry and laboratory medicine 1999; 37:1001-9.

Chapter 5*

CONFIRMING DIAGNOSIS OF HAZARDOUS AND HARMFUL ALCOHOL USE

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

Abstract

Objective. Conventional tests for alcoholism fail to confirm hazardous and harmful alcohol use (HHAU) accurately and objectively. In this study, we validated a Bayesian Alcoholism Test (BAT) for confirming the diagnosis of HHAU.

Study design and setting. 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 treatment seeking alcoholics, non-treatment seeking heavy drinkers and controls. Main outcome measures were test sensitivity and specificity, likelihood ratio's and receiver-operating characteristic (ROC) curves.

Results. Comparing alcoholics and controls, sensitivity of BAT (94%) was significantly higher than CDT (63%) and GGT (73 %). The area under the ROC 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).

* Under review: Korzec A, de Bruijn H, van Lambalgen M. Confirming Diagnosis of Hazardous and Harmful Alcohol Use: Diagnostic accuracy of a computer assisted diagnostic system compared to conventional markers of alcoholism

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).
Conclusion. BAT has better diagnostic properties than CDT and GGT for confirming HHAU.

INTRODUCTION

Alcoholism refers to a heterogeneous set of disorders. This set of disorders can be divided in two overlapping conceptual domains. The first domain contains psychiatric diagnoses and emphasizes addiction, social, psychological and physical damage such as alcohol dependence and alcohol abuse, often called Alcohol Use Disorders (AUD). The second domain emphasizes drinking patterns defined by amount of drinking and their effects on physical health, often referred as hazardous alcohol use or harmful use (HHAU) (1).

Alcoholism has severe consequences for society. 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 (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 under-diagnosis are denial on the part of patients (10,11), insufficient sensitivity of screening instruments in detecting patients with less severe alcoholism (12), 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), insurance (21), occupational (22) and pre-operative settings (23,24), there is a strong need of a 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 are a large number of diagnostic tests and when the relationship between the disease and the result of tests is of a probabilistic nature. Although the probabilistic computations involved are complex, with the advent of so-called Bayesian networks (25), mathematical and computational technology has now progressed so far as to make an expert system of the size we need 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 is 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 ambulatory psychiatric patients (Sint Lucas Andreas Hospital, Amsterdam). Only patients 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, were included in the control group (n=47). 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 admitted to a detoxification ward (Jellinek clinic, Amsterdam) and 29 patients attending an ambulatory alcoholism treatment center (Brijder stichting, Zaandam). Only patients with harmful use [defined as an alcohol intake of > 560 g ethanol/week (28)] in the last 4 weeks before 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 relation between alcohol and health and of the object of our study. We recruited a total of 68 men, from which 57 drank more than 28 alcoholic units per week. The remaining 11 subjects drank less than 28 alcoholic units per week during the last 90 days. Other characteristics of this group have been described previously (31).

One psychiatrist and six psychiatric residents who received prior training about the instruments collected the data. For each subject all data were collected on the same day. The readers of the criterion

assessments were blind to the results of the laboratory 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
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 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 fe-

male populations, primary care populations, clinical populations etc. 72 studies were included. Two reviewers appraised all articles for methodological content and results. If, on certain prevalence or condition, no studies were found, we obtained estimates 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://staff.science.uva.nl/~michiell/, for a summary table of data that were used for constructing 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.

An important fact to consider is that BAT does not contain DSM-IV AUD criteria. The only alcohol problems related questions within BAT are the CAGE questions and question about the level of response to alcohol (LRA). For the rest, all the components of BAT are either objective or contain questions about somatic data, which are unlikely to be lied about.

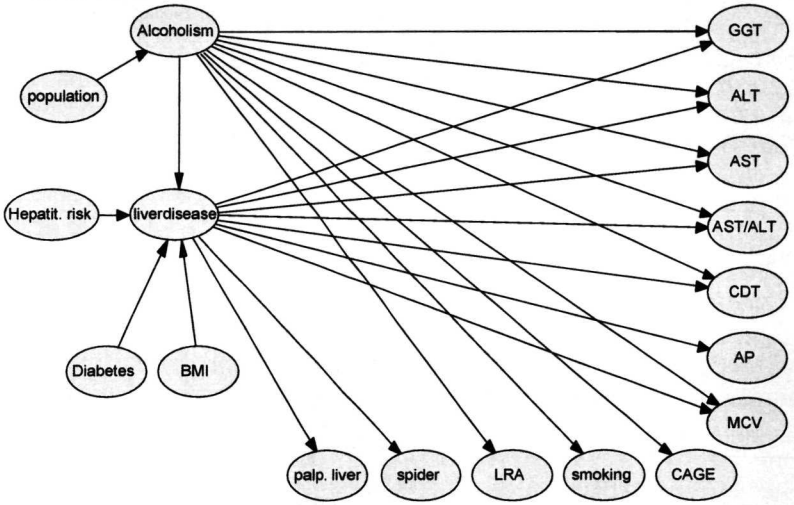
3. BAT components

All BAT components are parts of usual clinical practice when confirming an alcoholism diagnosis. The components concern clinical signs, biochemical test and some additional questions concerning differential diagnostic possibilities for elevated liver enzymes.

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.

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 liver disease	Fatty liver	Hepatitis	Liver cirrhosis
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

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: Cut down drinking, Annoyed by criticism about drinking, Guilty feelings about drinking, and Early 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), gammaglutamyltransferase (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: (ChronAlcol.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: 50, AST: 45, 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 test (40). MCV cutoff was based on several studies comparing social drinkers and alcoholics (41-43). The laboratory where the tests were performed recommended the other cutoffs. All biochemical tests were performed in the laboratory of the Sint Lucas Andreas Hospital, except CDT, which was performed at *bio*scientia, Ingelheim, Germany.

c. 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.

In the case of missing data (out of 182 included patients: CDT: 7; Cage 6: smoking: 6; LRA: 26), the data, which were present, were fed into BAT. For these subjects, BAT generated a result based on the data that were known.

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 (CI) for sensitivity, and specificity were computed using Wilson's method (44). Confidence intervals of likelihood ratios were computed using the score method (44). Difference between likelihood ratio's en 95% CI was computed using the ratio of two standardized ratio (44). Receiver operating characteristic 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 mean age of the heavy drinkers group (49.3) was significantly higher than that of the alcoholics (43.6).

Cutoff level BAT

BAT outcomes are, in principle, probabilities of a patient having HHAU. The clinician must convert the continuous probabilities yielded by BAT into a binary decision whether he considers the patient to suffer from HHAU or not. It is up to the clinician when he adopts a rule to convert a probability into a yes/no statement. The

sensitivity and specificity of BAT obviously depend on the decision rule chosen.

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)[1]	43.6 ± 7 (28-58)[2]
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)
[1]Age difference between controls and heavy drinkers and controls and harmful users n.s.
[2]Significant age difference between heavy drinkers and alcoholics: 5.7 (95% CI 2.7-8.7)

It must also be noted that the diagnostic markers found in the literature are represented by dichotomous variables. Hence, also for purposes of comparison BAT must be reformulated as a binary test.

The cutoff range which gave the best accuracy (total of true positives and true negatives divided by all subjects, here 95.6%) was between 41% and 50%. Because we aimed to design a confirmation test we chose 50% as cut off level.

We emphasize however, that this decision rule is to a certain extent arbitrary and that some contexts may require a different rule, e.g. if higher specificity is desired.

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).

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

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

( ) 95% confidence intervals
* 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)

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. 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

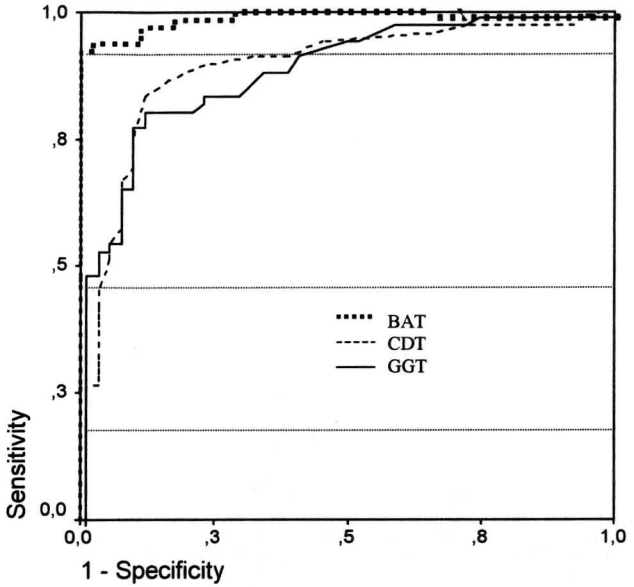
Of the 67 alcoholics, BAT did not recognize four subjects, against 24 subjects having normal CDT values and 18 subjects having normal GGT values.

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 was significantly higher ($p < 0.005$) than for CDT and for GGT. 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%.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).

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

Table 3.
Sensitivities, specificities and likelihood ratios of BAT without Cage, BAT with only CDT and GGT, CDT and GGT for diagnosing harmful and hazardous alcohol use, comparing treatment seeking alcoholics and controls

	Sensitivity	Specificity	Likelihood ratio +	Likelihood ratio -
BAT without CAGE	85 (74,7 - 91,7)	97,9 (89 -100)	40 (7,6 - 226)	0,15 (0,09 - 0,26)
BAT with only CDT-GGT	58 (46-69)	97,9 (89 -100)	27 (5-156)	0,43 (0,31- 0,55)
CDT(n=110)	63.1 (50.9 -73.8)	93.3 (82.1- 97.7)	9.5 (3.5 - 27.9)	0.40 (0.30 - 0.54)
GGT(n=114)	73.1 (61.5 -82.3)	91.5 (80.1-96.8)	8.6 (3.6 - 22.0)	0.29 (0.19-0.43)
Difference and 95% CI for the difference BAT without Cage - CDT	21.5†* (8,4-33,9)	4.4† (-6 - 15.9)	4.2 (4.1-4.3)*	0.38 (0.32-0.46)*
Difference and 95% CI for the difference BAT without Cage - GGT	11,9* (0-23,7)	6.4 (- 2.9 - 17.6)	4.6 (4.5-4.8)*	0.52 (0.42-0.63)*
Difference and 95% CI for the difference BAT- BAT with only CDT and GGT	9* (1,4-18,1)	0	1.6 (1.6-1.6)*	0.14 (0.11-0.19)*
Difference and 95% CI for the difference BAT without Cage - BAT with only CDT and GGT	26,9* (15,6-37,6)	0	1.5 (1.4-1.5)*	0.36 (0.30-0.43)*

( ) 95% confidence intervals

* 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)

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 4 from the entropy values, BAT gives more information than GGT.

