The Diagnostic Validity of Screening Tests and Laboratory Markers in Alcohol Use Disorders

Selami GÜL, Yıldız AKVARDAR, Gültakin TAŞ, Pınar TUNCEL

SUMMARY

Objective: To determine the validity of screening tests and markers suggested in the early detection of alcohol use disorders.

Method: Male patients aged 20-65 were recruited from a psychiatry clinic. They had been consuming alcohol at least once a week. They were not drug users and had no other diseases that affected the markers mentioned in this study. According to the DSM-IV diagnostic criteria, the diagnostic validity of screening tests [Michigan Alcoholism Screening Test (MAST), CAGE (Cut down, Annoyed, Guilty, Eye-opener), Alcohol Use Disorders Identification Test (AUDIT)] and markers [Mean corpuscular volume (MCV), aspartate aminotransferase (AST), alanine aminotransferase (ALT), gamma-glutamyltransferase (GGT), carbohydrate-deficient transferrin (CDT)] was investigated in alcohol dependent patients (n=29), alcohol abusers (n=28) and social drinkers (n=28).

Results: For the diagnosis of dependence, the sensitivity and specificity of the tests were respectively 69% and 100% for GGT, 41% and 93% for CDT, 100% and 57% for MAST, 97% and 86% for CAGE, and 100% and 75% for AUDIT. For the diagnosis of abuse, the sensitivity and specificity of the tests were respectively 32% and 100% for GGT, 32% and 93% for CDT, 86% and 57% for MAST, 75% and 85.7% for CAGE, and 89% and 75% for AUDIT. The widest area under the ROC curve belonged to GGT as a marker and AUDIT as a screening test.

Conclusion: According to the results of this study, the marker that has the highest validity is GGT, while AUDIT and CAGE are the best screening tests.

Key Words: Alcohol use disorders, screening tests, markers, validity

INTRODUCTION

Alcohol use disorders may lead to important problems in health, social and economic issues. Early diagnosis and effective intervention provide reasonable decrease in the quantity of these problems (Fleming et al, 1997). However, physicians diagnose less than half of patients and treat even less of them (Cherpitel et al, 1996, Dawson et al, 1992). One of the difficulties in early diagnosis and treatment is inadequate and unreliable self report of patients due to rejection and fear from being stigmatized. Some other obstacles may be negative attitude towards patients with alcohol use disorder, doubts about the success of treatment and inadequate knowledge of physicians.

Despite higher incidences of alcohol use disorders in Western societies (Farrell et al, 2001, Kessler et al, 1994), rates in Turkey are relatively low. Alcohol abuse prevalence was found as 0.8% (1.7% in males and 0.01% in females) in subjects over 18 in "Turkish Psychiatric Profile" project which was performed in 1997 (Kılıç, 1998). In Istanbul, rate for alcohol use was found as 33.5%, rate for risky alcohol use which was determined by CAGE (Cut down, Annoyed, Guilty, Eye-opener) screening test was found as 6.8% in 1997 (Akvardar et al, 2003a). It was reported that alcohol production and consume among especially younger individuals was growing in Turkey (Akvardar et al, 2003b, Türkcan, 1999).

Many practical and valid screening tests which depend upon self-report in order to
determine alcohol use disorders have been developed. Some of them are CAGE, Alcohol Use Disorders Identification Tests (AUDIT) and Michigan Alcoholism Screening Test (MAST) (Sharpe, 2001). Screening tests depend upon honesty and remembering abilities of patients. Rejection and fear from being stigmatized decrease the reliability of these tests. As a consequence a need for laboratory markers arise as objective tests (Sharpe 2001, Musshoff and Daldrup, 1998).

Carbohydrate-deficient transferrin (CDT) has been proposed as a promising marker recently (Salaspuro, 1999). It has been found to be effective from conventional markers like gamma-glutamyl transferase (GGT), aspartate-amino transferase (AST), alanine-amino transferase (ALT) and mean corpuscular volume (MCV) (Allen et al, 1994, Stibler, 1994). Sensitivity (82%) and specificity (97%) of CDT were reported to be high. But, results of later researches are controversial (Reynaud et al, 2000, Stowell et al 1997, Mikkelsen et al, 1998, Schmidt et al, 1997, Yersin et al, 1995).

