Chemometric evaluation of nine alcohol biomarkers in a large population of clinically-classified subjects: pre-eminence of ethyl glucuronide concentration in hair for confirmatory classification

Analytical Bioanalytical Chemistry, 401, 7, 2011, DOI: 10.1007/s00216-011-5314-7
V. Pirro, V. Valente, P. Oliveri, A. De Bernardis, A. Salomone, M. Vincenti
volume 401, Springer, 2011, 2153-2164

Abstract

Abstract An important goal of forensic and clinical toxicology is to identify biological markers of ethanol consumption that allow an objective diagnosis of chronic alcohol misuse. Blood and head hair samples were collected from 175 subjects—objectively classified as non-drinkers (N=65), social drinkers (N=51) and active heavy drinkers (N=59)—and analyzed to determine eight traditional indirect biomarkers of ethanol consumption: aspartate aminotransferase (AST), alanine aminotransferase (ALT), gamma-glutamyltransferase (γ-GT), alkaline phosphatase (ALP), mean corpuscular volume (MCV), carbohydrate-deficient transferrin (CDT), and cholesterol and triglycerides in blood] and one direct biomarker [ethyl glucuronide (EtG) in head hair]. The experimental values obtained from these determinations were submitted to statistical evaluations. In particular, Kruskal–Wallis, Mann–Whitney and ROC curve analyses, together with principal component analysis (PCA), allowed the diagnostic performances of the various biomarkers to be evaluated and compared consistently. From these evaluations, it was possible to deduce that EtG measured in head hair is the only biomarker that can conclusively discriminate active heavy drinkers from social and non-drinkers, using a cut-off value of 30 pg/mg. In contrast, a few indirect biomarkers such as ALP, cholesterol, and triglycerides showed extremely low diagnostic abilities and may convey misleading information. AST and ALT proved to be highly correlated and exhibited quite low sensitivity and specificity. Consequently, either of these parameters can be discarded without compromising the classification efficiency. Among the indirect biomarkers, γ-GT provided the highest diagnostic accuracy, while CDT and MCV yielded high specificity but low sensitivity. It was therefore concluded that EtG in head hair is the only biomarker capable of supporting a confirmatory diagnosis of chronic alcohol abuse in both forensic and clinical practice, while it was found that γ-GT, CDT, MCV, and AST—whether used alone or in combination—do not allow the conclusive classification of subjects according to ethanol consumption. However, a diagnostic strategy combining these four parameters could be formulated in order to create a multivariate model capable of screening suspected active heavy drinkers.

Introduction

Alcoholism is the most frequent addiction encountered in many Western countries. Therefore, an important goal of forensic and clinical toxicology is to identify biological markers of ethanol consumption that are capable of objectively supporting the diagnosis of chronic excessive alcohol consumption [1, 2]. A wide variety of biochemical and hematological parameters are affected by alcohol consumption and used as potential biomarkers in order to determine (i) binge drinking, (ii) hazardous or harmful alcohol consumption, (iii) alcohol dependence, or (iv) abstinence [3]. These parameters are usually classified as direct biomarkers if their levels directly reflect the alcohol intake, or indirect biomarkers if their values are related to organ (most often the liver) damage resulting from repeated heavy ethanol exposure [4].

In Europe, several indirect biomarkers are collectively used to monitor excessive alcohol consumption—for example, when deciding whether to relicense a driver and during workplace impairment testing—although a unique recognized standard protocol for selecting the most appropriate biomarkers does not exist, and no cut-off level has been defined for them. On the other hand, indirect parameters do not appear to be applied systematically in North America [4].

Among indirect biomarkers, the onesmost commonly used are aspartate aminotransferase (AST) and alanine aminotransferase (ALT), gamma-glutamyltransferase (γ-GT), alkaline phosphatase (ALP), mean corpuscular volume (MCV), cholesterol (in particular high-density lipoprotein cholesterol, HDL-C), triglycerides, and carbohydrate-deficient transferrin (CDT) [3]. In particular, the latter was considered the most clinically significant marker of alcohol misuse until a few years ago [5]. All the markers mentioned so far are determined in either blood plasma or serum, and provide informative results over a relatively short retrospective period [6]. Also, many biological factors such as age, gender, body mass index, and ethnicity have been shown to cause wide inter- and intra-individual variability in their normal values [3]. These indirect biomarkers also present low specificity and sensitivity because they can be affected by several factors that are not correlated with ethanol consumption, such as the presence of metabolic disorders, nonalcoholic hepatitis, and biliary diseases. A main consequence of this lack of specificity and sensitivity is that these biomarkers do not have sufficient diagnostic accuracy to identify alcohol misuse. To improve their diagnostic potential, several studies have proposed that mathematical combinations of indirect biomarkers should be considered. For example, some authors have suggested various mathematical equations that combine γ-GT and CDT values in order to increase sensitivity without compromising specificity [7–12]. More complex statistical methods have also been proposed to improve both sensitivity and specificity using an Early Detection of Alcohol Consumption (EDAC) test, based on a combination of various routine laboratory tests [13, 14].

The determination of direct ethanol metabolites in various biological matrices is currently the most accredited strategy for proving chronic alcohol abuse. Unlike indirect biomarkers, ethyl glucuronide

(EtG), ethyl sulfate (EtS), and fatty acid ethyl esters (FAEEs) appear to be specific and sensitive direct biomarkers. Their concentrations in all biological matrices are likely to show a direct correlation with alcohol intake, with a minor chance of positive interference from spurious factors [15]. Furthermore, EtG and FAEEs can be easily detected in the keratin matrix too. Determining EtG and FAEEs in hair samples provides information on previous alcohol intake over a long retrospective period [16], a feature which has proved to be very useful in forensic practice [17–19]. Hair samples are easily collected by a noninvasive procedure and stored at room temperature.

Lately, EtG determination in hair has become the criterion most widely used to diagnose alcohol addiction. The mechanism of the incorporation of EtG into hair is not completely understood yet, but uptake from sweat appears to be involved [20], without any influence of the melanin content [15, 22]—in other words, natural hair color and degree of pigmentation do not induce any bias; nor does ethnicity [21]. In contrast, exposure to stressful environmental conditions, washing habits, and the use of special cosmetic products do appear to represent possible sources of bias [15, 23, 24].

Recently, an international consensus conference fixed the EtG cut-off level for identifying chronic alcohol abusers at 30 pg/mg [25], although some toxicologists have raised warnings about the general applicability of this cut-off level [2, 26]. For instance, it is suggested that pubic and axillary hair samples should not be employed as matrices instead of head hair [27, 28]. A relatively large inter-individual variability in EtG synthesis and metabolism is also expected, which possibly affects the conversion of ethanol into EtG, and consequently the concentration of it in all biological matrices [2]. Although high biological variability is described in the literature, this has not prevented wide acceptance of hair EtG data in a forensic context [17, 18].

In general, it has been demonstrated that since EtG is a biomarker that originates directly from ethanol, it is an excellent indicator of alcohol intake, but careful interpretation of the data is advised, and it is suggested that additional evidence, possibly including traditional alcohol biomarkers, should be considered too [17]. In the present study, head hair and blood samples were collected from 175 subjects that were clinically classified as non-drinkers (N=65), social drinkers (N=51), and active heavy drinkers (N=59). AST, ALT, γ-GT, CDT, MCV, cholesterol, triglycerides, and ALP were determined in blood samples, while EtG was measured in head hair samples. The experimental values obtained were mathematically transformed and normalized in order to perform statistical evaluations. Kruskal–Wallis, Mann–Whitney, and ROC curve analyses, together with principal component analysis (PCA), were utilized to evaluate the diagnostic power of each biomarker to identify chronic alcohol abuse, as well as to compare the biomarkers.