Table 4.
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 (n= 68)*	BAT+ (n=23) 100%	CDT+ (n=23) 100%	GGT+ (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 (95% CI)	0.6921 (0.6723 - 0.7130)	0.8367 (0.6334-1)	0.9566 (0.7321-1)

*3 missing values of CDT

Correlation alcohol intake and test results

Using pooled data of all 182 subjects, BAT had a significantly 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).

BAT without Cage, and Bat with only CDT and GGT

As we plan to use BAT in populations that are prone to deny alcohol problems, we computed BAT results when Cage is not used as

component. The reason for doing so is that Cage is the only BAT-item that is evidently influenced by the willingness of the subject to tell the truth about alcohol problems. In comparison to BAT, the sensitivity of BAT without Cage sensitivity dropped from 94% to 85 %. It was still significantly better than CDT, but not better than GGT. Furthermore, we checked what the BAT results would be if we would simplify it by only using CDT and GGT as BAT components. The sensitivity of BAT with only CDT and GGT dropped to 58% (Table 3).

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 our population, this test has better diagnostic properties than the regular tests. Both positive and negative likelihood ratio's of BAT were superior to those of CDT and GGT. Therefore, BAT has better properties to rule in or rule out the diagnosis of HHAU, which makes it more appropriate for confirming the diagnosis.

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 HHAU (48) is that BAT can be easily accommodated for other populations 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.

First, 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. As all patients and subjects collaborated only if they agreed to an examination, there might be a selection bias, which could reduce external validity of our results. Our intent was to enable to confirm the diagnosis also in denying subjects. One could hypothesize that the denying subjects would not agree to participate with this study. However, there is no reason to assume that this selection would favor BAT and impede CDT or GGT. Another arguments that selection bias was not of a big impact on our study is that the sensitivity and specificity values of CDT and GGT found in our study, 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.

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 that this part of our study did not have enough power to detect the difference. A second reason might be that wine drinkers are not representative of heavy drinkers in general. They smoke less than the general population and they might have less drinking problems, as indicated by Cage, than usual populations of heavy drinkers (31). 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. The examination takes 30 minutes. Feeding in the data and producing the BAT results takes approximately a minute 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. Archives of Internal Medicine 1999; 159:1681-1689.
2. Reynaud M, Gaudun-Colombel AF, Le Pen C. Two methods of estimating costs linked to alcoholism in France. Alcohol and Alcoholism 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. Journal of Studies on 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. Archives of Family Medicine. 2000; 9:814-821.
9. Fiellin DA, Carrington Reid M, O'Connor PG. Screening for alcohol problems in primary care. Archives of Internal Medicine 2000; 160:1977-1989
10. Samet JH, O'Connor PG. Alcohol abusers in primary care: readiness to change behavior. The American Journal of Medicine 1998; 105:302-306
11. Ness DE, Ende J. Denial in the medical interview; recognition and management Journal of American Medical Association 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 and Alcoholism 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. British Medical Journal 2002; 325:870-872.
14. Feinstein AR. Clinical Epidemiology. Philadelphia: WB Saunders. 1985.
15. Kristenson H. Methods of intervention to modify drinking patterns in heavy drinkers. Recent Developments in Alcoholism 1987; 5:403-423.

16. Wick HD. Alcohol problems in general practice - the value of biological markers.
Schweizerische Rundschau fur Medizin Praxis 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. American Journal of Medicine 1988; 85:466-471.
18. Conigliaro J, Lofgren RP, Hanusa BH. Screening for problem drinking. Impact on physician behavior and patients drinking habits. Journal of general internal medicine 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 and Alcoholism 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 and Alcoholism 2002; 37(3):289-94.
23. Kitchens JM. Does this patient have an alcohol problem? Journal of the American Medical Association 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 subthreshold 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. Alcoholism: Clinical and Experimental Research 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, 4th edn. American Psychiatric Association Press, Washington, DC. 1994.
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. European Addiction Research 2003 ;9:182-187.
32. Allen JP, Fertig JB, Litten RZ, Sillanaukee P, Anton RF. Proposed recommendations for research on biochemical markers for problematic drinking. Alcoholism: Clinical and Experimental Research 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 and Alcohol Dependence 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 and Alcohol Dependence 1997; 47:161-9.
36. Mayfield D, Mcleod G, Hall P. The CAGE questionnaire: validation of a new alcoholism screening instrument. American Journal of Psychiatry 1974; 131:1121-1123.
37. Conigrave KM, Saunders JB, Reznik RB, Whitfield JB: Prediction of alcohol-related harm by laboratory test results. Clinical Chemistry 1993;39:2266-2270.
38. Korzec A, Arndt T, Bär M, Koeter MWJ: Trisialo-Fe_2-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. Clinical Chemistry 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 and Alcoholism 2001;36:406-412.
41. Chalmers DM, Chanarin I, Macdermott S, Levi AJ. Sex related differences in the hematological effects of excessive alcohol consumption. Journal of clinical pathology 1980;33:3-7.

42. Sillanukee P, Seppa K, Lof K, Koivula T. CDT by anion-exchange chromatography followed by RIA as a marker of heavy drinking among men. Alcoholism: Clinical and Experimental Research 1993;17:230-233.
43 Seppa K, Sillanaukee P, Koivula T. Abnormalities of hematological parameters in heavy drinkers and alcoholics. Alcoholism: Clinical and Experimental Research 1992;16:117-121.
44. 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. Alcoholism: Clinical and Experimental Research 2001;25:309-316.
49. Conigrave KM, Saunders JB, and Whitfield JB. Diagnostic tests for alcohol consumption. Alcohol and Alcoholism 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.

Chapter 6*

DIAGNOSING ALCOHOLISM IN DRIVERS UNDER INFLUENCE: comparing different diagnostic procedures

Abstract

Background In several European countries Drivers Under Influence (DUI) suspected of alcoholism, are mandatory referred for diagnostic examination. However, no generally accepted diagnostic procedure is available to establish the diagnosis alcoholism in this population.

Objective: The aim of this study was to compare four diagnostic procedures for confirming the diagnosis alcoholism: a standard fully structured interview (CIDI), a restrictive diagnostic procedure (RDP), a Bayesian alcoholism test (BAT) and a standard clinical diagnostic procedure (CDP).

Subjects and Methods Subjects were 116 DUI's referred for a diagnostic examination. Data were collected for all diagnostic procedures (CIDI, RDP, CDP and BAT). Results of the four diagnostic procedures were compared both quantitatively and qualitatively.

Results BAT identified 52,6 % of the total population as alcoholic, CDP 50%, RDP 27,8% and CIDI 7,8%. The prevalence of BAT differed significantly from RDP and CIDI, but not from CDP. The agreement between BAT and CDP was high (Kappa 0,78, 95% CI: 0,66- 0,89). All diagnostic procedures were significantly correlated with the average amount of drinking (alcohol units/week). Only BAT was significantly correlated with the highest number of alcohol units in one day.

Submitted: Korzec A, de Bruijn, Koeter MWJ, van den Brink W. Diagnosing alcoholism in drivers under influence:comparing different diagnostic procedures

Conclusion Different diagnostic procedures for diagnosing DUI result in widely ranging prevalence rates of alcoholism. The results of BAT and CDP are most closely related to prevalences found in standard clinical practice. CIDI results in unlikely low prevalence rates. The advantage of BAT is that it is more objective, each subject is diagnosed in the same - objective - way.

INTRODUCTION

Alcoholism refers to a heterogeneous set of disorders. Two overlapping conceptual frameworks are used to approach this set of disorders. The first approach comprises the psychiatric diagnoses alcohol dependence and alcohol abuse (Alcohol Use Disorders: AUD), and emphasizes loss of control and alcohol related social, psychological and physical consequences. The second approach emphasizes drinking patterns defined by amount of drinking and their effects on physical health, and is often referred as hazardous alcohol use (HAU) (1). The choice of the most appropriate diagnostic approach depends on the goal of the diagnostic procedure.

Diagnosing alcoholism is not different from diagnosing other diseases. In making a diagnosis, clinicians begin with an estimate of the a priori probability about the presence of a disease. Depending on this estimate, the clinician uses a diagnostic test for screening or for confirmation. In screening the aim is case finding and finding as many cases as possible. High sensitivity of the test (few false negative test results) is more important than high specificity (few false positive test results). In contrast, confirmation aims at a definite diagnosis. Very high specificity (very few false positive test results) is more important than high sensitivity (few false negative results).

In evaluating diagnostic procedures for alcoholism in legal settings, one must consider the differences in the context, compared to health care settings. In health care, the main diagnostic aim is to enhance health. Therefore, it is important to identify all alcoholic patients. In order to minimize the risk of missed diagnoses (and miss treatment possibility) a high sensitivity of diagnostic procedures is important.

In a forensic setting, such as a mandatory medical examination in a Drivers Under Influence (DUI) population, the aim is not to enhance health but to enhance traffic safety. Because diagnosis may be challenged in court, diagnosis is restricted to certain or definite cases. In legal settings, high specificity of diagnostic tests is important, because incorrect diagnoses have unacceptable legal consequences.

Understanding the legal dilemma is essential in choosing between the different diagnostic procedures. The dilemma is to find a bal-

ance between two opposite aims. On the one side, the requirement is to enhance traffic safety (for the common good) - each missed diagnosis endangers traffic safety and may have serious consequences for other people. On the other side, the requirement is to protect the rights of the individual - each incorrect diagnosis has serious consequences (for the individual) such as losing employment after being disqualified from driving.

Another issue to consider in evaluating diagnostic procedures in alcoholism is that most screening instruments for AUD and HAU are unfit if patient deny their drinking. In forensic settings, such as a mandatory diagnostic evaluation of alcoholism in DUI's, one can expect a very high probability of denial (2).

Summing up, the ideal diagnostic procedure for confirming HAU or AUD in forensic settings has to meet two requirements: (A) the diagnostic procedure should be highly specific and based on plausible reasons for the diagnosis alcoholism, in order to justify the confiscation of the drivers license; (B) the diagnostic procedure should identify cases objectively, independent of the subjects' willingness to admit alcohol use and related problems. Biochemical tests for alcoholism can be used in order to deal with the problem of denial (3). However, as these markers show only poor to moderate specificity, they do not fulfill the first requirement and therefore cannot be used in forensic settings without extra information to confirm the diagnosis of alcoholism.

