Alcohol & Alcoholism, Vol. 26, No. 3, pp. 337–346, 1991
Printed in Great Britain

0735–0414/91 $3.00 + 0.00
Pergamon Press plc

A COMMUNITY SCREENING TEST FOR HIGH ALCOHOL CONSUMPTION USING BIOCHEMICAL AND HAEMATOLOGICAL MEASURES

FRANK VANCLAY, BEVERLEY RAPHAEL*, MICHAEL DUNNE, JOHN WHITFIELD,† TERRY LEWIN‡ and BRUCE SINGH§

Department of Psychiatry, University of Queensland, Brisbane; †Clinical Biochemistry Department, Royal Prince Alfred Hospital, Sydney; ‡Discipline of Psychiatry, University of Newcastle, Newcastle and §Department of Psychological Medicine, Monash University, Melbourne, Australia

(*Received* 12 *March* 1990; *accepted for publication in revised form* 12 *March* 1991)

Abstract — A discriminant function based on a number of biochemical and haematological tests from an extended multiple biochemical analysis and full blood count, together with weight, smoking status and systolic blood pressure is developed. The function was far more effective at detecting high alcohol use (> 40 g ethanol per day) than serum gamma-glutamyl transpeptidase (GGT) or the Short Michigan Alcoholism Screening Test (SMAST) in a community sample of adult males. When classifying high alcohol consumption by GGT only, several division criteria were considered, the most effective being at 40 i.u./l. In terms of identifying high alcohol consumers, rather than alcoholics, the SMAST was no better than GGT, and both had unacceptably low sensitivity (49%, 51%) and poor performance on other measures, thus limiting their use as community screening tools. The discriminant function, however, had an estimated community sensitivity of 78%, was similarly high on other performance measures, and would perform satisfactorily as a community screening tool, particularly in situations where there was a tendency for individuals to under-report their alcohol consumption.

INTRODUCTION

There has been considerable discussion on the impact of alcohol consumption on a number of biochemical and haematological measures. Because there may be considerable incentive for individuals not to be honest about their drinking behaviour, either because of potential external consequences such as employment opportunities, or because of individual reasons, self report of alcohol consumption may not be a satisfactory measure of true alcohol consumption. Furthermore, it is desirable to be able to identify individuals at risk of problem drinking so that preventative measures can be taken as soon as possible, and before social support structures and the individual's social standing are adversely affected by alcoholism. Since many biochemical and haematological measures are affected by chronic and acute ethanol ingestion (Nemesanszky *et al.*, 1988; Hillers *et al.*, 1986*b*; Whitfield *et al.*, 1978*a*, *b*), it is possible that these may be used to determine individuals likely to be high risk drinkers (Whitfield *et al.*, 1981).

Previous research has tended to concentrate on the ability of individual markers to predict alcoholism or high alcohol consumption. The marker most commonly considered is serum gamma-glutamyl transpeptidase (GGT). However, although the probability of being a heavy drinker increases with increasing elevation of GGT, it lacks power as a screening test (Barrison *et al.*, 1987; Chick *et al.*, 1981; Whitehead *et al.*, 1978), and a number of studies have shown that standard questionnaires, in particular the Michigan Alcoholism Screening Test (MAST), its variations, and the

*Author for correspondence at: Department of Psychiatry, University of Queensland, Clinical Sciences Building, Royal Brisbane Hospital, Herston, Queensland 4029, Australia.

CAGE, are more accurate than GGT in identifying alcoholics (Bell and Steensland, 1987; Kristenson and Trell, 1982; Bernadt *et al.*, 1982). Furthermore, elevated levels of GGT could be due to a variety of clinical conditions or drugs, and normal levels vary according to sex, age and body mass (Penn and Worthington, 1983; Whitfield *et al.*, 1978*a*). However, there is some opinion that GGT still has value as a screening measure (Kristenson *et al.*, 1981; Peterson *et al.*, 1983), or at least as a supplement to self-report alcoholism questionnaires (Chick *et al.*, 1981; Gjerde *et al.*, 1987).

Other biochemical and haematological measures such as the erythrocyte mean corpuscular volume (MCV), and aspartate aminotransferase (AST), which have also been proposed as alcoholism indicators, are likely to have similar problems and are not as effective in detecting alcoholics as GGT (Baxter *et al.*, 1982). However, by using a combination of different measures, it is possible that a correct classification of high risk alcohol use could be made that is relatively free of confounding by other variables, and that is more sensitive and specific than any one measure on its own.