Our objective was to determine whether head hair EtG can be used as a unique biomarker to support the confirmatory diagnosis of chronic alcohol abuse in forensic and clinical practice. Our second goal was to determine which of the indirect biomarkers can be utilized to screen for alcohol misuse. Although the combined determination of both EtG and FAEEs is expected to produce the most accurate diagnosis of alcohol abuse [19], the high number of samples passed to our

laboratory (over 7000 hair samples per year), together with the availability of an ISO/IEC 17025:2005 accredited procedure for hair EtG, and the slightly higher reliability of EtG than FAEEs (i.e., fewer false-positive results) [18, 29], led us to consider EtG alone as a confirmation biomarker.

Experimental

Chemicals, reagents, and standard solutions Ethyl glucuronide and ethyl glucuronide-d5 (EtG-d5), used as the internal standard (IS), were acquired from Medichem® (Steinenbronn, Germany). Standard solutions of EtG and EtG-d5 were prepared in methanol at a concentration of 10 µg/mL and stored at −20 °C. Working solutions were prepared by progressive dilution. Calibration samples were obtained by spiking the blank head hair samples with the working solutions. Blank head hair samples were obtained from laboratory personnel who were non-drinkers (two subjects); these samples were stored at room temperature and used as surrogate matrix. Acetonitrile (CHROMASOLV®), dichloromethane (CHROMASOLV®), methanol (CHROMASOLV®), and formic acid were obtained from Sigma–Aldrich (St. Louis, MO, USA). Water was produced by a Milli-Q System (Millipore, Billerica, MA, USA). All of the chemicals and reagents used were of analytical purity grade. The determinations of AST, ALT, γ-GT, MCV, cholesterol, triglycerides, and ALP were performed using appropriate kits in accordance with the International Federation of Clinical Chemistry (IFCC) recommendations.

Determination of AST, ALT, γ-GT, MCV, cholesterol, triglycerides, and ALP in blood samples Serum activities of AST, ALT, γ-GT, cholesterol, triglycerides, and ALP were measured using a Roche Cobas Integra 800® autoanalyzer (Roche Diagnostics, Basel, Switzerland). MCV was measured by an Advia® 2120 hematology autoanalyzer (Siemens Healthcare Diagnostics, Milan, Italy). The upper reference limits used as cut-off values were 40, 50, 64, and 150 IU/L, respectively for AST, ALT, γ-GT, and ALP; 100 fL for MCV, and 200 mg/dL for cholesterol and triglycerides.

Determination of CDT in serum samples

%CDT was determined using the HPLC reagent kit from BioRad (Munich, Germany). Blood samples were prepared following the manufacturer's instructions. Briefly, the sample was centrifuged at 3500 rpm for 10 min. Next, 150 µL of serum were added to 450 µL of the solvent mixture; this was then agitated for few seconds and incubated at room temperature for 30 min. The resulting solution was centrifuged at 11,300 rpm for 10 min and the supernatant (500 µL) was transferred to a vial for HPLC–DAD analysis. The chromatographic separation was performed using an Agilent Technologies (Milan, Italy) series 1100 instrument interfaced to a G1315B diode-array detector (DAD). A wavelength of 460 nm was utilized for CDT detection. LC separation was

achieved by an anion-exchange cartridge maintained at +40 °C and eluted with a ternary buffer gradient delivered at a flow rate of 1.4 mL/min. For each
sample, 100 µL of supernatant were injected. The total runtime was 10 min. After integrating transferrin glycoform peaks using the ChemStation software (Agilent®), the relative percentage of low-carbohydrate transferrin (%CDT; consisting of asialo-, monosialo-, and disialotransferrin isoforms) was calculated in accordance with the Working Group on Standardization of CDT Measurement guidelines [30]. An upper reference limit of 2.0% was used as the cut-off value [31]. The present method was fully validated and accredited in accordance with ISO/IEC 17025:2005 rules. Linearity range, precision, carry over effect, and detection and quantification limits (LOD and LOQ) were evaluated.

Determination of EtG in head hair samples

Determinations of EtG in head hair samples were executed using a procedure described in the literature [32] under previously published analytical conditions [28]. Briefly, all hair samples were washed twice using methylene chloride and methanol (3 ml, 3 min) in sequence and then dried. Each sample was cut into small pieces (1–2 mm length) and weighed. About 50 mg of hair were added to 25 ng of EtG-d5 as well as 720 µl of a 35:1 v:v water and methanol mixture. Next, the samples were centrifugated (4000 rpm, 5 min) and incubated overnight at room temperature. Fifteen hours later, the samples were ultrasonicated with the UCI-150 ultrasonic cleaning bath (Raypa©, Ankara, Turkey) for 90 min, and then 100 µl of the liquid phase were transferred to a vial for LC-MS/MS analysis. Analyses were performed using an Agilent HPLC 1100 liquid chromatograph interfaced to an Applied Biosystems API 4000 triple-quadrupole mass spectrometer operating in ESI–negative ion mode. LC separation was performed using a LiChroCART Purospher STAR RP-18E column (150 mm × 4.6 mm i.d.) with a particle size of 5 µm. Two MS/MS transitions were utilized to identify and determine EtG and its internal standard. The EtG transition 221→75 was used for quantification.

The method was validated and accredited in accordance with ISO/IEC 17025:2005 rules. Linearity range, precision, accuracy, trueness, detection and quantification limits (LOD and LOQ), recovery, selectivity, specificity, and carryover effects were evaluated. Identification criteria for the analyte were established based on international [33, 34] and national [35] guidelines. Laboratory performance in hair EtG analysis was also monitored through its regular participation in interlaboratory proficiency tests organized by the Centre Universitaire Romand de Médecine Légale (University Hospital of Geneva, Switzerland) for the Society of Hair Testing. Linearity was observed in the range from 10 to 500 pg/mg with a correlation coefficient (R2) of 0.995. Intra-assay precision was satisfied, as the coefficient of variation (CV%) was below 10.0% at an EtG concentration of 10 pg/mg, and below 5.0% at 30 and 50 pg/mg. Both accuracy and trueness, expressed as the percent bias, were found to be lower than 15% at 10, 30, and 50 pg/mg. LOD

and LOQ values, calculated from analyses of multiple blank head hair samples, were 3 and 10 pg/mg, respectively. The detection of 3 pg/mg was experimentally confirmed by testing five spiked blank hair samples. An extraction recovery of 97.1±1.4% was estimated for samples spiked with EtG at a concentration of 30 pg/mg. Selectivity, specificity, and carryover tests all proved successful.

Study protocol

Blood and head hair samples were collected between October 2009 and December 2010 from two Alcohol Abuse Treatment Services located in Ciriè and Torino (Piedmont, Italy). One hundred seventy-five patients (30 women and 145 men aged 22–74, mean=44.6) who were participating in treatment or clinical trials were involved in the study. Using medical records, the patients were classified as active heavy drinkers (N=59, 10 women and 49 men aged 23–74, mean=46.9), social drinkers (N=51, 5 women and 46 men aged 22–74, mean=40.5) or non-drinkers (N=65, 11 women and 54 men aged 22–73, mean = 45.6). Not all of the participating social and non-drinkers presented previous alcohol dependence problems. The subjects were classified by medical doctors on a clinical basis, following OMS criteria DSM-IVand ICD-10, and using a combination of periodical blood tests, objective clinical symptoms, a questionnaire (essentially AUDIT), and personal interviews. In particular, most heavy drinkers had a full-blown medical history of alcohol abuse. General information on the estimated daily intake of ethanol during the past six months was gathered for each patient. Subjects who consumed more than 60 g ethanol per day were classified as active heavy drinkers [25]. Non-drinkers included people who had abstained from alcohol consumption for the last six months at least: some of them were former patients who were still under strict control; others were teetotallers who had volunteered to participate to the study. Social drinkers were mainly patients of the Addiction Services who were undergoing therapy for addictions other than alcohol (i.e., gambling, doping, and psychoactive drugs) and had volunteered to take part in the present study. In general, only patients who could be attributed to a certain group were included in the study; patients with an uncertain classification were discarded. Information on both recent and previous use of other legitimate and illicit drugs was also obtained. Medical history data on the occurrence of hepatic diseases (i.e., HCV, viral hepatitis, and hepatic cirrhosis) were also collected, which showed that 58 out of the 175 patients suffered from hepatic diseases. Each subject underwent venous whole blood sampling into two standard tubes for MCV and AST, ALT, γ–GT, CDT, cholesterol, triglycerides, and ALP determinations. The blood collected was stored at room temperature for less than 2 h and immediately processed to determine biochemical and hematological parameters. Hair samples were cut as close as possible to the scalp using freshly disinfected scissors. The samples were stored at room temperature and analyzed within 10 working days. Only hair segments that referred to a period of time in which the subjects maintained a regular alcohol daily intake were analyzed, and only the proximal 0–6 cm

segment was considered whenever a longer head hair sample was collected. Short head hair samples were analyzed in their entirety.