In order to deal with the above-mentioned requirements, we developed a Restrictive Diagnostic Procedure (RDP), based on biochemical, clinical and psychological instruments (4). RDP is a diagnostic algorithm, based on clinical history and laboratory markers, with several decision points, aiming to identify only definite cases of AUD and HAU. RDP however has three limitations: (a) RDP is mostly based on clinical experience and not on hard empirical data; (b) in earlier research, RDP identified only 31 % of a DUI population as alcoholics (4); and (c) RDP does not make use of the a priori prevalence of alcoholism in the population in which the diagnostic test is being performed.

In order to deal with these limitations of RDP, we devised a confirmatory diagnostic instrument, the Bayesian Alcoholism Test (BAT). BAT is an expert system. This expert system yields a prob-

ability for the patient to suffer from HAU after the imputation of data about the estimated prevalence of the disease in the target population, and data of a particular patient (e.g. values of selected blood markers and objective clinical signs). In an earlier paper, we described the development and first validation study of BAT in three populations: alcoholics, heavy drinkers and controls. We found that BAT had better confirmation properties than conventional biochemical markers for identifying HAU (5).

In the present study, RDP and BAT will be compared with two standard diagnostic procedures: (1) a fully structured interview (Composite International Diagnostic Interview: CIDI), and (2) a routine clinical approach (Clinical Diagnostic procedure: CDP) in a population of DUI's.

SUBJECTS AND METHODS

The medical ethics committee of the St. Lucas Andreas Hospital approved the study protocol. All participants of the study gave informed consent; the research was carried out according to the provisions of the Declaration of Helsinki of 1975, as revised in 1996.

Subjects

The study population consisted of 177 consecutive male DUI's who were referred for diagnostic evaluation between June 1998 and August 1999 after driving under the influence of alcohol. Of these, 61 subjects were excluded because they refused to participate in the study (n=50) or because of incomplete clinical or biochemical data (n=11), leaving a study population of 116.

In accordance with Dutch traffic regulations the following 4 groups were included for referral and examination: (1) DUI's with at least one arrest with a Blood Alcohol Level (BAL) ≥2.1‰ (high BAL group n = 29); (2) DUI's with al least four arrests with any BAC above 0.5 ‰ within 5 years, or three such DUI arrests and earlier educational course on drinking driving (many arrests group n = 11); (3) people who refused to cooperate with breath analysis (refusal group n = 34); and (4) former DUI's who apply for re-granting their driving license after losing their license for 12 months because of a

diagnosis of alcoholism (n = 42). The first three groups are mandatory referrals and are summed in the tables as first-examination group (n=74). The last group is self-referred and is called the re-examination group (n=42).

Standardized clinical data collection

All DUI's were examined and diagnosed by the same physician (AK). The examination was recorded in a standardized clinical report, of which a part was used for the legal procedure on behalf of the Dutch Traffic Test organization. The clinical report of each subject consisted of history taking, instruments to assess AUD, alcohol intake and patterns of alcohol use, physical examination and biochemical measurements.

History taking focused on clinical signs of alcoholism and on the presence of non-alcoholic causes that can elevate the same biochemical markers that are raised in alcoholism. The latter included questions about current and past illness, specifically diabetes, liver diseases, blood transfusions and intravenous drug use, anemia, and the use of prescribed drugs.

Instruments to assess alcoholism symptoms. To assess whether the subject had AUD according to DSM IV or ICD 10 the Composite International Diagnostic Interview (CIDI-2.1), section J on alcohol was administered. The CIDI is a reliable and valid, fully structured diagnostic interview that enables diagnosis to be computer-generated according to ICD-10 and DSM-IV- criteria (7,8). For the screening of alcohol problems the CAGE questions were used (acronym based on its four questions: Cut down drinking, Annoyed by criticism about drinking, Guilty feelings about drinking, and Early drink [drinking in the morning]) (9).

Alcohol intake and patterns of alcohol use over the last three months were assessed using the Timeline Followback (TLFB-90). This is a retrospective self-report survey that allows for the collection of information on drinking behavior in the 90 days before the assessment (10,11). The amount of alcohol was documented in alcohol units (AU); a standard drink in the Netherlands containing approximately 10 grams of ethanol.

Biochemical measurements:
Venous blood samples for determination of hemoglobin (Hb), Hematocrit (Ht), red blood cell count (E), Mean Cell Volume (MCV), carbohydrate-deficient transferrin (CDT), Gamma glutamyltransferase (GGT), aspartate aminotransferase (AST), alanine aminotransferase (ALT) and alkaline phophatase (AP) were taken. Serum samples for CDT were frozen within 4 hours after collection and stored at - 20°C until use. CDT was analyzed in duplicate, using a commercial kit, ChronAlcol.D. (Sangui Biotech Inc., U.S.A.). This test has been validated both analytically and clinically (12,13).

Measurement of serum GGT, ALT, AST and AP was executed within 4 hours with VITROS (Ortho Clinical Diagnostics) at 37°C. Hb, Ht, E, MCV were kept at room temperature and analyzed within 4 hours with Technicon H2 analyzer, Bayer. For ALT, AST, GGT, MCV and AP we used the cut off recommended by the clinical laboratories in the region where the test were performed. The reference limit of CDT was ≥ 3 U/l, AP ≥ 135U/l, GGT ≥ 65 U/l, ALT ≥ 50 U/l, AST ≥ 45 U/l, MCV ≥ 97 fl.
Physical examination included blood pressure, liver palpation, observation of skin abnormalities indicative for liver dysfunction (spider naevi, erythema palmare) and neurological examination in order to find dysfunctions indicative of polyneuropathy or withdrawal symptoms.

Diagnostic procedures
As the diagnostic window of biochemical markers does not exceed 2-3 months, the emphasis in the different diagnostic procedures is on current diagnosis (from the time of the examination until 3 months backward).

Data from clinical reports of every subject were processed in five diagnostic procedures: CIDI, RDP, BAT, BAT without Cage and CDP. The diagnostic procedures are not fully independent but are essentially different. The differences are summarized in table 1.
1. Composite International Diagnostic Interview (CIDI) is described above.
2. Restrictive Diagnostic Procedure (RDP) uses a simple algorithm and intents to identify only definite cases. AUD or HAU diagnosis was only made if either a) CIDI was positive, *or* b) if simultaneously

two or more biochemical tests were elevated, (CDT and GGT, CDT and ALT, CDT and AST, CDT and MCV>97, GGT and MCV>100, MCV and ALT or MCV and AST), *or* c) simultaneous elevation of one biochemical test and presence of at least one clinical sign or clinical symptom were present: CDT and hepatomegaly, and CDT and cage≥2. When there was a possible non-alcoholic cause for positive biochemical and clinical signs, diagnosis was not made. An earlier version of the RDP is described in detail in a previous paper (4). In the earlier version of RDP we used SCID; in this study we used CIDI as it generates also ICD-10 AUD disorders.

3. Bayesian Alcoholism Test (BAT) is a probabilistic expert system, based on Bayesian statistics. BAT is designed as 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 p. 80). The à priori probability of the disease (estimated prevalence of the DUI population; left side of figure 1: population) is modified by combining information of the a priori probability with diagnostic data, such as clinical signs or biochemical markers, from a particular patient (right and bottom side of figure1) resulting in a posterior probability of the disease for a specific patient. BAT uses the probabilistic relationship between the disease and all known signs/ markers simultaneously in order to calculate the posterior probability for each subject to suffer from the disease. For further details on BAT, we refer the interested reader to our previous study on BAT (5). In the case of alcoholism, starting from the à priori prevalence, BAT combines all data, shown in figure 1, on each subject and yields a posterior probability for the patient to suffer from HAU. On the basis of earlier research in a different but fully comparable DUI population we estimated the a priori prevalence (base rate) of HAU in our DUI population to be approximately 50% (4). In that study we found that the best accuracy of BAT was obtained when the cut off was set at a BAT-score of 50 (5).

In order to correct for a negative influence of denying subjects we also calculated probabilities of HAU with a version of BAT without Cage, thereby avoiding the risk of false negative CAGE-input.

4. Clinical Diagnostic Procedure (CDP) is based on clinical reasoning. In this diagnostic procedure a diagnosis was reached by clini-

cal judgement after evaluation of all available data, according to usual clinical practice. Biochemical markers, CIDI, historical data, clinical signs, and instruments to assess alcohol problems and maximal amount of drinking in one day (TLFBMAX) were used. Histories included time and circumstances of arrest. A police report of the Blood alcohol Level (BAL), data of earlier DUI arrests and reports of earlier medical examinations after DUI were available. All data and clinical signs were assessed as either diminishing the chance of current AUD, increasing the chance of current AUD, or confirming AUD diagnosis. A positive diagnosis of current AUD was made if the above described CIDI and RDP procedures resulted in an AUD diagnosis or if several AUD probability -increasing -data were present without the presence of potential confounding effects of physical, non-alcohol related, illnesses or drugs.

Table 1 shows which items were used for the different diagnostic procedures.

CIDI results only in AUD diagnosis, BAT and BAT without CAGE result only in HAU diagnoses. The other procedures (CDP, RDP) result in either an AUD diagnosis, a HAU diagnosis or both. We hereafter refer to both HAU and AUD as "alcoholic".

Statistical analysis

Group characteristics of DUI's were compared using ANOVA. When differences between the four groups were significant, post hoc comparisons were made using Tukey honestly significant difference. Frequencies of diagnoses according to the different diagnostic systems were calculated. The differences were calculated using McNemars test and confidence intervals for the differences were calculated using Wilson's test (14).

Differences in diagnostic groups concerning drinking parameters and characteristics of DUI arrests were assessed using T test. Levene's test of equality of variances was performed and used for T-test. Agreement and differences between diagnostic procedures were described with kappa. The reasons for differences between the diagnostic procedures were described qualitatively. All statistics were performed with Statistics Package for Social Sciences (SPSS for Windows, 11.0, 2000) and the statistics software Confidence Interval

Analysis (CIA), version 2.05.

Table 1
Differences and similarities of the 5 diagnostic procedures: Items used in a standard fully structured interview (CIDI), a restrictive diagnostic procedure (RDP), a standard clinical diagnostic procedure (CDP) and a Bayesian alcoholism test (BAT) with and without Cage.