A predictive tool is only required in situations where alcoholism is not clinically obvious and where there is potential for the individuals to disguise their true drinking habits. Therefore, the development of a marker to detect high alcohol consumption should be based on a community sample. Much of the previous research has focused on diagnosed alcoholics and therefore does not represent the population in which a community screening test might be diagnostically useful.

This study examined a whole range of biochemical and haematological measures collected by an extended Multiple Biochemical Analysis, including liver enzymes, and full blood count on a community sample of fathers of adolescent boys. A discriminant analysis was undertaken to derive a predictive formula that may be used to identify high drinking adult males. The sensitivity, specificity and other performance measures of the classification model are compared to GGT at various division criteria, and to the Short Michigan Alcoholism Screening Test (SMAST) (Selzer *et al.*, 1975).

METHODS

Sample selection

This analysis is a part of a prospective longitudinal study of adolescent lifestyles conducted in the lower Hunter Valley region of New South Wales, Australia. The original study (phase 1) was undertaken in 1983 and comprised nearly 4000 adolescents drawn from 23 local high schools, together with their mothers (maternal figure living with adolescent) and fathers (paternal figure living with adolescent). Ninety per cent of all enrolled students participated in the study. Fifty-three per cent of adolescents had at least one parent responding. For this study, only data relating to the fathers of adolescent boys were utilized.

Fathers were classified into two groups representing high and low alcohol consumption, with high alcohol consumption being a daily average of 40 g or more. A total of 260 fathers, including all fathers in the high alcohol group ($n = 118$), and a representative random sample of fathers in the low alcohol group ($n = 142$), were invited to participate in phase 2 of the study which comprised intensive interviewing and comprehensive testing including full biochemical and haematological analysis, which was carried out during 1984 and 1985. With refusals and longitudinal attrition, 201 fathers participated in some portion of phase 2, although only 153 fathers agreed to blood tests. Participation rates between the high and low alcohol groups were not significantly different, although it is possible that a lower proportion of high alcohol fathers agreed to participate at phase 1 of the study.

Measurement of alcohol consumption

Alcohol consumption was measured by estimate/frequency technique and by retrospective diary for the seven days prior to the interview. The variable measuring alcohol consumption that was used in this study is a combination of estimate and diary methods. Diary methods are more reliable (Poikolainen *et al.*, 1985) but are applicable only for frequent drinkers. Therefore, for individuals who usually drink at least one day per week, alcohol consumption was based on their diary providing that the diary week was typical in terms of

the number of drinking days. For non-frequent drinkers, i.e. those who usually drink less than one day per week, alcohol consumption was based on their estimated consumption using the estimate/frequency technique. Alcohol was originally determined in terms of grams of ethanol per week. However, individuals were classified into two alcohol consumption categories, less than 280 g per week (40 g per day), and 280 g or more per week.

Alcohol consumption in the low alcohol group ranged from 0 to 235 g ethanol per week (mean = 51; S.D. = 61), with 18 teetotallers (21% of the group). For the high alcohol group, consumption ranged from 288 to 935 g (mean = 474; S.D. = 161). There were no other discernable differences between the two groups. The mean age of fathers was 45 (S.D. = 6), with a range of 33 to 66.

Previous attempts to predict the actual quantity of alcohol consumption have tended not to be successful, although it is likely that multivariate techniques would be successful at classifying individuals as belonging to high or low alcohol consumption groups (Whitfield *et al.*, 1981).

The purpose of this research is to develop a screening test to detect high alcohol consumption in situations where individuals are not likely to be honest about their drinking behaviour. Since this study uses self-reported alcohol consumption, some measurement error could be influencing this study. However, in the methodology of this study there was no *a priori* reason for fathers to be deliberately deceitful about their alcohol consumption. Furthermore, as the discriminant function is only based on those fathers whose alcohol consumption category is the same for the two time periods, fathers would have to be consistent in their drinking and in their deceit in order to be included in the study. Fathers were not aware of the alcohol classification scheme. To the extent the misrepresentation of alcohol consumption would influence these results, it would serve to reduce the reported performance measures. The results reported for this study are therefore conservative.

The Short Michigan Alcoholism Screening Test (SMAST) (Selzer *et al.*, 1975) was part of a self-administered questionnaire administered at phase 2 which covered many aspects of drinking behaviour, lifestyle and personality measures.