The study protocol was approved by the recognize Ethics Committee at San Luigi Gonzaga University Hospital [36, 37] (Turin, Italy). All patients provided written informed consent before attending the study, and an anonymous code was attributed to each participating subject in order to respect privacy regulations.

Statistical evaluations and ROC curve analyses

Inter-individual samples typically yield independent data; therefore, the Kruskal–Wallis nonparametric hypothesis test and the post hoc Mann–Whitney test were employed to verify the occurrence of statistically significant differences between drinker categories. In the Kruskal–Wallis test, the null hypothesis H0 affirms that there are no significant differences between at least two drinker groups. For the post hoc Mann–Whitney test, the null hypothesis H0 states that there are no significant differences between the two groups under examination. A significant level (a two-tailed P value) of 0.05 (CI=95%) was chosen for all statistical tests. When the experimental P value proved to be smaller than the critical P value, the hypothesis H0 was rejected. The effects of age and a medical history of hepatic diseases were evaluated for the same data set. The correlations between age and the levels of the biomarkers were studied in the whole data set and within the non-drinkers, social drinkers, and active heavy drinkers categories. For each biomarker and for all categories, Fisher weights (a measure of the ratio of the between-class variance to the within-class variance, representing the discrimination power) were also calculated in order to evaluate the separation between subjects who had suffered from previous hepatic diseases and those who had not. A Fisher weight of >1 indicates separation between categories [38]: the greater the value of the Fisher weight, the better the separation based on the considered variable. ROC curve analyses were performed using the whole data set in order to estimate the sensitivity and specificity at a certain cut-off value as well as the overall accuracy when discriminating active heavy drinkers from social and nondrinkers. ROC curves were constructed by plotting (1 − specificity) versus sensitivity. Both variables were modeled on the heavy drinkers category: the sensitivity was defined as the ratio of the number of true positive results to the total number of heavy drinkers, while the specificity was the ratio of the number of true negative results to the sum of the social drinker and nondrinker categories. Its complement (1 − specificity) is therefore the ratio of the number of false-positive results (i.e., non-drinkers classified as heavy drinkers) to the sum of the social drinker and nondrinker categories [38]. The area under the ROC curve (AUC) is an experimental measure of the test accuracy: an AUC value of 0.80 was selected as an acceptable limit of accuracy for screening biomarkers, while a value of 0.90 was chosen as an acceptable limit of accuracy for confirmatory biomarkers [39]. To generate the EtG ROC curve, a numerical value

between 0 and the LOD (3 pg/mg) was randomly generated and assigned to the samples with nonmeasurable (below the LOD) EtG concentrations. Correlation coefficients and Fisher weights were calculated using the free chemometric package V-PARVUS 2010 (Genoa, Italy) [40].
All statistical tests and ROC curve analyses were performed using the software package SPSS® (SPSS Inc., Chicago, IL, USA; version 17.0 for Windows).

Principal component analysis
The original data matrix had 175 rows (subjects) and nine columns (biomarkers). No values were missing from the data set. Numerical values between 0 and the LOD (3 pg/mg) were randomly generated and assigned to the samples with nonmeasurable EtG concentrations to ensure that the data variance was not underestimated. Since all of the variables showed skewed distributions, a logarithmic (base 10) transformation was applied to each of them. Skewed distributions are very usual for clinical and toxicological variables, and a logarithmic transformation is commonly used to obtain a normal distribution from highly skewed data [41]. Column autoscaling was also executed as a further pretreatment in order to remove systematic differences among variables resulting from dissimilar scales and measurement units. Principal component analysis (PCA) was performed on the autoscaled data. PCA is one of the most widely used explorative techniques, as it provides information on both the sample distribution and the correlations among variables [42]. PCA is able to concentrate a considerable fraction of the information contained in the original data into a reduced number of new uncorrelated variables (the principal components, PCs). The scores are the projections of the samples onto the new axes, and they provide information about the sample distribution. The loading values and the cosines of the angles between each variable and the PCs highlight which of the original variables are the most important in the definition of each PC. Fisher weights were calculated for each PC in order to estimate their discriminative abilities—their abilities to separate two categories. Also, two reduced matrices—with six and five columns (variables), respectively—were submitted to PCA. PCA was performed using the software package SPSS®. Fisher weights were calculated using the free chemometric package V–PARVUS 2010.

Results and Discussion

Statistical evaluation and ROC curve analysis The means and standard deviations calculated for each biomarker and category in the whole data set are reported in Table 1. To compare these data distributions, box plots for each biomarker and category are shown in Fig. 1, where the y-axis is represented on a logarithmic scale. The original data set is available in the “Electronic supplementary material” as Table S1. Upon applying the Kruskal–Wallis test, the null hypothesis H0 (no difference) was rejected for AST, ALT, γ-GT, CDT, MCV and EtG, confirming that significant differences exist in at least one of the drinker categories for these parameters. To

understand which of the three population sets was accountable for the difference, the post hoc Mann–Whitney test was applied. Using this test, it was possible to verify ($P<0.05$) that the AST, γ-GT, and EtG distributions show statistically significant differences between any pair drawn from the three categories (heavy drinkers, social drinkers, and non-drinkers). ALT could not discriminate heavy from social drinkers, while CDT and MCV showed significant differences between heavy drinkers and both social and non-drinkers but no differences between social and non-drinkers. Experimental P values from all of the statistical tests are reported in Table 2. A series of correlation coefficients were calculated (Table 3) in order to check whether the experimental values were affected by potential sources of bias such as the age of the subjects and/or a medical history of hepatic disease. No correlation was found between age and the direct or indirect alcohol biomarkers. Cholesterol provided the highest correlation with age for all classes of drinkers, with r coefficients of 0.234, 0.326, and 0.225, respectively. Likewise, none of the biomarkers investigated was significantly influenced by a history of previous hepatic disease, since all of their Fisher weights—calculated for each category between subjects who had and those who did not have a medical history of hepatic diseases—were found to be <1. It is worth noting that the highest Fisher weights were obtained for the non-drinkers category with AST (0.64), ALT (0.46), and γ-GT (0.26), respectively, while all other Fisher weights (Table 3) were below 0.2. This result is logical, as the potential previous hepatic damage recorded by AST, ALT and γ-GT is only discernible whenever a direct and relevant cause, such as heavy alcohol consumption, is not occurring. It is worth noting that head hair EtG, MCV, and CDT are the biomarkers with the lowest Fisher weights (all below 0.05), which demonstrates that they are not influenced by previous hepatic diseases, as expected. AUC, sensitivity, and (1 − specificity) values calculated for the whole data set are reported in Table 4. From these data, it is clear that EtG in head hair is the most sensitive and specific biomarker for identifying chronic alcohol misuse. At the recognized cut-off level of 30 pg/mg, sensitivity and specificity are 91.5% and 97.4%, respectively, far above the corresponding values for any other biomarker. This indicates that an active heavy drinker has a 91.5% chance of being identified as an alcohol abuser, whereas a non-heavy drinker has a 2.6% probability of being misidentified as an alcohol abuser. EtG also presents an AUC value of 0.982. This indicates that, within a binary classification scheme, a randomly selected positive case has a 98.2% probability of obtaining an higher score (i.e., EtG value) than a randomly selected negative case. Aside from EtG, none of the indirect biomarkers show sufficient diagnostic accuracy to justify their use as single indicators of chronic alcohol misuse for either screening or confirmatory purposes. Among the indirect biomarkers, γ-GT is the only one that yields an AUC value that is higher than the acceptable limit for screening purposes; it is the indirect biomarker aside from EtG in head hair that shows the highest sensitivity, although this is still unsatisfactory (66.1%), and its specificity is also very low (81.0%). On the contrary, MCV and CDT present acceptable specificity (96.6% and 90.5%, respectively) but extremely low sensitivity and AUC values that are below the acceptance limit of