	CIDI	RDP	CDP	BAT	BAT without Cage
CIDI	+	+	+	-	-
CAGE	-	+	+	+	-
Medical History	-	+	+	-	-
Physical Examination	-	+	+	+	+
Biochemical data	-	+	+	+	+
Info about DUI arrest	-	-	+	-	-
Drinking Patterns (a.o. Tlfbmax)	-	-	+	-	-
Base rate Estimate	-	-	-	+	+
"algoritm"	+	+	-	+	+
AUD Diagnosis	+	+	+	-	-
HAU Diagnosis	-	+	+	+	+

BAL: Blood alcohol Level; AU: alcohol units; Tlfbmax: maximal used AU in one day in last three months. HAU: Hazardous Alcohol Use. AUD: Alcohol Use Disorder

RESULTS

Sample characteristics of the DUI subgroups

Table 2 shows, not surprisingly, that the second group ("many arrests group") had significantly more arrests in the last 5 years than the other groups ($p<0.000$). Also the BAL of the first group ("high BAL group") was significantly higher than the mean BAL in the other groups ($p<0.000$). There were no significant differences between the groups in terms of drinking parameters (reported AU/week, Tlfbmax and percentage of drinking days).

Table 2.
Group characteristics of 4 different groups of 116 drivers under influence. 1) DUI's with at least one arrest with a Blood Alcohol Level (BAL) ≥2.1‰ (n=29) 2) four DUI's arrests with any BAC above 0.5 ‰ within 5 years (n= 11), 3) refusal to cooperate with breath analysis (n=34). (Examination group n= 74). 4). DUI's that apply for re-granting the driving license after previous DUI, medical examination and loss of permanent driving license for 12 months because of diagnosis of alcoholism (re-examination group n=42).

	Study sample					
	High BAL group 1: (n=29)	Many arrests group 2: (n=11)	Refusal group 3: (n=34)	First examination groups 1-3: (n=74)	Re-examination group 4: (n=42)	p
Age (years)	42 (8,3)	35 (7,7)	40 (12,3)	40,5 (10,4)	42 (9,2)	n.s.
BAL at last arrest (‰)	2,33** (0,20)	1,31 (0,26)	1,32 (0,43)	1,84 (0,59)	1,82 (0,61)	0,000 †
No DUI arrests last 5 yrs	1,48 (0,74)	4,36* (1,69)	1,66 (0,81)	2,0 (1,4)	1,21 (1,29)	0,000 †
Reported AU/week	14,86 (19,81)	20,45 (16,62)	7,90 (8,86)	12,5 (15,7)	9,70 (15,70)	0,061
Tlfbmax	9,1 (7,45)	11,7 (7,0)	6,6 (7,87)	8,3 (6,5)	7,8 (7,0)	0,106
% of drinking days	33,1 (31,8)	52,2 (35,1)	27,96 (28,1)	33,5 (31,2)	33,86 (34,72)	0,283

Values are mean ± SD
† significant difference between the groups at p<0,000
* significant more arrests than other groups p<0,000
** significant higher BAL than other groups p<0,000
BAL: Blood alcohol Level; AU: alcohol units; Tlfbmax: maximal used AU in one day in last three months

Frequencies of diagnoses according to the different diagnostic systems

Table 3 shows that BAT (53%) and CDP (50%) identify significantly more subjects as alcoholics than CIDI (8%) and RDP (28%). There are no significant differences between BAT, BAT without Cage and CDP.

The prevalence of alcoholism (respectively HAU or AUD) according to the different diagnostic procedures was not significantly different between the DUI groups, except that according to RDP group 2 had a significantly lower prevalence of alcoholism than group 1, 3 and 4.

Table 3.
HAU and/or AUD diagnosis for different diagnostic procedures in 4 subgroups of 116 Drivers under influence.

Diagnostic procedure	CIDI	RDP	CDP	BAT without cage	BAT
Group 1 (n=29)	3 (10,3%)	13 (44,8%)	18 (62,1%)	18 (62,1%)	18 (62,1%)
Group 2 (n=11)	0 (0,0%)	1[1] (9,1%)	7 (63,6%)	6 (54,5%)	6 (54,5%)
Group 3 (n=34)	5 (14,7%)	8 (23,5%)	15 (44,1%)	17 (50%)	15 (44,1%)
Group 1-3 (n= 74)	8 (10,8%)	22 (29,7%)	40 (54,1%)	41 (55,4%)	39 (52,7%)
Group 4 (n=42)	1 (2,4%)	10 (23,8%)	18 (42,9%)	20 (47,6%)	22 (52,4%)
Total (n=116)	9 (7,8%)*	32 (27,6%)*	58 (50%)	61 (52,6%)	61 (52,6%)

*Differences between diagnostic systems for all subjects at p<0,05
CIDI<BAT (difference 0,44, 95%CI 0,34- 0,53), RDP<BAT (difference 0,26, 95%CI 0,16-0,35),
CIDI<BATwc (difference 0,44, 95%CI 0,33-0,53), RDP<BATwc (difference 0,25, 95%CI 0,15-0,34)
CIDI<CDP (difference 0,41, 95%CI:0,32-0,50), RDP<CDP (difference 0,23, 95%CI: 0,14-0,32)
CIDI<RDP (difference 0,2, 95%CI: 0,13-0,28)
No significant differences between BAT, BAT without Cage and CDP.
[1] Differences between subgroups at p<0,05. Only with RDP group 2 has lower prevalence than group 1, group 3 and group 4. No other significant differences between subgroups on any of the diagnostic systems.

Differences in diagnostic groups concerning drinking parameters and characteristics of DUI arrests indicative for HAU

As shown in table 4, there were significant differences of mean reported amount of drinking (AU/week), between positive or nega-

tive diagnosis, in all diagnostic procedures except CIDI. There were also significant differences of maximal amount of drinking between positive or negative diagnosis for BAT and BAT without Cage, but not for RDP and CIDI. This difference was not calculated for CDP, since maximal amount of drinking was incorporated in CDP. There were also significant differences of percentage of drinking days for all diagnostic procedures except CIDI.

Table 4.
Differences of drinking parameters between positive and negative diagnosis in CIDI, RDP, BAT, BAT without Cage and CDP

Diagnostic procedure	CIDI		RDP		CDP		BAT without Cage		BAT	
Diagnosis	+	-	+	-	+	-	+	-	+	-
Mean AU/week	23,0	10,5	17,7	9,1*	18,7	4,3*	18,1	6,3*	16,9	5,5*
TLFBMAX	11,0	7,6	9,8	7,1	10,0	5,7	9,4	6,1*	9,1	6,4*
Percentage of drinking days	54,2	32,3	47,2	28,4*	47,3	19,7*	44,5	21,6*	42,7	23,6*
BAL ‰	1,9	1,8	1,9	1,8	1,8	1,9	1,8	1,9	1,8	1,9
Mean no of arrests in last 5 years	1,9	1,8	1,6	1,8	2	1,5	1,8	1,7	1,9	1,6

Values are mean

*Differences between positive and negative diagnosis for all diagnostic procedures at $p<0,05$

Difference mean AU/week in CIDI is 12,5 (95% CI: -8,0 - 32,9), in RDP is 8,6 (95% CI: 7,3-16,5), in CDP is 14,4 (95% CI: 9,2-19,6), in BAT without Cage is 9,9 (95% CI: 4,5 - 15,3), and BAT is 11,4, (95% CI: 6,2 -16,6).

Difference mean TLFBMAX in CIDI is 3,4 (95% CI: -1,4 - 8,2), in RDP is 2,8 (95% CI: - 0,08 - 5,7), in BAT without Cage is 3,3 (95% CI: 0,8 -5,9) and in BAT is 2,73, (95% CI: 0,2 -5,3).

Difference in mean percentage of drinking days in CIDI is 21,9 (95% CI: -4,9 - 48,8), in RDP is 18,7 (95% CI: 3,1-34,4), in CDP is 27,6 (95% CI: 15,8- 39,4), in BAT without Cage is 22,9 (95% CI: 10,8 -35,0), and in BAT is 19,1 (95% CI: 6,7 - 31,4).

No significant differences between mean Blood alcohol levels and between mean number of arrests for all diagnostic procedures

There was no significant difference in Blood Alcohol level and in number of arrests in the last 5 years between positive or negative diagnosis for all diagnostic procedures.

Agreement and differences between diagnostic procedures
The agreement and differences between the diagnostic procedures are described in Table 5. The kappa between BAT, BAT without Cage and CDP are all above 0,74, indicating high agreement.

Table 5.
Agreement and differences between different diagnostic procedures. BAT: Bayesian Alcoholism test; BATwc: BAT without cage. CDP: Clinical Diagnostic Procedures. RDP: Restrictive diagnostic Procedure. CIDI: Composite International Diagnostic Interview.

Comparisons	++	--	+-	-+	Kappa and 95% CI
BAT- BATwc	55	49	6	6	0,793 (0,681-0,904)
BAT- CDP	53	50	8	5	0,776 (0,661- 0,891)
BAT- RDP	27	50	34	5	0,343 (0,075- 0,511)
BAT- CIDI	7	53	54	2	0,075 (-0,099-0,249)
BATwc-CDP	52	49	9	6	0,741 (0,619-0,864)
BATwc-RDP	27	50	34	5	0,343 (0,075- 0,511)
BATwc- CIDI	6	52	55	3	0,042 (-0,132-0,216)
CDP- CIDI	8	57	50	1	0,121 (-0,060-0,301)
CDP- RDP	27	31	53	5	0,379 (0,211-0,548)
RDP- CIDI	9	84	23	0	0,362 (0,128-0,595)

Qualitative differences between BAT, BAT without Cage and CDP
The qualitative differences between BAT and CDP are summarized in table 6. The subjects that scored negative on BAT and positive on CDP (5 subjects) had higher reported alcohol consumption (mean difference 11 AU/week; 95% CI 1,4-21,1). This result could be biased as there is also higher TLFBMAX (difference 5,9 AU/day; 95% CI 0,07-11,7), and TLFBMAX was incorporated in CDP. In the opposite case when BAT scored positive while CDP scored negative (8 subjects), the GGT was more often (4 subjects) elevated above cut off.