Blood collection, biochemical and haematological analysis

Fathers attended field stations at local community health centres. Blood collection occurred in the morning following at least 12 h of fasting. Blood was collected by clean venepuncture by registered nurses. Multiple biochemical analysis covering electrolytes and enzymes was undertaken by Hunter Biochemistry Services, Royal Newcastle Hospital, and haematological analysis was conducted by the Department of Haematology, Royal Newcastle Hospital. The nurses who collected blood samples also took pulse and blood pressure readings, and participants' height and weight measurements.

Since there is a trade-off between sensitivity and specificity (and consequently, the predictive value of the test), and because there is no clinical basis to determine a division criterion for GGT to indicate high consumption of alcohol, division criteria of 30 (Whitehead *et al.*, 1978; Barrison *et al.*, 1987), 35 (Bernadt *et al.*, 1982), 40 (Whitfield *et al.*, 1978*b*; Baxter *et al.*, 1982), and 50 (Bell and Steensland, 1987; Chick *et al.*, 1981; Latcham, 1986; Gjerde *et al.*, 1987) were considered.

Statistics

Alcohol consumption was determined at phase 1 and at phase 2 of the study. Alcohol data were only available for 149 out of the 153 fathers who gave blood. Only the 125 fathers whose alcohol consumption category was the same for both time periods were used in the discriminant analysis. The discriminant analysis was based on the 125 fathers who provided blood samples and whose alcohol consumption was consistent for the two phases of the study. However, discriminant analysis requires complete data for all variables being considered as potential predictors. With the large number of potential predictors being considered (42, Table 1), even with only a small ratio of missing data per variable, the cumulative impact of missing data resulted in only 95 cases with complete data being available to the

Table 1. Alcohol and biochemical and haematological measures

Measure	Unit	Alcohol group Low (n = 75) (mean values)	Alcohol group High (n = 37) (mean values)
Sodium	mmol/L	139.8	139.8
Potassium	mmol/L	4.28	4.19
Chloride	mmol/L	102.4	102.4
Bicarbonate	mmol/L	28.1	27.4
Urea	mmol/L	5.52	4.91*
Creatinine	mmol/L	0.092	0.090
Urate	mmol/L	0.364	0.427‡
Calcium	mmol/L	2.343	2.371
Phosphate	mmol/L	0.923	0.947
Protein	g/L	72.3	73.7
Albumin	g/L	42.3	43.5*
Bilirubin (ln)	μmol/L	2.447	2.408
Alkaline phosphatase	U/L	66.1	73.2
Gamma-glutamyl trans (ln)	U/L	3.147	3.774‡
Creatine kinase	U/L	103.0	104.5
Aspartate aminotrans (ln)	U/L	2.848	3.085‡
Lactate dehydrogenase (ln)	U/L	5.225	5.271
Glucose	mmol/L	4.91	5.11
Transketolase	U/L	0.526	0.558
Thiamin pyro phosphate effect	%	9.5	8.5
Total cholesterol	mmol/L	5.80	5.95
Triglycerides (ln)	mmol/L	0.2016	0.5401†
HDL cholesterol	mmol/L	1.136	1.298*
Total/HDL cholesterol	ratio	5.49	4.85
White blood count	× 10E9/L	6.55	7.07
Red blood count	× 10E12/L	5.129	5.128
Haemoglobin	g/dL	15.55	15.99*
Haematocrit	L/L	0.4540	0.4656
Mean corpuscular volume	fL	88.51	91.69‡
Mean cell haemoglobin	pg	30.37	31.52‡
Mean cell Hb concentration	g/L	342.2	343.5
Red cell distribution (ln)	%	2.5921	2.5881
Platelets	× 10E9/L	270.6	283.2
Mean platelet volume	fL	98.3	97.1
Lymphocytes	× 10E9/L	33.29	33.56
Per cent lymphocytes	%	2.14	2.35
Pulse rate	beats/min	75.2	74.3
Systolic blood pressure	mm Hg	125.5	137.4‡
Diastolic blood pressure	mm Hg	85.6	92.8†
Height	metres	1.750	1.735
Weight	kgs	77.1	82.9*
Body mass index	kgs/m^2	25.10	28.16†
Smoker	0 = no, 1 = yes	0.20	0.49†

*$P < 0.05$, † $P < 0.01$, ‡$P < 0.001$.

discriminant analysis. Since there is a potential that this subset would be different from the full dataset, the discriminant analysis was run numerous times consecutively, deleting non-entered variables with missing data from the predictor list till a final solution was reached. The final solution was based on 112 cases.