0.80. AST appears to perform slightly better than ALT, whereas ALP exhibits good specificity (96.6%) but negligible sensitivity (13.6%). Figures 2 and 3 show the differences between %CDT and head hair EtG—the two biomarkers utilized in most European countries to assess chronic alcohol abuse. Figure 2 displays the ROC curves for EtG and CDT, which highlight how much EtG exceeds CDT in terms of sensitivity, specificity, and overall diagnostic power. In Fig. 3, head hair EtG is plotted against CDT on a logarithmic scale for all 175 subjects studied. The additional horizontal and vertical lines represent the cut-off values for EtG (30 pg/mg) and CDT (2.0%); these lines divide the plane into four quadrants that identify the number of true and false positive (or negative) results in each category. For example, the red dots (heavy drinkers) above each cut-off line identify true positive classifications, whereas the ones below the cut-off lines correspond to incorrect classifications. It is evident that EtG misidentifies only four heavy drinkers out of 59 (three of which are close to the cut-off level), while CDT produces as many as 30 false negative results with the same pool of subjects. Likewise, only three of the social drinkers (51 yellow dots) and non-drinkers (65 green dots) out of the 116 are classified as chronic abusers according to head hair EtG, while seven false positive identifications are yielded by CDT determinations. Nevertheless, it is worth noting that (i) 64 out of 65 non-drinkers are correctly classified by both biomarkers, and (ii) all 65 non-drinkers had EtG hair concentrations of 5 pg/mg or less; i.e., below the 7 pg/mg cut-off value that is internationally recognized as indicating abstinence [18, 43].

Principal component analysis

As a first approach, principal component analysis was applied to the whole data set (175 subjects × 9 variables). The first principal component (PC1) explained 64.8% of the total variance, while the second component (PC2) represented 17.6% of the total variance. The score and loading biplot for the two lower-order PCs is shown in Fig. 4. As the biplot shows, two groups of objects are present. These groups indicate similarities among samples on the basis of measured variables associated with alcohol intake. The differences between the subjects in terms of ethanol consumption are almost fully represented by the second principal component. In fact, PC2 is able to discriminate active heavy drinkers from social drinkers and non-drinkers; this is also confirmed by the value of the Fisher weight between active heavy drinkers and non-heavy drinkers, which is 6.1 for PC2. The loadings within the plot indicate that the biomarker most involved in the definition of PC2 (characterized by the highest absolute loading value along this axis), and therefore the biomarker that most strongly differentiates among drinker categories, is EtG in head hair, followed by CDT, γ-GT, and MCV. Since EtG derives directly from ethanol metabolism, we can deduce that PC2 variability is related to alcohol intake. Loading values confirm that CDT, MCV, and γ-GT are the most valuable of the indirect biomarkers, as they significantly contribute to group formation, thus providing useful information for classification purposes. Figure 4 conversely shows that PC1 is not useful for discriminating active heavy drinkers from social drinkers and non-drinkers. Accordingly,

the Fisher weight for PC1 is equal to 0.08. The distribution along PC1, which explains more than half of the variance contained in the original data, is most likely related to high inter-individual variability, which is commonly encountered when investigating such nonspecific clinical parameters. AST, ALT, and γ-GT are the biomarkers most involved in the definition of PC1; since these enzymes reflect hepatic damage, the occurrence of transitory or permanent hepatic stress may be responsible for the observed data dispersion along PC1. This assumption is clearly confirmed by considering the distribution of subjects in Fig. 5, especially within the nondrinkers category. The highest PC1 score values are all related to subjects who are suffering from hepatic disease or have a previous history of hepatic damage (A2: deep green dots vs. A1: light green dots). From Fig. 5, it is also clear that increases in the AST, ALT, and γ-GT levels due to ethanol consumption hide this trend for social drinkers and active heavy drinkers. In the PC2 direction, data dispersion is almost independent of the occurrence of hepatic disease. In particular, the maximum separation between subgroups 1 (light green dots) and 2 (deep green dots) is orthogonal to the EtG loading direction. The group separation between pairs of categories is even more evident from the inset of Fig. 5, where the centroid value for each category is depicted. The experimental loading values for ALP, cholesterol and triglycerides confirm that these three biomarkers do not make significant contributions to the definitions of PC1 and PC2. Indeed, after eliminating them from the data matrix, the separation of active heavy drinkers from social drinkers and non-drinkers improved slightly. This improvement was discernible in both the new PCA (see the “Electronic supplementary material,” Figure S1) applied to the reduced matrix with 175 rows and six columns (AST, ALT, γ-GT, CDT, MCV and EtG) and the PC2 Fisher weight, which increased from 6.1 to 6.4. A fourth PCA was applied to a different reduced matrix with 175 rows (subjects) and five columns (AST, ALT, γ- GT, CDT and MCV) from which EtG had been removed.

PC1 explained 73.6% and PC2 15.7% of the total variance. The score and loading biplot is displayed in Fig. 6. As already found from our examination of the whole data set, the discrimination between active heavy drinkers and non-heavy drinkers occurs almost totally along the second principal component, but its Fisher weight is only 1.16, a remarkable decrease from the value (6.4) obtained when EtG is included in the data set. It is also noticeable that the separation of the groups (see Fig. 6) is far less pronounced when the EtG parameter is not used than the separation observed in Fig. 4, confirming once again that head hair EtG is by far the most powerful biomarker for the identification of chronic alcohol misuse. Nevertheless, it is important to stress that PC2, which is a linear combination of the original variables, produces a Fisher weight (1.16) that is considerably higher than those obtained from the individual variables (AST: 0.33, ALT: 0.037, γ-GT: 0.88, CDT: 0.48, and MCV: 0.15). The loading values reported in Fig. 6 indicate that the biomarkers that contribute the most to the definition of PC2 are γ-GT, CDT, and MCV, while the parameters most involved in PC1 formation are AST, ALT, and γ-GT. In particular, AST and ALT

appear to highly correlated, meaning that they provide approximately the same information about the analyzed data set. Consequently, only one of these parameters can be usefully considered for classification purposes, possibly AST (see ROC, sensitivity and specificity analysis above).

Conclusions

Determining average alcohol consumption and recognizing chronic alcohol misuse have become increasingly important in many forensic and clinical activities, including driver licensing, workplace testing, and abuse rehabilitation. The objectives of the present study were to investigate which among several alcohol abuse biomarkers were the most efficient at singling out chronic heavy drinkers and identifying the best context (i.e., screening versus confirmation) to use them in. The concentration of EtG in head hair was found to be the most accurate biomarker of chronic alcohol misuse, so it is recommended for confirming initial judgments. In fact, EtG proved to be very sensitive and specific for the identification of active heavy drinkers, and guaranteed excellent diagnostic accuracy—with negligible false positive and false negative results—provided that an appropriate cutoff value is utilized. Further studies addressing the possible sources of false results and determining the optimal cut-off value for non-drinkers are underway in our as well other laboratories. In contrast, no other indirect biomarker appeared to have sufficient sensitivity to effectively identify chronic abusers. For example, the determination of CDT in blood samples leaves half of the chronic abusers undetected. Combining indirect biomarkers rather than using them individually significantly improves the ability to diagnose alcohol misuse, but the use of such a combination should still be confined to screening applications—namely, circumstances where the chance of overlooking some chronic abusers is acceptable. On the other hand, the several false positive results obtained when using indirect biomarkers indicates that the adoption of confirmatory procedures is necessary. Among the indirect biomarkers, γ-GT, CDT, MCV, and AST were found to provide useful information on alcohol consumption, whereas cholesterol, triglycerides, and ALP do not contribute relevant information, and may occasionally act as misleading factors. ALT proved to be closely correlated with AST, and it provides a similar (but weaker) contribution. In conclusion, a diagnostic strategy combining γ-GT, CDT, MCV, and AST can be proposed for screening evaluation, based on a multivariate model capable of highlighting suspected active heavy drinkers. Potential sources of bias, such as subject age or a medical history of previous hepatic disease, which are frequently claimed to be sources of bias in order to justify positive reports, were found to have only very minor effects on useful alcohol biomarker values. In particular, EtG, CDT, and MCV appeared to be totally unaffected by the occurrence of previous hepatic disease.