In this study, it was aimed to evaluate the validity of screening tests like CAGE, MAST, AUDIT, a new marker CDT and conventional laboratory markers like MCV, ALT, AST and GGT in order to determine alcohol use disorders.

**METHOD**

**Sampling**

Male patients who were between 20-65 years old and using alcohol at least once a week and did not quit drinking more than last three days and were present at general and psychiatry in and outpatient units of Dokuz Eylül University Faculty of Medicine between July 2002 and March 2003 were recruited into the study. Those who had positive history for diseases (hepatobiliary system, pancreatic, endocrine, renal, cerebrovascular or ischemic heart diseases, cancer, trauma, epilepsy) or drugs (anti-psychotics, tricyclic anti-depressants, valproic acid, carbamazepine, barbiturates, anticoagulants, anti-inflammatory agents and antifungals) which might influence laboratory markers were excluded. Patients were classified according to DSM-IV diagnostic criteria (American Association of Psychiatry, 1994) as alcohol dependency (n=29), alcohol abuse (n=28) or social drinker (without alcohol use disorder, control group, n=28).

**Procedures and materials**

Structured interview form which included demographic characteristics, medical histories, alcohol consumption, duration and incidence characteristics, substance and drug use specificities and alcohol and substance part of structured clinical interview form for DSM-IV axis I disorders (SCID-I, Özkürkçügil et al, 1999) has been implemented to all patients who had given written informed consent. Ethics Committee of Dokuz Eylul University had given approval.

In order to determine alcohol use disorder, all patients have undergone CAGE, MATT and AUDIT evaluations. CAGE is a brief and short evaluation which includes four questions in order to differentiate individuals with alcohol use disorders (Ewing, 1984). Point of intersection is recommended as two. MATT includes 25 questions which differentiates alcohol dependants from others when point of intersection is accepted between 5-9. It evaluates drinking problem, help seeking behaviors and losses related with alcohol (Coşkunol et al, 1995). In this study, point of intersection is accepted as 5. AUDIT, has been developed as a screening test by World Health Organization in order to determine harmful and risky alcohol consumption in first stage health facilities. Point of intersection is assumed as 8 or 9 and the test consists 10 questions (Saunders et al, 1993). It was reported as valid in diagnosis of alcohol consumption according to the relationship with MATT (Saatçioğlu et al, 2002). In this study self reporting forms were used because of easy implementation and point of intersection was accepted as 8. These tests were evaluated after DSM-IV interview for diagnosis in order to prevent influence of test results.

Following completion of queries, blood samples of patients were handled. MCV, ALT, AST and GGT values were determined immediately and 5 cc of specimens were centrifuged and stored at −20 °C for CDT detection. Measurements were performed after completion of all groups.

Recommended reference values of the kits
were used. Mean corpuscular volume (MCV) measurement was performed as a component of routine complete blood count in Coulter LH 750 (Beckman-Coulter) device. Reference values for MCV in males between 21-65 years old were 80.7-95.5 fL. Serum AST and ALT levels were measured with standard method in Roche/Hitachi 747 auto-analyzer. Reference values were 0-38 U/L for AST and 0-41 U/L for ALT. Serum GGT levels were measured with colorimetric method in Roche/Hitachi 747 auto-analyzer. Values above 50 U/L were accepted as positive.

Measurement of CDT levels was performed with Tina-quant CDT % (Roche diagnostics kit which depended on calculation of percentage in total transferrin level. Following saturation of serum with iron, total transferrin and CDT columns were measured separately with Roche/Hitachi 912 analyzer. Then CDT percentages were calculated. Values above 6% were accepted as positive.