All biochemical, haematological and primary health measures such as height, weight, pulse, blood pressure, and smoking status, were accepted as potential discriminating variables. All variables were screened for skew. Assigning a somewhat arbitrary criterion for skew coefficients greater than 1.5 resulted in GGT, AST, lactate dehydrogenase, triglycerides, bilirubin and red cell distribution being natural log-transformed. Box's *M*-test of differences in the covariance matrices between groups, a test of multi-normality, was not significant, although it was significant when these variables were not logged. In order to simulate a community screening situation, no other data transformations or case exclusions were accepted. Discriminant analysis was undertaken using Wilks' method of variable selection. The default values of *F*-to-enter and *F*-to-remove of 1.0 were used, indicating that a variable would be entered if the ratio of between groups variance to within groups for that variable was greater than 1.0.

Discriminant analysis was undertaken using SPSS-X Version 3.1 installed on a VAX 8550 mainframe running VMS 5.1. Despite this analysis producing a linear discriminant function, the SPSS-X discriminant analysis procedure is based on separate co-variance matrices (cf. Ryback *et al.*, 1982).

At phase 1, 13% of the full complement of 838 responding fathers of adolescent boys were in the high alcohol group. Given that it is quite likely that the proportion of male adults in the Hunter Valley in the high alcohol group is larger than this, both because high drinking fathers were perhaps less likely to respond than low drinking fathers, and because fathers in general are less likely to be in the high alcohol group than non-fathers, a notional 5% was added to make the prior probability for the high alcohol group 0.18, and 0.82 for the low alcohol group. An accurate measure of the prior probability of high alcohol consumption is not absolutely necessary, but does improve the predictive power of the discriminant analysis. Comparisons were made between a number of different modification values, with the prior probability of 0.18, 0.82 returning the highest proportion of cases correctly classified and being intuitively acceptable.

The performance measures that were calculated are defined as follows. Sensitivity was defined as the percentage of true high alcohol consumers selected by the screening test. Specificity is defined as the percentage of true low alcohol consumers selected as being in the low alcohol group. The Predictive Value of a Positive Test (PV+) represents the percentage of the group selected by the screening test that are in fact in the high alcohol consumers. The Predictive Value of a Negative Test (PV−) represents the proportion of the group selected as being in the low alcohol group that are, in fact, low alcohol consumers. While sensitivity and specificity can be determined for every sample and can be generalized, PV+ and PV− are based on the number of falsely classified cases, and therefore are dependent on distribution of cases. Therefore, they cannot usually be determined for many studies, because of inadequate sampling. PV+ and PV− can only be determined using an unstratified community sample, or by the use of weighting factors to represent the stratified groups in their correct proportions. For a community screening test, the most important performance measures are the sensitivity and PV+.

RESULTS

Fathers' alcohol category was significantly associated with a number of biochemical and haematological measures (Table 1).

GGT (natural log transformed) was the variable that differentiated the alcohol consumption categories the most with a Wilks' lambda of 0.761, and had far greater differentiating power than AST ($\lambda = 0.897$) or MCV ($\lambda = 0.867$). In fact, urate ($\lambda = 0.779$), mean cell haemoglobin ($\lambda = 0.884$) and systolic blood pressure ($\lambda = 0.884$) had greater or equivalent differentiating power to AST or MCV. GGT on its own however, did not perform well as a marker of high alcohol

consumption, no matter which division criterion was used (Table 2). The SMAST performed only marginally better than GGT.

The discriminant analysis performed very well, far exceeding GGT and the SMAST, and obtained a sensitivity of 81%, a PV+ on this data set of 88%, with a projected community sensitivity of 78% and PV+ of 57% (Table 3). Other measures of the performance of the function were also high. The eigenvalue (ratio of between groups to within groups variability in the discriminant scores) was 1.72; the canonical correlation was 0.79; Wilks' lambda was 0.37, which was highly significant ($P < 0.0001$), and 63% of variance in the discriminant scores was attributable to the grouping variable.

In SPSS-X, the discriminant analysis is based on the canonical discriminant function. Unfortunately, the complex mathematical nature of canonical discriminant analysis limits the ease of use of the classification procedure (Appendix). However, Fisher's linear discriminant functions, which are inherently more portable, are also provided. Classification using Fisher's functions is based on comparing the scores on the two functions. The case is classified as belonging to the group whose function returns the largest discriminant score. The two functions appear in Table 4.