Acknowledgements

The authors wish to thank Bruno Sciutteri and Emanuele Bignamini of the Alcohol Abuse Treatment Service Department of ASO–TO2 (Turin, Italy) and Silvia Aloi of the Alcohol Abuse Treatment Service Department of ASO–TO4 (Ciriè, Turin, Italy) for their keen cooperation. We also wish to thank all the laboratory personnel for preparing all the samples. The generous financial contribution from the Compagnia di San Paolo (Turin, Italy) for renovating analytical instrumentation is gratefully acknowledged (grant 411/PV–2009.1993). The authors are also indebted to the Regione Piemonte for its continuous financial support.

References

1. Kerekes I, Yegles M, Grimm U, Wennig R (2009) Alcohol Alcohol 44:62–66
2. Tagliaro F, Bortolotti F, Viel G, Ferrara SD (2011) Forensic Sci Int 207:e69–70
3. Conigrave KM, Davies P, Haber P, Whitfield JB (2003) Addiction 98:31–43
4. Marques P, Tippetts S, Allen J, Javors M, Alling C, Yegles M, Pragst F, Wurst F (2009) Addiction 105:226–239
5. Høiseth G, Morini L, Polettini A, Christophersen A, Mørland J (2009) Alcohol Clin Exp Res 33:812–816
6. Liniger B, Nguyen A, Friedrich–Koch A, Yegles M (2010) Traffic Inj Prev 11:123–126
7. Hietala J, Koivisto H, Anttila P, Niemelä O (2006) Alcohol Alcohol 41:528–533
8. Bianchi V, Ivaldi A, Raspagni A, Arfini C, Vidali M (2010) Alcohol Alcohol 45:247–251
9. Antilla P, Järvi K, Latvala J, Blake JE, Niemelä O (2003) Clin Chim Acta 338:45–51
10. Anton RF, Lieber C, Tabakoff B (2002) Alcohol Clin Exp Res 26:1215–1222
11. Thiesen H, Hesse M (2010) Addict Behav 35:260–262
12. Chen J, Conigrave KM, Macaskill P, Whitfield JB, Irwig L (2003) Alcohol Alcohol 38:574–582
13. Harasymiw J, Seaberg J, Bean P (2004)Alcohol Alcohol 39:329–335
14. Harasymiw J, Seaberg J, Bean P (2006) J Addict Dis 25:59–63
15. Morini L, Politi L, Polettini A (2009) Addiction 104:915–920
16. Pragst F, Balikova MA (2006) Clin Chim Acta 370:17–49
17. Kintz P (2011) Forensic Sci Int 207:e71
18. Albermann ME, Musshoff F, Madea B (2011) Anal Bioanal Chem 400:175–181
19. Prags F, Rothe M, Moench B, Hastedt M, Herre S, Simmert D (2010) Forensic Sci Int 196:101–110
20. Albermann ME, Musshoff F, Madea B (2010) Anal Bioanal Chem 396:2441–2447
21. Paul R, Kingston R, Tsanaclis L, Berry A, Guwy A (2008) Forensic Sci Int 176:82–86

22. Kharbouche H, Steiner N, Morelato M, Staub C, Boutrel B, Mangin P, Sporkert F, Augsburger M (2010) Alcohol 44:507–514
23. Offidani C, Strano Rossi S, Chiarotti M (1993) Forensic Sci Int 63:105–108
24. Morini L, Zucchella A, Polettini A, Politi L, Groppi A (2010) Forensic Sci Int 198:23–27
25. Kintz P (2010) Forensic Sci Int 196:2
26. Palmer RB (2009) Semin Diagn Pathol 26:18–27
27. Lee S, Han E, In S, Choi H, Chung H, Chung KH (2011) Forensic Sci Int 206:19–21
28. Pirro V, Di Corcia D, Pellegrino S, Sciutteri B, Vincenti M, Salomone A (2011) Forensic Sci Int 210:271–277
29. Gareri J, Appenzeller B, Walasek P, Koren G (2011) Anal Bioanal Chem 400:183–188
30. Jeppsson JO, Arndt T, Schellenberg F, Wieldens JMP, Anton RF, Whitfield JB, Helander A (2007) Clin Chem Lab Med 45:558–562
31. Bianchi V, Arfini C, Helander A (2007) Clin Chem Lab Med 45: A161
32. Morini L, Politi L, Groppi A, Stramesi C, Polettini A (2006) J Mass Spectrom 41:34–42
33. EC (2002) Commission Decision 2002/657/EC implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results. Off J Eur Commun L221/8–36
34. SOFT/AAFS (2006) Forensic toxicology laboratory guidelines. Society of Forensic Toxicologists/American Academy of Forensic Sciences, Mesa/Colorado Springs
35. GTFI (2010) Linee guida per i laboratori di analisi di sostanze d'abuso con finalità tossicologico–forensi e medico–legali. Gruppo Tossicologi Forensi Italiani, Pavia
36. Dilts DM (2010) J Clin Oncol 28:3799–3800
37. Wang-Gillam A, Williams K, Novello S, Gao F, Scagliotti GV, Govindan R (2010) J Clin Oncol 28:3803–3807
38. Forina M, Oliveri P, Lanteri S, Casale M (2008) Chemometrics Intell Lab Syst 93:132–148
39. Luna Herrera J, Martínez-Cabrera G, Parra-Maldonado R, Enciso-Moreno JA, Torrez-López J, Quesada-Pascual F, Delgadillo-Polanco R, Franzblau SG (2003) Eur J Clin Microbiol Infect Dis 22:21–27
40. Forina M, Lanteri F, Armanino C, Casolino C, Casale M, Oliveri P (2010) V–PARVUS 2010. University of Genoa, Genoa. http:// www.parvus.unige.it. Accessed Jan. 27th, 2011
41. Leardi R, Armanino C, Lanteri S, Alberotanza L (2000) J Chemometrics 14:187–195
42. Vandeginste BGM, Massart DL, Buydens LMC, De Jong S, Lewi PJ, Smeyers–Verbeke J (2008) Supervised pattern recognition. In: Vandeginste BGM, Rutan SC (eds) Handobook of chemometrics and qualimetrics: part B. Elsevier, Amsterdam, pp 88–158, 207–213, 236–237
43. Pragst F, Yegles M (2008) Ther Drug Monit 30:255–263

Table 1 The means and standard deviations calculated for each biomarker and category in the whole data–set

Biomarker	Overall	Non–drinkers	Social drinkers	Heavy drinkers
AST (IU/L)	43±47	25±16	40±33	65±66
ALT (IU/L)	47±48	30±25	59±63	54±47
γ–GT (IU/L)	121±264	40±61	51±68	271±408
CDT (%)	1.9±1.6	1.3±0.3	1.6±0.8	2.7±2.4
MCV (fL)	90.1±11.7	88.7±14.8	87.8±6.3	93.8±10.6
EtG (pg/mg)	67.9±133.3	3.5±1.2	8.0±9.2	190.8±173.3
Cholesterol (mg/dL)	188±104	190±45	164±173	208±62
Triglycerides (mg/dL)	157±251	138±99	170±208	166±374
ALP (IU/L)	87±48	80±31	82±27	97±70

Table 2 Experimental P–values for Kruskal–Wallis and *post–hoc* Mann–Whitney tests performed for each biomarkers in the whole data–set