Table 6.
Specific reasons for discrepancies between BAT and CDP diagnoses

BAT HAU diagnosis, CDP no diagnosis		
Subject and DUI group	BAT score	Clinical data
1. Examination group	60	ALT 55 U/l, GGT 125 U/l, smoking 6 sig/d 1 DUI arrest
2. Examination group	67	ALT 64 U/l, GGT 71 U/l, smoking 20 sig/d, LRA 5 AU, Hepatitis present
3. Examination group	54	Cage=2, CDT 2,9, therapist writes patient stopped drinking 6 months ago
4. Re-examination group	59	Cage=2, CDT 2,7 U/l, smoking 10 sig/d, TIfbmax 10 au
5. Re-examination group	77	Cage=3, 4 DUI arrests last 5 years, Cage answer interpreted as unreliable
6. Re-examination group	98	CDT missing, AST 46 U/l, AST/ALT ratio 1,1, MCV 97 fL, BMI 36, tolerance 5 AU, cirrhosis.
7. Re-examination group	55	GGT 68 U/l, hepatomegaly
8. Re-examination group	65	ALT 61 U/l, GGT 80 U/l, BMI 31, smoking 4 sig/d, LRA 7 au.
CDP diagnosis, BAT HAU no diagnosis		
Subject and DUI group	BAT score	Clinical data
1. Examination group	45	CDT 2,9, MCV 99 fL, tlfbmax 7 au
2. Examination group	25	Cage=2, smoking 5 sig/d, tlfbmax 20 au, 3 DUI arrests in last 5 years
3. Examination group	25	CDT missing, 2 arrests in last half year
4. Examination group	49	CDT 2,8, AAG 1025
5. Examination group	31	CDT 2,6, tlfbmax 8 AU, 5 dui arrests in last 5 years

BAT cut off 50, CDT 2,6%, GGT cut off 65 U/l, ALT 50 U/l, AST 45 U/ L MCV 97 U/l, BMI 25, LRA 5 AU, Cage ≥2

The subjects that scored negative on BAT without Cage and positive on CDP (6 subjects) had higher Cage values (mean difference 1,6; 95% confidence interval 0,4-2,8), and higher amount of DUI arrests in the last 5 years (mean difference 2,0, 95% CI 0,7-3,3). In the opposite case when BAT without Cage scored positive while CDP scored negative (9 subjects), GGT had higher values (mean difference 33 U/l; 95% CI 5-61) and was more often elevated above cut-off values (6 subjects)

DISCUSSION

In this study we used several diagnostic procedures to identify AUD or HAU in a population of DUI's. We found that BAT and BAT without Cage and CDP identified the highest number of cases (52,6%, 52,5% and 50% respectively). RDP identified 27,8% and CIDI only 7,8%. There was a high agreement between BAT, BAT without Cage and CDP. For all diagnostic procedures except CIDI there was a significant difference in the average amount of drinking (AU/week) and in percentage of drinking days between subjects that scored positively and subjects that scored negatively. BAT, BAT without Cage and CDP, but not RDP and CIDI, were associated with the maximally used alcohol units in one day. Blood alcohol level and number of arrests in the last 5 years did not differentiate between positive or negative diagnosis in all procedures.

Furthermore, we found that there were almost no differences among the different subgroups within the population. The group with four or more prior arrests and the group with high BAL's did not differ on most diagnostic tests. Surprisingly, the examination group did not differ from the re-examination group. One would expect that subjects in the re-examination group would apply for re-granting their driver's license after they stopped or reduced their drinking and would score more often negative on the diagnostic procedures. It should be noted, however, that this group consists of only those subjects who did get an alcoholism diagnosis in their first examination.

Our study was an exploratory study, with the purpose of examining how the different diagnostic procedures performed and whether any assumptions can be made as to which procedure is better. As there is no gold standard, we cannot definitely indicate

which test was better. We will however consider the pros and con's of the diagnostic procedures as to which might be the most suitable in the specific context of forensic DUI's examination.

A first consideration is that the diagnostic procedures used in the present study are not fully independent. CIDI is incorporated in RDP, and both RDP and CIDI are incorporated in CDP. Not surprisingly, additional data resulted in more AUD diagnosis.

A second consideration is that the different diagnostic procedures identify different diagnostic categories. CIDI produces AUD diagnosis, while RDP, CDP and BAT produce a combination of AUD and HAU diagnosis. BAT without Cage gives only HAU diagnosis. Therefore, it is not surprising that BAT had better correlation with drinking patterns than CIDI.

It can be assumed that CIDI underestimates AUD diagnosis as it identifies only those DUI's who are aware of, and are willing to be open about, their alcohol problems. Also RDP might result in under-diagnosis because physical signs of alcoholism are late symptoms of alcoholism and some alcoholics may not show elevations on biochemical tests. Especially in young subjects biochemical markers have a low sensitivity for the detection of HAU (15-17).

An earlier study indicated that the prevalence in a DUI population should be in the range of 50 to 75% (4). BAT and BAT without Cage give prevalences that are consonant with this indication. Furthermore, BAT and BAT without Cage have the advantage of objectivity. Clinical intuitions, which have been proven to be often deceptive, do not interfere with the results of BAT. For example, in a subject with cirrhosis or hepatitis, laboratory abnormalities are attributed by the clinician to the possibility of non-alcoholic liver diseases and therefore the clinician is not sure enough to make an alcoholism diagnosis. BAT is constructed in such a way that the laboratory abnormalities might be attributed to non-alcoholic liver diseases, alcoholism or both. Maybe the clinician is too lenient here, as alcoholism is the most probable cause for the liver problems in this population. Another example is the fact that BAT scores HAU positive when GGT is increased in combination with hepatomegaly, and if the alternative reasons for such an increase are negative. Clinically this seems counterintuitive, as an elevated GGT seems a too weak confirmation of alcoholism and hepatomegaly is not a very

precise measurement. However, in a population with such a high prevalence of HAU, it is far more probable that the increased GGT and the hepatomegaly are due to excessive alcohol use than to any other unknown cause.

Another important issue when considering our results is the certainty of the diagnosis. As stated in the introduction: understanding the legal dilemma is essential in choosing between the different diagnostic procedures. The dilemma is to find a balance between two opposite aims. On the one side, the requirement to enhance traffic safety (for the public); each missed diagnosis endangers traffic safety. On the other side there is the interest of the individual driver. One must consider that the legal context of evaluation of medical diagnostic procedure falls under administrative law. Diagnostic procedures in this context are part of an administrative legal procedure to evaluate whether the subject has the right to have a driving license. While in criminal law the burden of proof must be given "beyond reasonable doubt", in administrative law less strict proof is required. The proof is set at a "plausible" or "most likely" level and proportional to the great "common good" that is at stake.

For traffic safety, the main question is the likelihood that a DUI will drive under the influence of alcohol again, or the recidivism risk. When driving under the influence of alcohol, the chance of "being caught" by the police is very low. This means that group 2, the group with at least 3 prior arrests, is a group that drives often under influence of alcohol and did not diminish that behavior after their first or second arrest. One could assume that this group is a group with a fairly high recidivism risk. However, this group has no higher positive scores on any of the diagnostic procedures. With RDP, this group even had a significantly lower prevalence than the group with one prior arrest with high BAC. With CIDI, no one of this group scored positively.

Based on the findings of the current study and our previous research with the BAT (5) we can conclude that in a population of denying subjects, the only diagnosis that can be made is HAU. The results of our previous study indicate that BAT has better diagnostic properties than the conventional biochemical markers for identifying HAU. Based on the prevalences found and the discrepancies

between BAT and CDP, the apparent advantage of objectivity with BAT seems to be confirmed.

References

1. Carrington Reid M, Fiellin DA, O'Connor PG. Hazardous and harmful alcohol consumption in primary care. Archives of Internal Medicine 1999;159:1681-1689
2. Mischke HD, Venneri RL. Reliability and validity of the MAST, Mortimer Fuilkins Questionnaire and CAGE in DWI Assessment. Journal of studies on Alcohol 1987; 48, 492-501.
3. Conigrave KM, Saunders JB, and Whitfield JB. Diagnostic tests for alcohol consumption. Alcohol and Alcoholism 1995; 30:13-26.
4 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 and Alcoholism 2001; 36: 594-602
5 Korzec A, de Bruijn C, van Lambalgen M. Confirming Diagnosis of Hazardous and Harmful Alcohol Use. Diagnostic accuracy of a computer assisted diagnostic system compared to conventional markers of alcoholism. (under review)
6 Feinstein A.R. Clinical Epidemiology. Philadelphia: WB Saunders. 1985.
7 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 and Alcohol Dependence 1997;47:195-205.
8 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 and Alcohol Dependence 1997;47:161-9.
9 Mayfield D, Mcleod G, Hall P. The CAGE questionnaire: validation of a new alcoholism screening instrument. American Journal of Psychiatry 1974;131:1121-1123.
10 Allen JP, Fertig JB, Litten RZ, Sillanaukee P, Anton RF. Proposed recommendations for research on biochemical markers for problematic drinking. Alcoholism: Clinical and Experimental Research 1997; 21:244-247.
11 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.
12 Korzec A, Arndt T, Bär M, Koeter MWJ: Trisialo-Fe_2-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.

13 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. Clinical Chemistry 2000;46:483-492.
14 Altman DG, Machin D, Bryant TN, Gardner MJ (eds). Statistics with confidence. BMJ Books sec edition 2001.
15 Nystrom M, Peraslo J, Salaspuro Carbohydrate-deficient transferrin (CDT) in serum as a possible indicator of heavy drinking in young students. Alcoholism: Clinical and Experimental Research 1992; 16, 93-97.
16 Chan AWK, Leong FW, Schanley DL, Welte JW, Wieczorek W, Rej R, Whitney RB. Transferrin and mitochondrial aspartate aminotransferase in young adult alcoholics. Drug and Alcohol Dependence 1989; 23:13-18.
17 Salaspuro, M. Carbohydrate-deficient transferrin as compared to other markers of alcoholism: a systematic review. Alcohol 1999; 19(3):261-271. Elsevier Science Inc.

Chapter 7

SUMMARY AND GENERAL DISCUSSION

Summary

The aim of this thesis was to enhance the validity of clinical diagnostics of alcoholism. More specifically the aim was to provide the clinician with a method to confirm the diagnosis of alcoholism. This method was eventually used in populations of drivers under influence (DUI's).

We first studied the discriminant validity of Alcohol Use Disorder (AUD) diagnoses according to DSM-IV within a population of well-functioning male heavy drinkers (chapter two). This study was conducted in order to explore whether it is possible to infer Alcohol Use Disorders from biochemical tests, clinical signs and clinical symptoms indicative of hazardous alcohol use. No significant differences were found between individuals with AUD and those without AUD.

In two different studies among DUI's, we used different methods to obtain a prevalence-estimate of alcoholism in a DUI-population. We found that the prevalence of alcoholism in the DUI population is around 50% (chapter three and six). We also found that interviews like SCID and CIDI are inadequate instruments to diagnose alcoholism in DUI's as they result in serious under-diagnosis.

In the study described in chapter four we compared the diagnostic accuracy of two tests for hazardous alcohol use (HAU): one %CDT-test including asialo-, monosialo-, disialo- and trisialo-isoforms and one without the trisialo-isoforms, and found that the CDT test without trisialo-isoforms had greater diagnostic accuracy.