DISCUSSION

In keeping with other studies (Kristenson and Trell, 1982; Bell and Steensland, 1987; Bernadt *et al.*, 1982; Whitehead *et al.*, 1978; Barrison *et al.*, 1987; Chick *et al.*, 1981; Latcham, 1986), GGT did not perform well as a marker of high alcohol consumption in its own right in this community study. The highest estimated sensitivity for GGT used as a community screening test for high alcohol consumption was 69%, which occurred for a division criterion of 30 i.u./l, but had an unacceptably low PV+ of 35%. Division criteria of 40 and 50 yielded PV+ values of 47%,

Table 2. Performance measures for GGT and the SMAST for this sample and community estimates

	GGT criterion (i.u./l)				
	30	35	40	50	SMAST
Sample*					
Sensitivity	70.7	53.7	53.7	36.6	57.5
Specificity	75.0	83.3	90.5	92.9	89.5
PV+	58.0	61.1	73.3	71.4	74.2
PV−	84.0	78.7	80.0	75.0	80.0
Community*					
Sensitivity	68.9	53.3	51.1	35.6	49.2
Specificity	72.1	80.8	87.5	91.3	86.3
PV+	35.2	37.9	47.3	47.4	42.7
PV−	91.3	88.7	89.1	86.6	90.9

*Sample figures are the actual results based on fathers whose drinking was consistent across the two phases. For GGT, $n = 125$. To provide comparability, the SMAST sample figures are based on the same group as for GGT, except for 9 fathers who failed to complete the SMAST. Community figures are estimates derived using the phase 2 alcohol classification only (GGT, $n = 149$; SMAST, $n = 180$), with fathers in the low consumption category weighted to reflect the actual proportion of low alcohol consumers in the community (82%). The weighting factor represents the differential sampling fraction between the low and high alcohol consumption groups. For GGT, the weighting factor was 1.97. For the SMAST the weighting factor was 2.45.

Table 3. Classification of alcohol consumption status by discriminant analysis

Sample*		Predicted group	
Actual group		Low alcohol	High alcohol
Low alcohol (n = 75)		71	4
High alcohol (n = 37)		7	30
Sensitivity	= 30/(30+7)	= 81.1%	
Specificity	= 71/(71+4)	= 94.7%	
PV+	= 30/(30+4)	= 88.2%	
PV−	= 71/(71+7)	= 91.0%	
Community*		Predicted group	
Actual group		Low alcohol	High alcohol
Low alcohol (n = 187)		163	24
High alcohol (n = 41)		9	32
Sensitivity	= 32/(32+9)	= 78.0%	
Specificity	= 163/(163+24)	= 87.2%	
PV+	= 32/(32+24)	= 57.1%	
PV−	= 163/(163+9)	= 94.8%	

*Sample figures are the actual results based on fathers whose drinking was consistent across the two phases, whose data were complete for the variables in the discriminant function, and who defined the function (n = 112). Community figures are estimates derived using fathers with complete biochemical data and using the phase 2 alcohol classification (n = 135), with fathers in the low consumption category weighted by 1.99 to reflect the actual proportion of low alcohol consumers in the community (82%).

but had unacceptably low sensitivity, 51% and 36% respectively. Because PV+ does not increase between division criteria of 40 and 50, but sensitivity deteriorates considerably, we concur with Hambidge (1987) that a division criterion of 50 is excessively high. Since the sensitivity does not change greatly between division criteria of 35 and 40, but the PV+ improves considerably, 40 appears to be the most desirable division criterion if GGT is to be used as a measure on its own. However, a criterion of 40 would fail to detect half the high alcohol consumers (sensitivity), and as a community screening test, over 52% of the selected group would be false positives (PV+).