Biomarker	Kruskal–Wallis	Mann–Whitney		
		social vs non–drinkers	heavy vs non–drinkers	heavy vs social drinkers
AST (IU/L)	< 0.0005	< 0.0005	< 0.0005	0.014
ALT (IU/L)	< 0.0005	0.002	< 0.0005	0.813
γ–GT (IU/L)	< 0.0005	0.022	< 0.0005	< 0.0005
CDT (%)	< 0.0005	0.063	< 0.0005	< 0.0005
MCV (fL)	< 0.0005	0.169	< 0.0005	< 0.0005
EtG (pg/mg)	< 0.0005	0.001	< 0.0005	< 0.0005
Cholesterol (mg/dL)	0.182			
Triglycerides (mg/dL)	0.535			
ALP (IU/L)	0.661			

Table 3 AUC, sensitivity and (1–specificity) values *

Biomarker	AUC value	sensitivity	(1–specificity)
AST (IU/L)	0.728	0.475	0.181
ALT (IU/L)	0.612	0.356	0.198
γ–GT (IU/L)	0.834	0.661	0.190
CDT (%)	0.771	0.508	0.095
MCV (fL)	0.730	0.169	0.034
EtG (pg/mg)	0.982	0.915	0.026
Cholesterol (mg/dL)	0.579	0.525	0.422
Triglycerides (mg/dL)	0.450	0.119	0.198
ALP (IU/L)	0.528	0.136	0.034

* Sensitivity and (1–specificity) values were calculated at the following cut-off level: AST: 40 IU/L, ALT: 50 IU/L, γ–GT: 64 IU/L, CDT: 2.0%, MCV: 100.0 fL, EtG: 30 pg/mg, cholesterol and triglycerides: 200 mg/dL, ALP: 150 IU/L

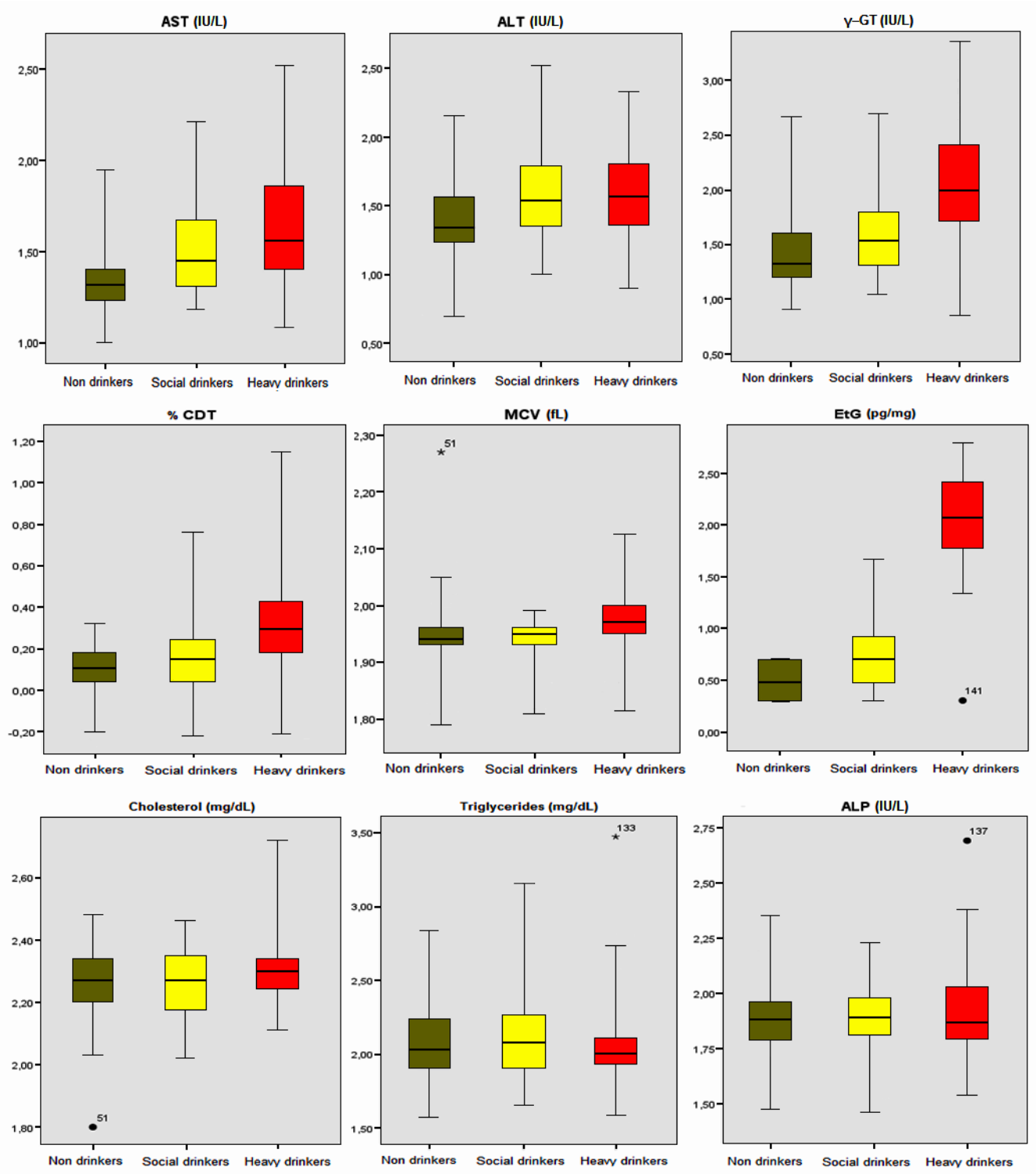


Fig. 1 Box plots for each biomarker and category. The data are reported on logarithmic scales (base 10)