**Statistical analysis**

SPSS for Windows 11.0 and Epi Info 6 programs were used for statistical analysis. Demographics, alcohol consumption amount, frequency and duration and mean values of these quantities in three groups were compared with Kruskal-Wallis variance analysis and in binary groups with Mann-Whitney-U test. p levels below 0.05 were accepted as statistically significant. In order to determine the group which forms the significance, Bonferroni correction of p level (<0.016) was used. Validity evaluation was performed with alcohol and substance use disorders part of SCID-I interview based DSM-IV diagnosis. Sensitivity reflects the rate of patients with DSM-IV diagnosis determined by the test, and specificity reflects the rate of patients who have not been diagnosed as alcohol use disorder and determined by the test. Defining subjects as healthy or patient with discrete limits may lead to problems in interpretation of diagnostic tests (Aksakoğlu, 2001). For this reason ROC curves which provided linear comparison of sensitivity and specificity were used.

**FINDINGS**

**Social demographics and characteristics of alcohol consumption**

There were no statistically significant differences between three groups according to age or level of education (Table 1). Amount and frequency of alcohol consumption were
significantly higher in dependency and abuse groups versus controls. And duration of alcohol consumption was longer in dependency group versus others.

Mean values for MCV, AST, ALT and GGT levels were significantly different in all three groups (Table 1). Differences in MCV, AST and ALT levels were originated from dependency group, and in GGT levels were originated from dependency and abuse groups. There were no significant differences between mean CDT levels. Highest mean values in screening tests were present in dependency group, but the differences were not statistically significant.

**Diagnostic value of the tests**

In order to determine alcohol dependency, GGT had the highest predictive sensitivity (69%), and CDT had the second (41%) (Table 2). Specificity of all markers were high. Sensitivities of screening tests were higher and close to each other in alcohol dependency. Specificity of CAGE was highest (86%) and MATT was lowest (43%) in this group of patients. Combination of GGT+CAGE showed 100% sensitivity and 79% specificity in determining alcohol dependency. When GGT+CAGE+MCV combination was used sensitivity was 100% and specificity was 86%. In comparative analysis with ROC curve, area under curve (AUC) values for laboratory markers varied between 0.66-0.95 in dependency (Table 2, Figure 1). The largest area was made by GGT and the smallest one belonged to CDT. Areas of screening tests were large and similar (AUDIT=0.99, CAGE=0.98, MATT=0.96, Table 2, Figure 2).

Most sensitive laboratory markers were CDT and GGT for alcohol use disorders (32%, Table 2). Sensitivity of CAGE (75%) were found to be lower than other screening tests (MATT=86%, AUDIT=89%), but it showed the highest specificity (86%).

In comparative analysis with ROC curve, area under curve (AUC) values for laboratory markers varied between 0.47-0.84 in alcohol use disorders (Table 2, Figure 3). The largest area was made by GGT and the smallest one belonged to ALT (GGT>MCV>CDT>AST>ALT). Areas of screening tests were large and similar and the largest area was made by AUDIT (0.91, Table 2, Figure 4).

Combination of GGT+MCV showed 43% sensitivity and 100% specificity in alcohol use disorder (Table 2). When CAGE added to this combination, sensitivity was found as 89% and specificity was found as 86%.

**DISCUSSION**

**Diagnostic effectiveness of laboratory markers**

GGT was the most sensitive and specific marker of alcohol dependency and abuse. It constituted the largest area under ROC curve. In alcohol dependents and heavy drinkers, sensitivity of GGT was reported between 50-90% and mean specificity was found as 70% (Musshoff and Daldrup, 1998). In our study sensitivity of GGT was consistent with reported results and specificity was higher.

CDT, which was intensively studied in recent years as an objective marker of alcohol consumption, showed lower sensitivity than GGT
in determination of alcohol dependency and abuse in our study, but its specificity was higher (92.9%). In some studies with patients whose daily alcohol consumption were more than 60 grams, CDT was found to be more effective versus conventional markers (Anton et al. 2002, Huseby et al. 1997, Lesch et al. 1996, Anton et al. 1994). But there are some other studies which conclude the opposite (Mikkelsen et al. 1998, Schmitt et al. 1998, Yrsin et al. 1995). Analysis of CDT in laboratory is more complex versus other conventional markers (Musshoff and Daldrup 1998, Allen et al. 2001). The reason for relatively low sensitivity may be due to complex and difficult nature of the method or ignorance of the amount of alcohol consumption.