Previous research has reported that various questionnaire methods for detecting alcoholism have higher sensitivity than GGT (Bernadt *et al.*, 1982; Kristenson and Trell, 1982; Bell and Steensland, 1987), but that they are poor at identifying high alcohol consumption in groups other than known alcoholics (Peterson *et al.*, 1983). In this community study, the sensitivity and PV+ of the SMAST at detecting high alcohol consumption in the community was no different to that of GGT at the 40 i.u./l criterion. However, the SMAST is a measure of alcoholism, not of high alcohol consump-

Table 4. Fisher's linear discriminant functions

	Low alcohol consumption	High alcohol consumption
GGT (ln U/L)*	16.58019	18.29798
Total/HDL chol	4.607653	3.614331
Triglycerides (ln mmol/L)	22.71204	25.96742
Systolic BP (mm Hg)	1.058830	1.135510
Smoker (0 = no, 1 = yes)	39.78796	42.98949
Chloride (mmol/L)	24.29264	24.78484
Urate (mmol/L)	139.8794	160.3261
Albumin (g/L)	9.094861	9.477791
Alk phosphatase (U/L)	0.5323822	0.5768011
Potassium (mmol/L)	2.370934	−0.005053316
Haemoglobin (g/dL)	−7.371042	−8.425513
Weight (kgs)	0.6385306	0.7003300
HDL cholesterol (mmol/L)	106.7612	111.0708
TPP effect (%)	−0.5129924	−0.6178344
Creatinine (mmol/L)	112.9086	72.13842
Bilirubin (ln μmol/L)	59.53700	60.91435
Red blood count (× 10E12/L)	44.51560	46.25953
(Constant)	−1813.541	−1897.391

*In order of entry to the discriminant function.

tion, and therefore it is not surprising that its sensitivity at detecting high alcohol consumers is low, although it is surprising that the PV+ was not higher. On the full data set (n = 180), there were 17 low alcohol consumers who were classified as being possible or probable alcoholics by the SMAST, representing 15% of the low alcohol group. An examination of these men reveals that they were previous high drinkers and have reduced their alcohol consumption because of alcohol related problems. Eleven of these 17 fathers had sufficient non-missing blood tests to be included in the classification using discriminant analysis. Three of these 11 fathers were classified by the discriminant function as being in the high alcohol group.

The resultant discriminant function based on 14 biochemical and haematological measures, together with weight, smoking status, and systolic blood pressure, was greatly superior to GGT and the SMAST at determining high alcohol consumption. Although some men who have changed their alcohol consumption over the period of the study were classified according to their new alcohol category, it is evident that the discriminant function is not sensitive to very recent changes in alcohol consumption. This is a good feature and does indicate that the reported sensitivities and specificities will be understated if the interest is in detecting those men who have prolonged high alcohol consumption.

Fisher's linear discriminant functions (Table 4), once produced, are very easy to use and do not need computer technology. The values from the biochemical analyses are multiplied by the appropriate coefficients. The resultant scores are summed for each of the Fisher's functions. The function with the largest score identifies the likely group membership for the individual being classified. While the multiplication and summation tasks are probably too complex to be undertaken by hand, at least when speed and reliability are considerations, the functions could be programmed into pocket programmable calculators or into the computer of the biochemical laboratory and calculated directly and printed on the laboratory's computer output along with the individual biochemical results.

There are a number of features of the discriminant function that may reduce its generalizability in other populations.

1. Smoking, for example, was highly correlated with drinking, with 49% of those in the high alcohol group regarding themselves as a smoker, and only 20% in the low alcohol group being a smoker. Since smoking has significant health impacts, and effects on many of the biochemical and haematological measures used in this study (Hillers *et al.*, 1986*b*), the extent to which the proportion of smokers in each alcohol group varies in other populations will limit the applicability of the function.

2. In this study, 18% of males were high alcohol consumers. However, in other communities, the proportion of males in the high alcohol group (the prior probability) may be different. While the canonical discriminant function (Appendix) allows the user to specify the prior probabilities, the Fisher functions (Table 4) would need to be recalculated.

3. Because many biochemical measures are affected by age and sex (Whitfield *et al.*, 1978*a*), a screening test based on this discriminant function may not be sensitive for younger age groups or for women (Bliding *et al.*, 1982). This study was undertaken on a group of fathers, the youngest of whom was 33. Although it would be possible to do a similar analysis for women, in this study only 1.2% of mothers drank more than 40 g of alcohol per day. With such a low proportion of women with high alcohol consumption, mass community screening would not be economic or efficient, and other methods to detect these women should be determined.

4. Because the discriminant function was generated from only 112 cases, and largely tested on these cases, the estimates of the coefficients in the discriminant function (Table 4, Appendix) and the performance measures are subject to some error. Nevertheless, it is clear that inclusion of variables such as blood pressure, smoking and weight, in a multivariate approach including laboratory data, does increase the probability of successful discrimination between the high and low alcohol consumption groups.