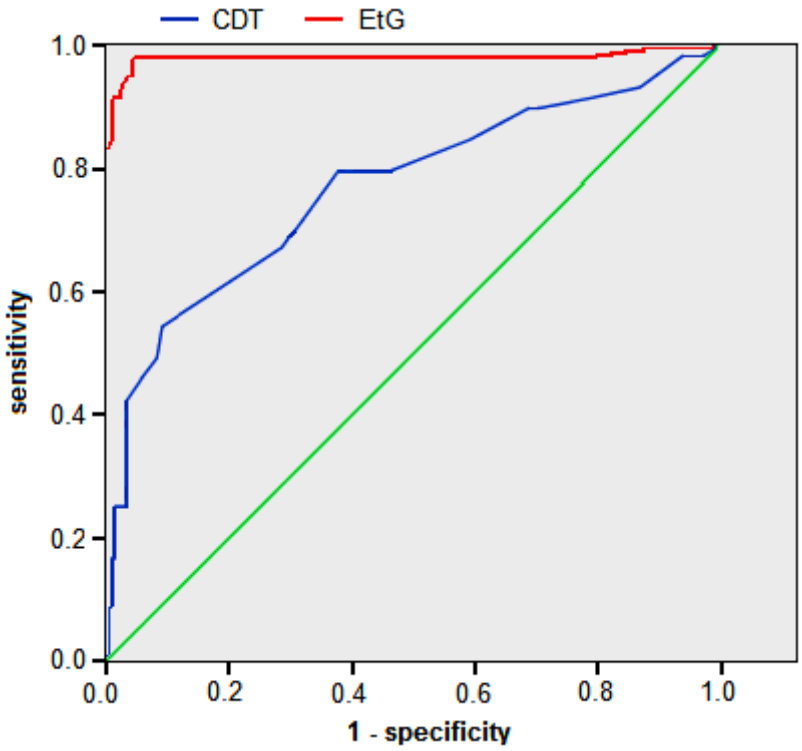


Fig. 2 Receiver operating characteristic (ROC) curves for head hair EtG and CDT

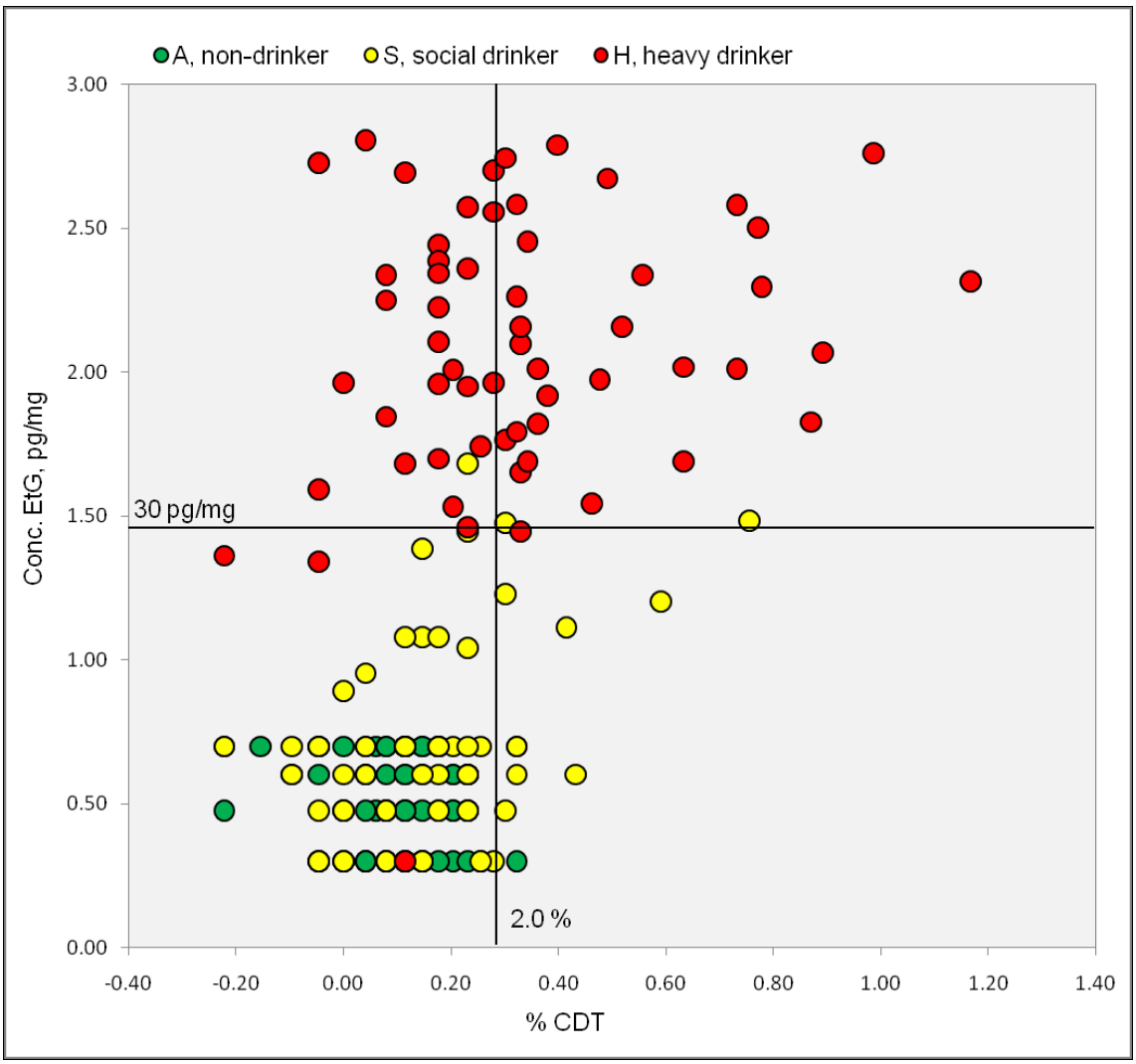


Fig. 3 Concentration of head hair EtG (pg/mg) versus %CDT for 175 patients clinically classified as non-drinkers (A, green dots, N=65), social drinkers (S, yellow dots, N=51), and active heavy drinkers (H, red dots, N=59). The data are reported on a logarithmic scale (base 10)

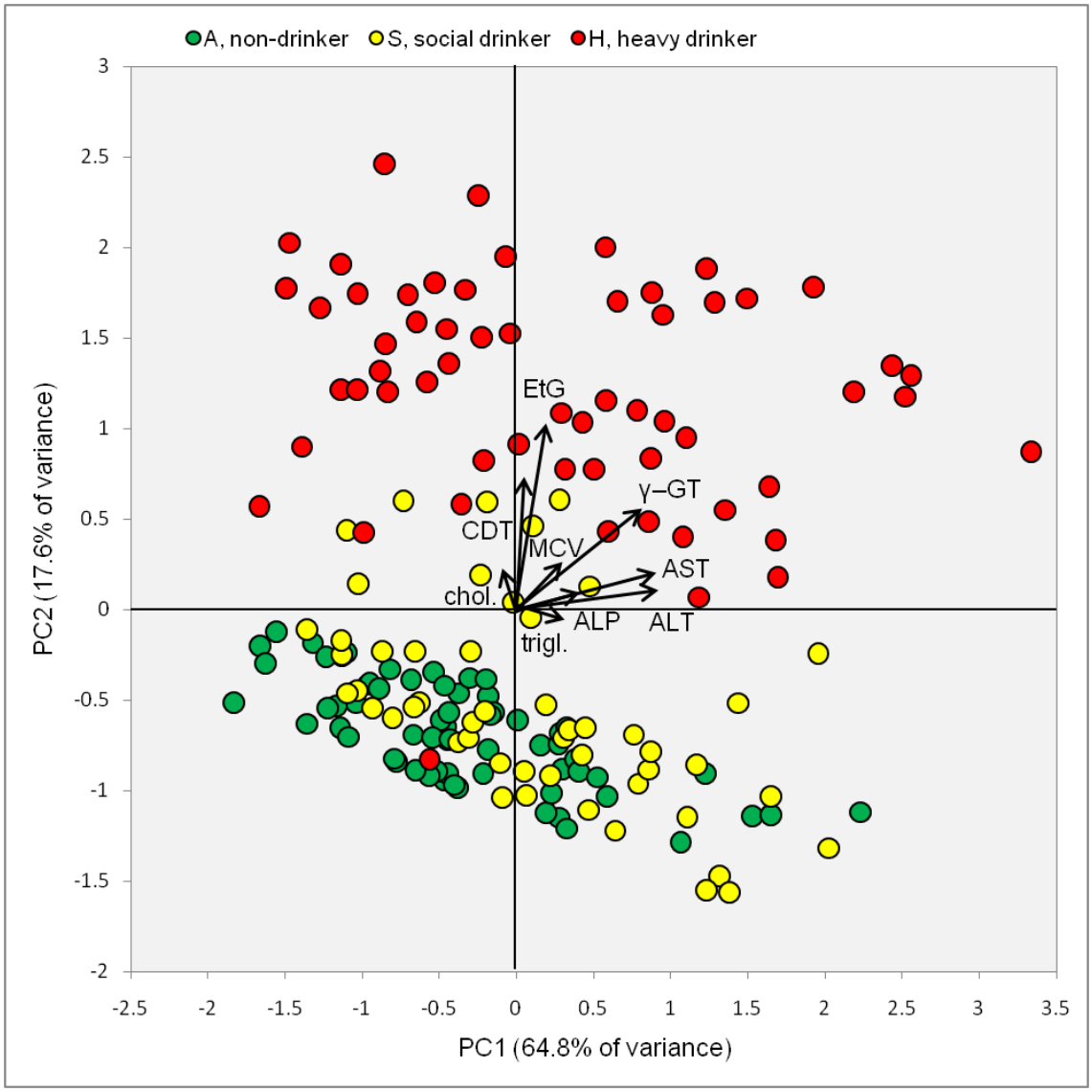
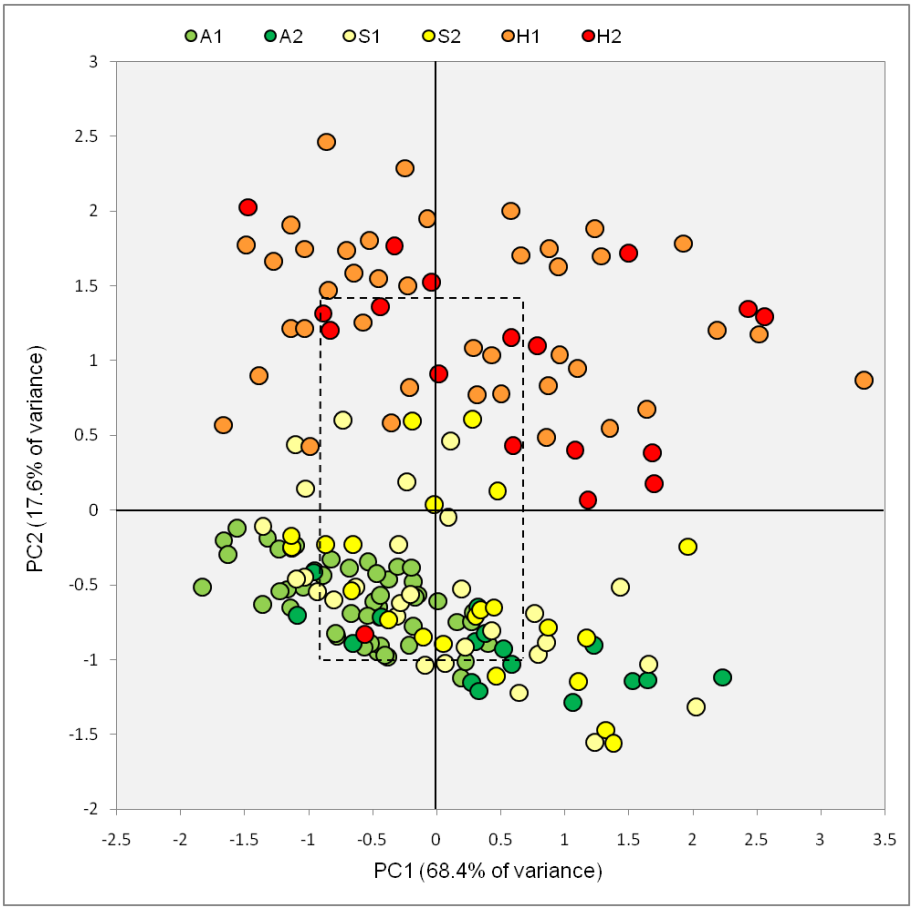