Reynaud and colleagues (2000) reported higher sensitivity rates for diagnostic efficacy of CDT, GGT and MCV in alcohol dependency, abuse or FIGURE 1. Effectiveness of laboratory markers in diagnosis of alcohol dependency according to analysis of ROC curve. Areas constituted by markers are 0.95 for GGT, 0.83 for AST, 0.76 for ALT and 0.66 for CDT.

Social drinker groups according to DSM-IV criteria (CDT 85%, GGT 80%, MCV 63% in dependency; CDT 67%, GGT 42%, MCV 24 % in abuse). Specificity of CDT and MCV was similar (97% and 96%), but GGT showed lower specificity rates than this study (76%). Alcohol consumption characteristics were not reported in this study, but there is a positive correlation between alcohol consumption amounts and laboratory markers (Gomez et al. 2001, Mikkelsen et al. 1998).

Mikkelsen and colleagues (1998) found CDT sensitivity as 47%, specificity as 90% and probability as 4.7 in their study which ignored the amount of alcohol consumption and performed in patients who were diagnosed as alcohol dependent according to SCAN (Schedules for Clinical Assessment in Neuropsychiatry). Mundle and colleagues (2000) found GGT sensitivity as 68% and specificity as 95%; CDT sensitivity as 46% and specificity as 97% to determine alcohol dependency according to DSM-III-R criteria. Consistency of our results with these studies thought to be due to use of structured diagnostic interviews, performance of diagnosis based evaluation, use of similar measurement method (CDT percentage) and lack of lower limit for alcohol consumption.

In our study AST and ALT showed low sensitivity in determining alcohol dependency and abuse and this finding was consistent with some other reports (Gomez et al. 2001, Sillanaukee et al. 1998, Bell et al. 1994). When specificity concept was taken into consideration controversial results were reported; Sillanaukee et al. (1998) and Bell et al. (1994) reported high
sensitivity as this study, but Gomez et al. (2001) found low results. ALT did not have diagnostic value in alcohol abuse and AST showed low diagnostic value.

In this study MCV showed low sensitivity (31% in dependency and 11% in abuse), high specificity (100%) and constituted a smaller area in ROC curve than GGT. Results of ROC analysis revealed similar (Schmitt et al. 1998) and lower (Bataille et al. 2003) levels when compared with GGT. Specificity of MCV is high, but its sensitivity is not satisfactory (Reynaud et al. 2000, Mundle et al. 2000, Schmitt et al. 1998, Yersin et al. 1995). For this reason, it is not useful in determining alcohol use disorder.

**Diagnostic effectiveness of screening tests**

In this study CAGE was found to be less sensitive, but more specific than other screening tests in alcohol abuse when cut-off value was accepted as ≥2. An ideal test should have sensitivity and specificity rates higher than 80% (Sharpe 2001). In our study if alcohol use disorders were evaluated as a single diagnostic group, CAGE reached these rates. When cut-off point was accepted as ≥2 for CAGE, sensitivity was reported between 61-100% and specificity was reported between 77-96% in some other studies (Cherpitel, 1997). Questioning guilty feelings due to alcohol consumption and thoughts about limit or give up alcohol use was commented as negative aspects of CAGE.

Alcohol use may cause guilty feelings following only one single use and healthy people may also think about limiting or giving up alcohol consumption (Sharpe, 2001).

**MATT** is a screening test which is recommended to determine alcohol use disorders and particularly evaluate lifelong medical, social and legislative problems due to alcohol use in dependency clinics. It was reported to demonstrate high sensitivity (87.7%) and moderate specificity (68.1%) in determining alcohol dependency and abuse (Teitelbaum and Mullen 2000). Our rates for sensitivity and specificity were similar. MATT constituted a relatively smaller area under ROC curve. MATT which has a greater extent than other screening tests, may be useful in earlier stages of evaluation.

**AUDIT** evaluates the amount of alcohol consumption besides determining dependency and problems related with alcohol use. This is important in order to demonstrate harmful alcohol consumption (Allen et al. 1997). It was reported to have sensitivity and specificity rates higher than 80% to determine harmful alcohol use(Saunders et al. 1993). Its sensitivity and specificity rates were reported to be high in people with severe psychiatric diseases followed on outpatient basis (Maisto et al. 2000). In our study, AUDIT had high sensitivity and reasonable specificity rates in determining dependency and abuse.
alcohol dependency at standard cut-off point. Its sensitivity was higher than laboratory markers and other screening tests, but specificity was lower than CAGE in determining alcohol abuse. AUDIT constituted the largest area at ROC analysis. It is well known that the amount of alcohol consumption is greater in males (Farrell et al. 2001). The reason for higher sensitivity rates of AUDIT which questions the frequency and amount of alcohol consumption may be due to inclusion of only male subjects.