5. The discriminant analysis presented includes commonly requested biochemical and

haematological measures and health status indicators. However, one measure, thiamin pyro phosphate effect (TPP effect), is expensive and labour intensive to determine. TPP effect is the fourth-to-last variable to enter the function, and though significant, the individual contribution of this variable is likely to be slight. Consequently, there would be little impact on the predictive capacity of the discriminant function if TPP effect was not included, should TPP effect be considered too expensive or inconvenient to use in a general screening test. Although the discriminant functions presented are suggested as being appropriate for adoption elsewhere, ideally individuals or institutions interested in this methodology should develop their own function based on a sample of the relevant population. It may be the case that for other populations TPP effect will not be a significant variable in the function.

This study has shown that accurate identification of adult male high alcohol consumers is possible using a range of biochemical, haematological and other health related measures. A discriminant function is presented that may be used in community screenings of adult males, although individual factors in other populations may reduce the performance of this function. However, if large scale screening of high alcohol consumption were to take place, the study could easily be replicated to provide new parameter values for the predictor variables. While diary or other self-report measures may be more cost efficient, and equally reliable, a method for determining alcohol consumption is necessary in situations where individuals are likely to disguise their alcohol consumption or when it is unavailable or missing. The function could be incorporated into the standard computer output of biochemical laboratories, provided that information relating to smoking and blood pressure was also available.

Acknowledgements — Multiple biochemical analysis was undertaken by Dr John Dickeson, Director, Hunter Biochemistry Services, Royal Newcastle Hospital. Haematological analysis was undertaken by Dr Arno Enno, Director of Haematology, Royal Newcastle Hospital. Kath Heggart was responsible for the majority of phase 2 interviewing. Jane Walton and Anne Clarke collected the blood samples. Dr Gerard Byrne and Dr Russell Viner participated in discussions relating to the clinical importance of the various biochemical and haematological measures.

The research was supported by grants from the National Health and Medical Research Council, The Australian Associated Brewers, the Commonwealth Department of Health and Community Services, and the Victorian Health Promotion Council.

REFERENCES

Barrison, I. G., Ruzek, J. and Murray-Lyon, I. M. (1987) Drinkwatchers — description of subjects and evaluation of laboratory markers of heavy drinking. *Alcohol and Alcoholism* **22**, 147–154.

Baxter, S., Fink, R., Leader, A. R. and Rosalki, S. B. (1982) Laboratory tests for excessive alcohol consumption evaluated in general practice. *British Journal of Alcohol and Alcoholism* **15**, 164–166.

Bell, H. and Steensland, H. (1987) Serum activity of gamma-glutamyl-transpeptidase (GGT) in relation to estimated alcohol consumption and questionnaires in alcohol dependence syndrome. *British Journal of Addiction* **82**, 1021–1026.

Bernadt, M. W., Mumford, J., Taylor, C., Smith, B. and Murray, R. M. (1982) Comparison of questionnaire and laboratory tests in the detection of excessive drinking and alcoholism. *Lancet* **1**, 325–328.

Bliding, G., Bliding, A., Fex, G. and Tornqvist, C. (1982) The appropriateness of laboratory tests in tracing young heavy drinkers. *Drug and Alcohol Dependence* **10**, 153–158.

Chick, J., Kreitman, N. and Plant, M. (1981) Mean cell volume and gamma-glutamyl-transpeptidase as markers of drinking in working men. *Lancet* **1**, 1249–1251.

Gjerde, H., Amundsen, A., Skog, O. J., Morland, J. and Aasland, O. G. (1987) Serum gamma-glutamyl-transferase: an epidemiological indicator of alcohol consumption? *British Journal of Addiction* **82**, 1027–1031.

Hambidge, D. M. (1987) Gamma-glutamyl-transpeptidase and mean cell volume in alcoholics. *British Journal of Psychiatry* **150**, 568.

Hillers, V. H., Alldredge, J. R. and Massey, L. K. (1986*a*) Determination of habitual alcohol intake from a panel of blood chemistries. *Alcohol and Alcoholism* **21**, 199–205.

Hillers, V. H., Massey, L. K. and Alldredge, J. R. (1986*b*) Effects of alcohol use and other health practices on blood chemistries and blood pressure. *Alcohol and Alcoholism* **21**, 207–12.

Kristenson, H. and Trell, E. (1982) Indicators of alcohol consumption: comparisons between a questionnaire (Mm-MAST), interviews and serum γ-glutamyl transferase (GGT) in a health survey of middle-aged males. *British Journal of Addiction* **77**, 297–304.