The most important aim of this dissertation was the development of a confirmation test for diagnosing HAU. In chapter five, the development of a confirmation test, the Bayesian Alcoholism Test

(BAT), is described. Furthermore, BAT is validated and compared to single diagnostic tests in populations of treatment-seeking alcoholics, non-treatment-seeking heavy drinkers and non-alcoholic controls. We found that BAT has better diagnostic properties than CDT and GGT for confirming HAU.
Since our primary goal was to develop a confirmation diagnostic test in the context of medical examination of DUI's, we compared BAT to conventional methods used for diagnosing alcoholism in DUI's (chapter six). Because a gold standard for alcoholism does as yet not exist, we used alternative standards to validate BAT. The results of BAT and a Clinical Diagnostic Procedure (CDP) are most closely related to prevalences found in standard clinical practice. The advantage of BAT above CDP is that it is more objective because each subject is diagnosed in the same - objective - way.

General Discussion

As pointed out in the introduction, diagnosing alcoholism in the context of a legal situation raises several conceptual, epidemiological and clinical questions. We will first recapitulate these questions, then summarize the results of our studies, then discuss the implications of these results and end with suggestions for further research. The questions were:

1. How to define alcoholism?
2. What is the prevalence of alcoholism in a DUI population?
3. Which clinical arguments are used for the diagnosis of alcoholism and how valid are these arguments?
4. What is the value of the diagnostic tests used for the diagnosis of alcoholism in a DUI population?
5. Is it possible to design a diagnostic tool that, by combining probabilities of relationship between elevated biochemical markers and clinical signs, enhances the diagnostic ability to confirm whether a subject regularly uses a hazardous amount of alcohol?
6. Does such a diagnostic tool work in a real forensic situation where DUI's are examined for alcoholism?

Ad 1. How to define alcoholism?

Alcoholism can be approached as a mental disorder, resulting in Alcohol Use Disorder (AUD) diagnosis, or approached as alcohol intake that carries a health risk, resulting in a Hazardous Alcohol Use (HAU) diagnosis. In concurrence with the WHO/ISBRA collaborative project we chose an amount of 4 alcoholic units a day (for men) as threshold for HAU as this amount poses a risk for health (1,2).

For diagnosing DUI's, HAU diagnosis is more useful than AUD diagnosis as it can be diagnosed with objective biological tests that are not influenced by denying.

The question was whether one can infer AUD diagnosis from HAU diagnosis as is currently done in usual practice of forensic examination of DUI's.

In a population of well-functioning men with hazardous alcohol use we found no differences in drinking behavior and hazardous alcohol use indicators between subjects identified with a DSM IV AUD diagnosis (30 out of 57) and subjects without such a diagnosis. Even individuals with dependence can hardly be distinguished from those without AUD. Taking into consideration the methodological limitations of our study, we must question the possibility to infer current AUD-diagnoses by means of clinical signs and biochemical markers.

As it seems doubtful whether one can infer AUD from HAU and because AUD diagnoses are dependent on the cooperation of the subject, which is questionable in DUI's, the most logical choice is to decide for HAU diagnosis as alcoholism definition in the context of traffic safety.

Ad 2. What is the prevalence of alcoholism in a DUI population?

Bayes theorem is a mind-blowing experience for any clinician because it shatters many illusions of certainty in intuitive clinical diagnostics. The implication of Bayes theorem for diagnostics is described in the introduction.

The prevalence of alcoholism in a Dutch DUI population described in chapter three ranged from 8% to 82%, depending on the diagnostic criteria of alcoholism and the assessment procedure that was used. As there is no gold standard, the exact prevalence is in the dark.

The low prevalence rates obtained with SCID and CIDI (8%) are improbable because these instruments identify only those alcoholics who are aware of, and willing to be open about, their alcohol problems, which is unlikely in a DUI population, because this may lead to losing the driving license.

The high prevalence rates of alcoholism resulting from the population-based method (respectively 74% and 82%) are not very convincing either. There are several reasons for this conclusion. As shown in several studies, a DSM-IV abuse diagnosis, based solely on 'drinking in situations in which it is hazardous', does not signify that these subjects do always have a clearly distinguished alcohol problem. Both Vingilis (based on a literature search) and Hasin (based on research comparing a large sample of drinking drivers to controls and subjects with an abuse diagnosis without the criterion of driving after drinking) found that the percentage of 'hard core', or dependent alcoholics are a minority in the DUI population (5,6). This is in concurrence with our own clinical impression.

The prevalence found with the clinical diagnostic procedure (CDP) concurs best with earlier studies. For the total group, CDP identified 45% of DUI's as alcoholics; this is within the range of 25-50% (alcoholics ánd excessive drinkers) described by Vingilis in a review of prevalence studies among DUI's until 1989 and also within the range (25%-60%) of the studies, after 1989, described in chapter three table 1 (6-8). The prevalence found with CDP was different in subgroups of DUI's: in the examination group (the group that underwent their first examination after DUI) it was 58% and in the re-examination group (the group that was applying for re-granting a driving license after an earlier diagnosis of alcoholism) it was 36%. In our second study -chapter six- these percentages were 54% and 43% respectively. These percentages of the examination group also converge well with the national prevalence rates found by the Disqualification Division, as can be expected from the fact that they base their numbers on the psychiatric reports using the CDP. The prevalence found by CDP is dependent on the clinical arguments that are being used. The next step was therefore to evaluate the validity of these arguments.

Ad 3. Which clinical arguments are used for the diagnosis of alcoholism and how valid are these arguments?

In clinical reasoning, the clinician interprets medical history items, clinical signs and biochemical tests as either increasing or diminishing the probability on a positive diagnosis. In chapter three, the clinical diagnostic procedure for diagnosing alcoholism (AUD and HAU) is described. The good news about CDP is that it takes all available information (± 30 items consisting of historical data, clinical signs, biochemical measurements and instruments to assess alcohol problems) into account. The bad news is that the value and meaning of many items is not clear.

Because diagnosis in a legal situation can have serious consequences and sometimes has to be defended in court, there was a need for a diagnostic confirmation instrument. We therefore developed a clinical diagnostic system, the restrictive diagnostic procedure (RDP), with the goal of obtaining only definite cases of alcoholism (either AUD or HAU - see chapter three). For RDP we selected several 'robust' signs such as different combinations of elevated biochemical markers. In order to avoid false positive outcomes, we built 'safety valves' in the RDP-algorithm, controlling for other reasons for elevated bio markers (like drug use and the presence of non-alcoholic diseases). Items like high blood pressure, hand tremor, erythema palmare, smoking and level of response to alcohol (LRA) were not selected for RDP because of the relatively low specificity of these signs.

RDP identified 51 % of the alcoholics that were diagnosed with the usual clinical diagnostic procedure in our first DUI study, and 55% in our second DUI study when another AUD interview and a better CDT test was used. No comparison was possible with other confirmation tests, as there were none. Our clinical judgement was that RDP was a reasonable confirmation instrument. The fact that it identified only one half of the alcoholics that were identified with CDP was expected: higher specificity generally goes at the expense of lower sensitivity.

However, we also had to face some limitations of RDP: (a) RDP is mostly based on clinical experience and not on hard empirical data; (b) RDP has a too low sensitivity to be acceptable in the context of traffic safety where the danger for other traffic participants is at

stake; (c) RDP does not make use of the a priori prevalence of alcoholism in the population in which the diagnostic test is being performed, and d) RDP excludes signs which are not very specific but may have some information value for diagnosing alcoholism.

In order to enhance the validity of HAU diagnosis we had to leave behind the trust in clinical reasoning and solve two problems: 1. What reliable data are available on sensitivity and specificity values of markers of hazardous alcohol use, and 2. What is the best way to combine different diagnostic tests in order to maximize diagnostic information value? The results of the attempts to solve these problems are described in the next paragraphs.

Ad 4. What is the value of the diagnostic tests used for the diagnosis of alcoholism in a DUI population?

The value and accuracy of diagnostic tests is determined by several parameters: sensitivity and specificity of the diagnostic test, the prevalence of the disease in the population in which the test is used and the spectrum characteristics of the target population (9,10).

A variety of laboratory tests are available to assist in the diagnosis of hazardous alcohol use (11,12). However they have only moderate specificity and thus can not be used as confirmation instruments. Of all laboratory tests carbohydrate-deficient transferrin (CDT) has the best diagnostic accuracy, with an estimated specificity ranging from 80% to 96% depending on the selection of controls.

At the time our studies were performed there were many different CDT tests from which we had to choose. This was an important issue, CDT being the best available test for identifying hazardous alcohol use.

In chapter four we describe a comparison of two CDT tests. We found that the CDT test that includes trisialo-Fe_2-transferrin (%CDTri-TIA Axis, Norway) performed less well in terms of sensitivity than the CDT test that uses only the asialo-, monosialo- and diasialo- isoforms of carbohydrate-deficient transferrin (ChronAlcoI.D. Sangui Biotech Inc. USA). This result has also been also found in other studies (13-15). As a consequence the production of so-called "trisialo-tests" has been terminated, and replaced by a new CDT test, which excludes trisialo-Fe_2-transferrin (16).

Ad 5. Is it possible to design a diagnostic tool that, by combining probabilities of relationship between elevated biochemical markers and clinical signs, enhances the diagnostic ability to confirm whether a subject regularly uses a hazardous amount of alcohol?
In our main study, in chapter six, we aimed to build and validate an expert system that, by combining different diagnostic tests, enhances the diagnostic ability to confirm that an individual subject is suffering from HAU.

We developed an expert system, the Bayesian Alcoholism Test (BAT), to facilitate the confirmation of the diagnosis of HAU. As explained in the introduction, an expert system is a computer program that codifies existing general knowledge about a domain, 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 HAU.

The expert system allows answering queries of the following type: given values obtained for some diagnostic tests, what is the probability that a given patient suffers from HAU? What is the probability that the subject is suffering from another non-alcoholic disease? Also it is possible, in contrast to many other suggestions of combining diagnostic tests, to add a node that incorporates a prevalence estimate of the disease in the population where the test is applied. The limitations of other suggestions to combine laboratory tests or laboratory tests and clinical signs are described in the introduction.

a. Building BAT.
We first had to choose those history items, clinical signs and biochemical tests that are frequently related to alcoholism. Second, the items should be easy to measure and reliable (which made a sign such as erythema palmare problematic). Third, our selection was also motivated by those tests that could differentiate between alcoholism and conditions like liver disease or diabetes. The choice was initially made from 26 history items, 17 clinical signs and 5 biochemical tests that were able to distinguish social drinkers from alcoholics, described in the Alcohol Clinical Index (17). As this index is almost 20 years old we collected modern literature on items from the Alcohol Clinical Index and added other tests like CDT,

AST/ALT ratio and level of response to alcohol (LRA). The literature search also encompassed prevalence of diabetes and non-alcoholic liver diseases and causal probabilities between diabetes and non-alcoholic liver diseases and biochemical tests.