Fig. 4 PCA of 175 patients who were clinically classified as nondrinkers (A, green dots, N=65), social drinkers (S, yellow dots, N=51) and active heavy drinkers (H, red dots, N=59). A score and loading (nine parameters) biplot of PC1 and PC2 is shown



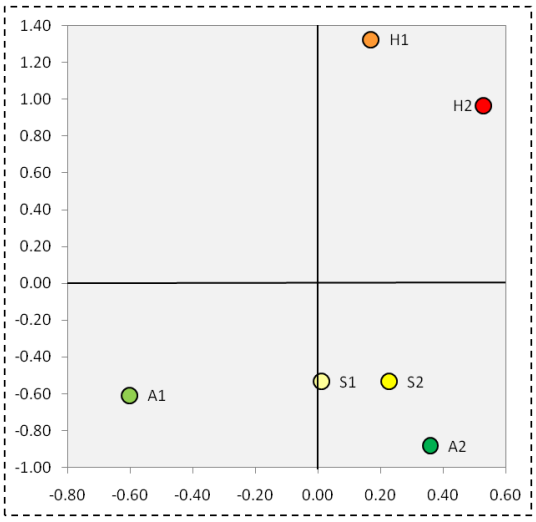


Fig. 5 PCA of 175 patients who were clinically classified as non-drinkers (A1, pale green, not affected by hepatic diseases; and A2, dark green, affected by hepatic diseases), social drinkers (S1, pale yellow, not affected by hepatic diseases; and S2, yellow, affected by hepatic diseases) and active heavy drinkers (H1, orange, not affected by hepatic diseases; and H2, red, affected by hepatic diseases). A score and loading (nine parameters) biplot of PC1 and PC2 is shown. The central area of the graph is expanded on the right to represent the centroid for each category

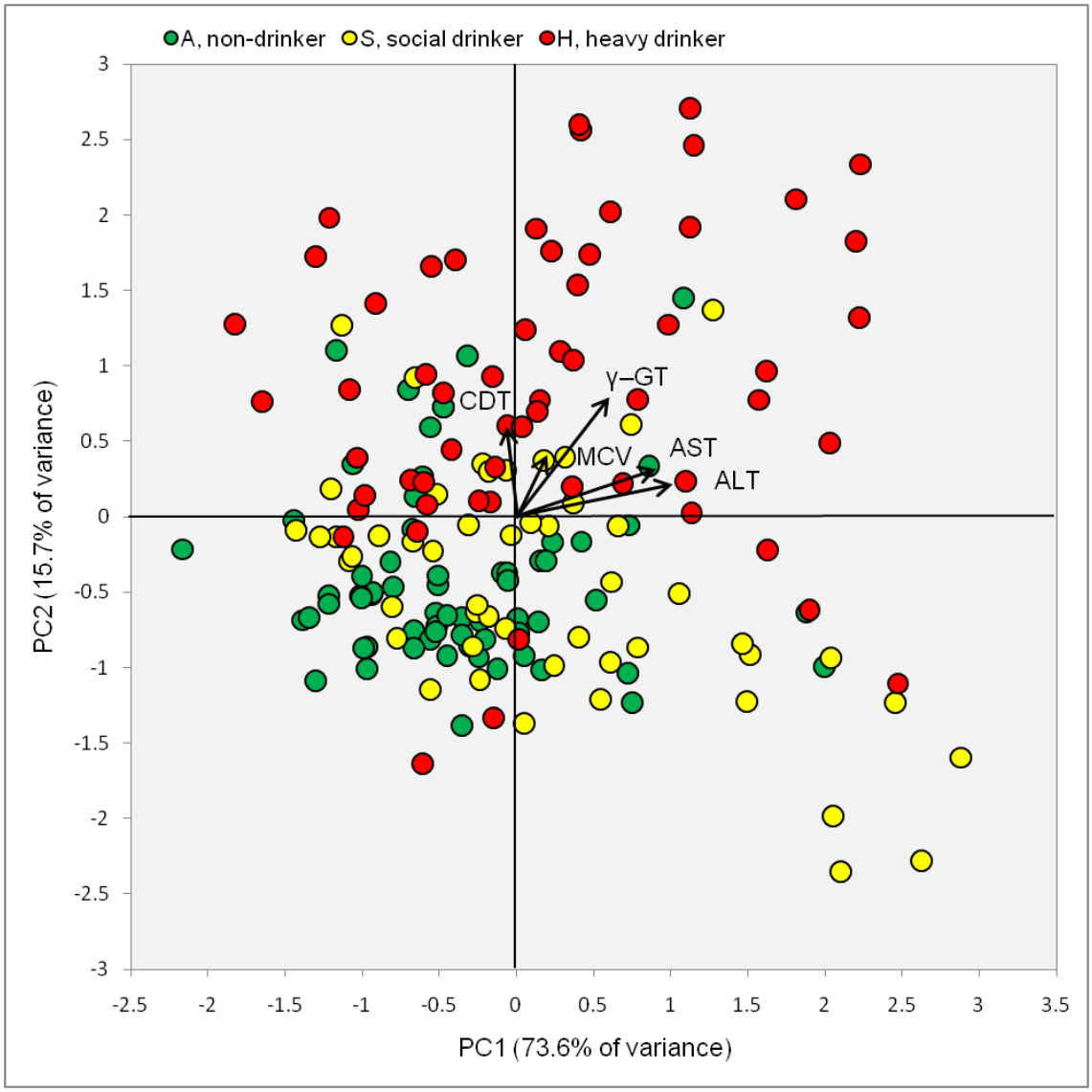


Fig. 6 PCA of 175 patients clinically classified as non-drinkers (A, green dots, N=65), social drinkers (S, yellow dots, N=51) and active heavy drinkers (H, red dots, N=59). A scores and loadings (five parameters) biplot of PC1 and PC2 is shown