In our study screening tests showed higher validity than laboratory markers in determining alcohol abuse. Similar results were reported in other studies. Siegfried and colleagues (2001) found lower specificity rates in a study which they performed in first stage health units, but the rate for CAGE in order to determine problematic drinkers was higher than CDT. Hermanson and colleagues (2000) found higher rates with AUDIT in transportation employees and Reynaud and colleagues (2001) found higher rates with CAGE in Emergency Room patients to demonstrate alcohol abuse than CDT or GGT.

Sensitivity of combinations may be higher than sensitivity of single tests. But in this instance specificity rate decreases (Sharpe 2001, Reynaud et al. 2001). In our study sensitivity of GGT-MCV combination was moderate-high and GGT-CDT combination was high in determining alcohol dependency. Combination of GGT-CDT showed sensitivity rates between 82-91% and specificity rates between 68-87% in determining male heavy drinkers and alcohol dependents (Anton et al. 2002, Mundle et al. 2000, Conigrave et al. 2002). Similarly, Mundle and colleagues (2000) reported that GGT-MCV combination showed 79% sensitivity and 91% specificity and GGT-CDT combination showed 82% sensitivity and 87% specificity in determining alcohol dependency.

When CAGE is combined with laboratory markers, determination rates for alcohol abuse increase (Reynaud et al. 2001). Combination of GGT-MCV and CAGE has the highest predictive capacity. Sensitivity and specificity of this combination were found to be over 80% in determining alcohol abuse. Combinations of AUDIT with laboratory markers revealed similar results achieved with AUDIT alone.

Laboratory markers which have weaker effectiveness than screening tests in determining alcohol abuse, may be used as objective diagnostic tools in medical services like psychiatry, emergency, internal medicine or surgery clinics where patients with alcohol use disorders are common. Moreover they may be used alone or in combination to discriminate alcohol status of risky drivers. In first stage health units where laboratory facilities are limited or in second or third stage health units (hospitals) CAGE or AUDIT may be used in order to determine alcohol use disorders. Moreover, questions of CAGE may be asked to include only the last year to provide more consistent results with diagnostic criteria. Combination of GGT+MCV+CAGE may be preferred because of high validity rates and cheap and easy implementation.

More severe symptoms like homelessness, unemployment, suicidal thoughts, incompliance with the therapy and frequent hospitalizations are more common in patients with additional psychiatric diagnoses (Breakey et al. 1998). For this reason screening tests and laboratory tests should be used more frequently in these patients. For screening purposes CAGE or AUDIT may be used routinely or combination of GGT-MCV-CAGE may be preferred in suspicious cases. Earlier diagnosis of alcohol use disorders with the help of screening tests and inexpensive laboratory markers may lead to decrease in quantity of problems caused by these disorders.

It is thought that the design of our study has two important characteristics: first, there is no limitation in the amount of alcohol consumption; second, because of DSM-IV criteria, there is no doubt in diagnoses. Alcohol use disorders may be present in mild or moderate drinkers as well (Farrell et al. 2001). As a consequence, markers may have less sensitivity in determining alcohol abuse, but are more objective because alcohol use disorders are not limited with physical problems. Assessment of alcohol abuse separately from dependency provides another positive aspect to our study. Alcohol abuse generally stays occult and becomes overt following onset of serious health problems (Miller 1991).
Inclusion of male patients only, low number of subjects and performance in a chosen clinic may be the limitations of our study and prevent the generalization of the results. There are a few studies which evaluate the validity of screening tests in determining alcohol dependency from our
country, but we did not find any study on alcohol abuse. There is a need for further researches which are focused on this purpose and performed in first stage health units and clinics where alcohol use disorders are seen more common.

REFERENCES