The variables 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 (chapter five 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 hazardous alcohol use as well for diabetes and for liver disease. The results of this procedure and literature selection is described in http://staff.science.uva.nl/~michiell.

b. Validation of BAT

We validated BAT in 3 populations: alcoholics, hazardous drinkers and non-alcoholic controls. We found that BAT had better diagnostic properties than CDT or GGT.

Comparing alcoholics with harmful use and controls (n=114), the sensitivity of the Bayesian Alcoholism Test was significantly higher, (94%), than the sensitivity of carbohydrate-deficient transferrin (63%), and of gamma-glutamyltransferase (73%). Specificity was high for the Bayesian Alcoholism Test (98%) but was not significantly different from the specificity of carbohydrate-deficient transferrin (93%) and specificity of gamma-glutamyltransferase (92%). Comparison of the ROC curves showed that BAT was superior to that of CDT and GGT. The area under the curve for BAT was 0,989 and was significantly higher ($p < 0.005$) than for CDT (0,909) and for GGT (0,902) (chapter five FIGURE 2).

In a population of heavy drinkers, the Bayesian Alcoholism Test could differentiate better than other markers between heavy drinkers above a harmful level (>56 AU/week), heavy drinkers with haz-

ardous use (28-56 AU/week) and heavy drinkers below a hazardous consumption (<28 AU/week).

Using pooled data of all 182 subjects, included in the study, the Bayesian Alcoholism Test had a better correlation coefficient (0.797) with the amount of drinking than carbohydrate-deficient transferrin (0.657), and gamma-glutamyltransferase (0.604).

The BAT system has several advantages above the usual diagnostic tests for excessive alcohol use.

First, our results indicate that, in the populations studied, this test has better diagnostic properties than the regular tests for populations of harmful users.

A second advantage is that it also produces a probability that the clinical and biochemical abnormalities are caused by another disease that sometimes gives clinical and chemical signs comparable to those caused by alcoholism.

The third advantage above other suggestions for using combinations of biochemical tests for confirmation of hazardous alcohol use, is that BAT can be easily accommodated for use in different populations with different spectra, and in different populations with varying prevalence of disease, without changing cut-off values of the used tests.

However, our study also has limitations that deserve attention. Firstly, our study results are applicable for men only. Secondly, the external validity of our study must be considered. We used a relatively small population, especially the controls. In addition, a relatively large percentage (27%) of the controls was abstaining from alcohol. It should be noted, however, that the sensitivity and specificity values of the usual hazardous alcohol use markers (CDT and GGT) in these populations were similar to those found in other studies (11,12). Thirdly, the majority of the conditional probabilities used in designing the diagnostic system are based on literature and expert estimates. Many studies had methodological shortcomings or produced inconclusive data. However, when new research and data on conditional probabilities in different populations become available, the properties of BAT can be further ameliorated.

The proof of the pudding is in the eating. We therefore applied BAT to a new population of DUI's.

Ad 6. Does BAT work in a real forensic situation where DUI's are examined for alcoholism?

The aim of this study was to compare four diagnostic procedures for confirming the diagnosis alcoholism: the standard fully structured interview (CIDI), our restrictive diagnostic procedure (RDP), the standard clinical diagnostic procedure (CDP) and our Bayesian Alcoholism Test (BAT) in a DUI population.

BAT identified 53 % of the total population as alcoholic, CDP 50%, RDP 28% and CIDI 8%. The agreement between BAT and CDP was high (P_0= 89%; Kappa 0,78, 95% CI: 0,66- 0,89). There was a significant difference in prevalence between BAT and RDP and between BAT and CIDI, but not between BAT and CDP. All diagnostic procedures were significantly correlated with the average amount of drinking (alcohol units/week). BAT was also significantly correlated with the highest number of alcohol units in one day. Comparing the subgroup with many previous arrests (>3 arrests), with the subgroup arrested with a high blood-alcohol-level (BAL), no significant differences in the results of the diagnostic procedures were found. This is a significant result as these items were thought to be of value in CDP. Many physicians estimate that a high BAL makes the diagnosis of alcoholism much more probable.

Different diagnostic procedures for diagnosing DUI result in widely ranging AUD and HAU prevalence rates. The results of BAT and CDP most closely resemble prevalence rates found in standard clinical practice. CIDI results in unlikely low prevalence rates. The advantage of BAT over CDP is that it is more objective, i.e. each subject is diagnosed in the same way. As a consequence, BAT diagnoses are probably easier to defend in court.

Conclusions and recommendations for future research

Confirming diagnosis of alcoholism in DUI's is possible but should be restricted to diagnosis of hazardous alcohol use. On the basis of our results, BAT is the most appropriate instrument to make this diagnosis in DUI's.

Examining the contributions of the single items to the diagnostic performance might further refine BAT. Regarding further validation of BAT, additional studies are necessary for women, phase III diagnostic studies (described at the end of the introduction) in popula-

tions with hepatic diseases without alcohol use and in general population samples to establish differentiating power in these populations. Furthermore it is interesting to investigate the diagnostic abilities of BAT for application in other populations, like life insurance examinations or in treated alcoholics. In other populations, one might find that BAT needs extra nodes.

Postscript: Alcoholism and traffic safety

DUI's as a group seem to have more overall traffic citations, more moving violations, more collisions and more suspended licenses than either an alcoholic or control group (6). This fact yields three questions on subgroups which can be associated with high risk of impaired driving:

1) Are alcoholics a hazard for traffic safety? A positive answer has face validity: alcoholics drink a lot each day and if they drive it seem obvious that they will engage more often in drunk driving. There is conclusive research that alcoholics, as a group, are involved in more collisions, are arrested more often for DUI and so forth. DUI's that meet the three criteria of high volume of alcohol intake, frequent drinking and alcohol dependence have a high rate of impaired driving incidents, an average of 5 per year (19). However, there are also studies suggesting that many alcoholics are not high-risk drivers because either they do not drive at all, or they do not while under influence of alcohol (6).

2) Is the question posed to clinicians by The Dutch Traffic Test Organization, whether there is a diagnosis of alcoholism the right one? The answer to that question is that it is not the right one, if alcoholism means a DSM IV Alcohol Use Disorder (AUD), because this diagnosis is based too much on information that can not be reliably obtained in a forensic setting. The question can, however, be answered if the clinician takes a HAU approach to the diagnosis of alcoholism and if he or she makes use of all the available objective information using the BAT as the final diagnostic procedure.

3) Which other subgroups than alcoholics have a high risk for relapse in DUI?

There is a vast amount of research on this question. Antisocial personality disorder, enhanced risk-taking after the use of alcohol and

binge drinking (20) are all associated with drunk driving but cannot be diagnosed reliably in the context of one medical examination.

From the perspective of DUI's all this is difficult to grasp. In the clinical practice of medical examinations of DUI's it is not unusual to see some DUI's who first lose the love of a partner and children because of heavy drinking to continue drinking, but decide to stop with alcohol after losing the driving license. Apparently some value mobility more than alcohol and marriage.

References

1. Saunders JB, Lee NK. Hazardous alcohol use: its delineation as a subthreshold disorder, and approaches to its diagnosis and management. Comprehensive Psychiatry 2000; 41:95-103.
2. 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. Alcoholism: Clinical and Experimental Research 2002; 26:332-339.
3. Hasin D, Paykin A: Alcohol dependence and abuse diagnoses: concurrent validity in a nationally representative sample. Alcoholism: Clinical and Experimental Research 1999;23:144-150.
4. Harrison PA, Fulkerson JA, Beebe TJ: DSM-IV substance use disorder criteria for adolescents: a critical examination based on a statewide school survey. American Journal of Psychiatry 1998;155:486-492.
5. Hasin D, Paykin A, Endicott J, Grant B: The validity of DSM-IV alcohol abuse: drunk drivers versus all others. Journal of Studies on Alcohol 1999; 60:745-755.
6. Vingilis, E. Drinking drivers and alcoholics: Are they from the same population?. In Research advances in alcohol and drug problems, Smart, R.G., et.al. eds, pp. 299-342. Plenum Press, New York 1983.
7. 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 and Alcoholism 2001; 36: 594-602.
8. Vingilis, E. Are drinking-drivers alcoholics? A review of the literature In High alcohol consumers and traffic: Proceedings of the International Workshop: High Alcohol Consumers and Traffic, Nov 28-30, pp.165-182. Inrets 1989.
9. Whiting P, Rutjes AWS, Reitsma JB, Glas AS, Bossuyt PMM, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy; a systematic review. Annals of Internal Medicine 2004;140:189-202.
10. Goehring C, Perrier A, Morabia A. Spectrum bias: a quantitative and graphical analysis of the variability of medical diagnostic test performance. Statistics in medicine 2004; 23:125-135.
11. Conigrave KM, Saunders JB, and Whitfield JB. Diagnostic tests for alcohol consumption. Alcohol and Alcoholism 1995; 30:13-26.
12. Salaspuro M. Carbohydrate-deficient transferrin as compared to other markers of alcoholism: a systematic review. Alcohol 1999; 19:261-271.
13. Arndt T. Carbohydrate-deficient transferrin as a marker of chronic alcohol abuse: a critical review of preanalysis, analysis, and interpretation. Clinical Chemistry 2001; 47:13-27.

14. Viitala K, Lähdesmäki K, Niemela O. Comparison of the Axis %CDT TIA and the CDTect method as laboratory tests of alcohol abuse. Clinical Chemistry 1998; 44:1209-15
15. Dibbelt L. Does trisialo-transferrin provide valuable information for the laboratory diagnosis of chronically increased alcohol consumption by determination of carbohydrate-deficient transferrin? Clinical Chemistry 2000; 46:1203-5.
16. Helander A, Fors M, Zakrisson B. Study of Axis-shield new %CDT immunoassay for quantification of carbohydrate-deficient transferrin (CDT) in serum. Alcohol and Alcoholism 2001; 36:406-412.
17. Skinner HA, Holt S, Sheu WJ, Israel Y. Clinical versus laboratory detection of alcohol abuse: the alcohol clinical index. British Medical Journal 1986; 292:1703-1708.
18. Korzec A, Hendrikx MMT, Warmink HH. Bestaat er een samnehang tussen alcoholisme en de kans op verkeersongevallen Nederlands Tijdschrift voor Geneeskunde 1995; 139:2020-2023.
19. Dawson DA Alternative definitions of high risk for impaired driving: the overlap of high volume, frequent heavy drinking and alcohol dependence. Drug and Alcohol Dependence 1999; 54:219-228.
20. Duncan DF. Chronic drinking, binge drinking and drunk driving. Psychological reports 1997; 80:681-682.