Kristenson, H., Trell, E. and Hood, B. (1981) Serum-γ-glutamyl-transferase in screening and continuous control of heavy drinking in middle-aged men. *American Journal of Epidemiology* **114**, 862–872.

Latcham, R. W. (1986) Gamma-glutamyl-transpeptidase and mean corpuscular volume: their usefulness in the

assessment of in-patient alcoholics. *British Journal of Psychiatry* **149**, 353–356.

Nemesanszky, E., Lott, J. A. and Arato, M. (1988) Changes in serum enzymes in moderate drinkers after an alcohol challenge. *Clinical Chemistry* **34**, 525–527.

Penn, R. and Worthington, D. J. (1983) Is serum γ-glutamyl-transferase a misleading test? *British Medical Journal* **286**, 531–535.

Peterson, B., Trell, E. and Kristenson, H. (1983) Comparison of γ-glutamyl-transferase and questionnaire test as alcohol indicators in different risk groups. *Drug and Alcohol Dependence* **11**, 279–286.

Poikolainen, K., Karkkainen, P. and Pikkarainen, J. (1985) Correlations between biological markers and alcohol intake as measured by diary and questionnaire in men. *Journal of Studies in Alcohol* **46**, 383–387.

Ryback, R. S., Eckardt, M. J., Rawlings, R. R. and Rosenthal, L. S. (1982) Quadratic discriminant analysis as an aid to interpretive reporting of clinical laboratory tests. *Journal of the American Medical Association* **248**, 2342–2345.

Selzer, M. L., Vinokur, A. and Rooijen, L. van (1975) A self-administered Short Michigan Alcoholism Screening Test (SMAST). *Journal of Studies in Alcohol* **36**, 117–126.

Whitehead, T. P., Clarke, C. A. and Whitfield, A. G. W. (1978) Biochemial and haematological markers of alcohol intake. *Lancet* **1**, 978–981.

Whitfield, J. B., Allen, J. K., Adena, M. and Gallagher, H. G. (1981) A multivariate assessment of alcohol consumption. *International Journal of Epidemiology* **10**, 281–288.

Whitfield, J. B., Hensley, W. J., Bryden, D. and Gallagher, H. (1978*a*) Effects of age and sex on biochemical responses to drinking habits. *Medical Journal of Australia* **2**, 629–632.

Whitfield, J. B., Hensley, W. J., Bryden, D. and Gallagher, H. (1978*b*) Some laboratory correlates of drinking habits. *Annals of Clinical Biochemistry* **15**, 297–303.

APPENDIX
CANONICAL DISCRIMINANT FUNCTION

Discriminant score = −29.35854

+0.6222934	GGT (ln U/L)
−0.3598450	Total/HDL cholesterol
+1.179309	Triglyceride (ln mmol/L)
+0.02777840	Systolic BP (mm Hg)
+1.159803	Smoker (0 = no, 1 = yes)
+0.1783070	Chloride (mmol/L)
+7.407107	Urate (mmol/L)
+0.1387218	Albumin (g/L)
+0.01609140	Alkaline phosphatase (U/L)
−0.8607357	Potassium (mmol/L)
−0.3819970	Haemoglobin (g/dL)
+0.02238772	Weight (kgs)
+1.561198	HDL cholesterol (mmol/L)
−0.03798052	Thiamin pyro phosphate effect (%)
−14.76958	Creatinine (mmol/L)
+0.4989618	Bilirubin (ln μmol/L)
+0.6317646	Red blood count (× 10E12/L)

The probability of an individual belonging to the high alcohol group is determined using the following formula.

$$P\ (\text{high}/D) = \frac{0.18 \times P(D/\text{high})}{0.18 \times P(D/\text{high}) + 0.82 \times P(D/\text{low})}$$

where:

P (high/D) is the posterior probability, i.e. the probability of the case belonging to the high alcohol group with a Discriminant score of D.

$P(D/\text{high})$ is the conditional probability, i.e. the probability of a Discriminant score D belonging to the high alcohol group.

$P(D/\text{low})$ is the conditional probability of the Discriminant score D belonging to the low alcohol group.

0.18, 0.82 are the prior probabilities.

The conditional probabilities are determined by referring to normal distribution tables using the following parameters.

	Mean	S.D.
High alcohol group	1.8485	0.992
Low alcohol group	−0.9119	1.004