PUBLICATIES

Hengeveld MW, Korzec A. Een geval van rouwimpotentie. Nederlands Tijdschrift voor Geneeskunde 1980; 124(9):297-9.

Korzec, A. Weerzien met J. Haley. Tijdschrift voor psychotherapie. 18/4 1992.

Korzec, A. Bestaat Hysterie en is dat erg? in: Koerselman, G.F., Korzec A. Hysterie, een vergeten diagnose ? Uitgave van Bureau Post Academisch Onderwijs Geneeskunde Amsterdam 1992.

Koerselman, G.F.Korzec, A. (red.) Hysterie, een vergeten diagnose. Uitgave van Bureau Post-Academisch Onderwijs Geneeskunde Amsterdam 1992.

Koerselman, GF, Korzec A. Tussen geneeskunde en geneeskunst. In van Es J.C., Keeman J.N., De Leeuw P.W. e.a. (red.). Het medisch jaar 1995 Houten/Diegem; Bohn, Stafleu, van Loghum 1995.

Korzec A., Hendrikx MMT., Warmink HH. Bestaat er een samenhang tussen alcoholisme en de kans op verkeersongevallen? Nederlands Tijdschrift voor Geneeskunde 1995;139:2020-2023.

Korzec A. Bestaat er een samenhang tussen alcoholisme en de kans op verkeersongevallen? (beantwoording ingezonden brief). Nederlands Tijdschrift voor Geneeskunde 1995; 139: 2758.

Koerselman GF, Korzec A. Op den Velde, W. Tonino, JT. Spoedgevallen in de psychiatrie. In: Keeman, J.N., Schadé,E. (red.) Spoedeisende Geneeskunde. Houten/ Antwerpen. Bohn, Stafleu, van Loghum. 1996

Korzec A. Kan een biologische visie op psychiatrische stoornissen de psychiatrische diagnostiek ingrijpend veranderen? In: PAOG reader: Wat is medische psychiatrie? 1996.

Korzec A. A Protocol Strategy for an accurate Diagnosis of Alcoholism, in: Mercier-Guyon C. (eds): Alcohol, Drugs and Traffic Safety vol I. Proceedings of the 14th International Conference on Alcohol, Drugs and Traffic Safety. CERMT. Annecy 1997.

de Bruijn H, Korzec A, Bär M, Koeter MWJ. Riskant drinkgedrag met of zonder alcoholisme diagnose. Tijdschrift voor Alcohol Drugs en andere Psychotrope stoffen 2000.

Korzec A, Bar M, Koeter MAW, Kievit W. Diagnosing alcohol use disorders in drinking drivers. Proceedings 15th International conference on Alcohol drugs and traffic safety. Stockholm. 2000 ISBN 91-88250-43-1.

Korzec A, Bar M, Koeter MAW, Kievit W. Diagnosing alcoholism in high-risk drinking drivers: comparing different diagnostic procedures with estimated prevalence of hazardous alcohol use. Alcohol and Alcoholism 2001; 36:594-602.

Korzec A, Arndt T, Bär M, Koeter MAW. Trisialo-Fe_2-transferrin does not improve the diagnostic accuracy of carbohydrate-deficient transferrin (CDT) as a marker of chronic excessive alcohol intake. Journal of Laboratory Medicine 2001; 25 (9/10): 407-410.

Korzec A, Arndt T, Bär M, Koeter MAW. %CDTri-TIA vs ChronAlcoI.D. in the diagnosis of excessive alcohol intake in New Research Abstracts no 657 American Psychiatric Association 2001 Annual Meeting, Mind meets brain. New Orleans, LA, May 5-10, 2001. American Psychiatric Association, Washington.

Arndt T, Korzec A, Bär M, Koeter MWJ. Diagnostic accuracy of carbohydrate-deficient transferrin (CDT) as a marker of alcohol abuse is not improved by including trisialo-Fe2-transferrin in CDT - A study on alcoholics and hazardous drinkers. In: Pragst F, Aderjan R (Hrsg.). GTFCh Symposiumsband 2001, Toxikologische Aspekte der Sterbehilfe - Neue Drogen: chemische, analytische und toxikologische Aspekte. Verlag Dr. Dieter Helm, Heppenheim, 2001:327-342.

Arndt T, Korzec A, Bar M, Kropf J. Further arguments against including trisialo-Fe2-transferrin in carbohydrate-deficient transferrin (CDT): a study on male alcoholics and hazardous drinkers. Med Sci Monit. 2002 Jun; 8(6):CR411-8.

de Bruijn H, Korzec A, Arndt T, van den Brink W. The discriminant validity of alcohol use disorder in well-functioning men with hazardous alcohol use. European Addiction Research. 2003; 9(4):182-7.

Bakker A, Korzec A. De gecompliceerde kliniek: als protocollen tekortschieten 2003. Benecke N.I. Amsterdam 2002.

de Bruijn C, Korzec A, Van den Brink W: Behandeling van alcoholisme: maatwerk. In: Bakker A, Korzec A: De gecompliceerde kliniek 2003. Benecke N.I. Amsterdam 2002.

Bakker A, Korzec A. De gecompliceerde kliniek: als protocollen tekortschieten 2004. Benecke N.I. Amsterdam 2003

Vakpublicaties

Brinkgreve C, Korzec A, Over minnenijd. De Gids 1976; 4:219-230
Korzec A. Was Freud een romanticus? Intermediair 1983; 14: 23-25
Korzec A. , Korzec M. Een revolutie in gedachten: biologisch denken verovert de psychiatrie. Intermediair 41. 1993.
Korzec A. Gezondheid als deugd, het gezondheidsbegrip bij Nietzsche. Praecox 5. 1994.
Korzec A. Het 'dat nooit' gevoel. Tijdschrift voor het leerlingenwezen. 1996:15 (1)
Korzec A. Dodelijke kritiek. Boekman 2003; 57:86-91.

SAMENVATTING

Het doel van deze dissertatie was het verbeteren van de validiteit van de klinische diagnose van alcoholisme. Meer specifiek geformuleerd: het was onze bedoeling om een methode te ontwikkelen die de klinicus zou kunnen gebruiken om de diagnose alcoholisme te bevestigen. Deze methode werd uiteindelijk toegepast in de populatie rijders onder invloed (drivers under influence - DUI's).

In de eerste plaats bestudeerden wij de discrimante validiteit van stoornissen in gebruik van alcohol (alcohol use disorders - AUD), zoals geclassificeerd in de DSM-IV. Wij bestudeerden dit binnen een populatie van goed functionerende, mannelijke wijndrinkers met een stevig gebruik (hoofdstuk twee). De studie was opgezet met het doel te verkennen of het mogelijk is om stoornissen in gebruik van alcohol (AUD) af te leiden uit biochemische tests en klinische symptomen die wijzen op riskant gebruik van alcohol. Er werden geen significante verschillen gevonden tussen individuen met AUD en diegenen zonder AUD.

Om een prevalentie schatting te verkrijgen van alcoholisme in de populatie rijders onder invloed (DUI's) gebruikten wij verschillende methoden in twee studies. We vonden een prevalentie van alcoholisme in de populatie 'rijders onder invloed' van ongeveer 50% (hoofdstuk drie en zes). Interviews zoals de SCID en de CIDI bleken inadequate instrumenten om alcoholisme te diagnosticeren in de DUI-populatie daar zij resulteren in een ernstige onder-diagnose.

In hoofdstuk vier beschrijven we de studie waarin de diagnostische accuraatheid van twee testen voor riskant alcoholgebruik (hazardous alcohol use - HAU) zijn vergeleken een % CDT-test waarin asialo-, monosialo-, disialo- en trisialo-vormen en een zonder trisialo-vormen. In deze studie bleek dat de CDT test zónder trisialovormen een hogere diagnostische accuraatheid vertoont.

Het belangrijkste doel van deze dissertatie was de ontwikkeling van een test om riskant alcoholgebruik te bevestigen. In hoofdstuk vijf wordt de ontwikkeling van een confirmatie test, de Bayesiaanse Alcoholisme Test (BAT) beschreven. De BAT is vervolgens gevalideerd en vergeleken met enkelvoudige diagnostische tests in een populatie van alcoholisten die behandeling zochten, een populatie van drinkers met een stevig gebruik die geen behandeling zochten,

en een non-alcoholische controlegroep. BAT bleek betere diagnostische parameters te hebben dan CDT en GGT om de diagnose riskant alcoholgebruik te bevestigen.

Aangezien ons belangrijkste doel was om een diagnostische test te ontwikkelen, teneinde bij rijders onder invloed de diagnose alcoholisme te kunnen bevestigen, vergeleken we de BAT met conventionele methoden die alcoholisme diagnosticeren bij rijders onder invloed (hoofdstuk zes). Omdat een gouden standaard voor alcoholisme niet bestaat hebben we alternatieve standaarden gebruikt om de BAT te valideren. De resultaten van de BAT en de klinische diagnostische procedure (CDP) hangen nauw samen met prevalenties zoals we die vinden in de standaard klinische praktijk. De BAT heeft het voordeel dat zij een meer objectieve meting is omdat ieder subject op dezelfde -objectieve- wijze wordt gediagnosticeerd.

CURRICULUM VITAE

Aleksander Korzec studeerde geneeskunde aan de Universiteit van Amsterdam. Na zijn artsexamen in 1973 specialiseerde hij zich tot 1978 in de psychiatrie in het Wilhelmina Gasthuis te Amsterdam bij Prof dr PC Kuiper. Hij werkte vervolgens als psychiater en universitair docent in het AMC te Amsterdam en tevens als psychiater in het ziekenhuis De Heel in Zaandam. Sinds 1989 is hij als psychiater verbonden aan het Sint Lucas Andreas ziekenhuis te Amsterdam, sinds 2000 als A-opleider. Hij was onder meer lid van het bestuur van de medische staf. Voorts was hij lid van de Commissie Wetenschappelijke Activiteiten van de Nederlandse Vereniging voor Psychiatrie. Sinds 2000 is hij lid van het Consilium Psychiatricum.

The subject of this thesis is the use of Bayes' Theorem in research of a confirmation test for the diagnosis of alcoholism. Bayes' Theorem is named after Rev. Thomas Bayes, an 18th-century mathematician. The problem addressed and solved by Bayes regarded gambling and inverse probabilities. It is now widely used in diagnostic research.
The application of Bayes' Theorem is a mind-blowing experience for any clinician because it shatters many illusions of certainty in intuitive clinical diagnostics.