Acta Genet Med Gemellol 35: 23-33 (1986)

Received 10 October 1985
Final 15 November 1985

Genetic Variation and Plasma Creatine Kinase Activity

J.B. Whitfield[1], N.G. Martin[2]

[1] *Department of Clinical Biochemistry, Royal Prince Alfred Hospital, Camperdown, NSW, Australia; and* [2] *Department of Human Genetics, Medical College of Virginia, Richmond*

Abstract. The distribution of plasma creatine kinase (CK; E.C.2.7.3.2) is known to be skewed, and this has made it difficult to analyse the sources of variation. We have studied plasma CK in 206 pairs of twins and have analysed the results after separating them into what appear to be two Gaussian frequency distributions. The results in the main distribution (CK < 300 iu/l) are apparently affected by genetic factors common to both men and women, and by environmental factors which are of much greater effect in men. The tendency for some men to have very high CK values may also have a genetic basis.

Key words: Creatine kinase, Genetic factors, Repeatability, Twins

INTRODUCTION

There is a considerable range of values for plasma CK, particularly in younger subjects and those who exercise frequently or strenuously [5,9]. This complicates the diagnostic use of CK measurements, to some extent in the assessment of possible myocardial infarction [12], and to a considerable degree in the detection of carriers of Duchenne muscular dystrophy [6].

A number of studies have shown, however, that there is a tendency for each person's CK to remain within their own individual range [13,16]. The sources of this difference between people are largely unexplored but one study on twins [10] provides evidence for heritability of this characteristic.

One problem which must be overcome in investigations of the reference range,

repeatability, or genetics of creatine kinase, is the non-Gaussian nature of the frequency distribution, since most methods depend on a reasonably normal curve. Attempts to find transformations which produce normality (in the statistical sense) have been unsuccessful [11] despite a range of mathematical manoeuvres. We believe that it may be more productive to consider the observed frequency distributions as the sum of two (or more) normal distributions, to estimate the means and standard deviations of each, and if they are sufficiently separated, to exclude results which are probably not derived from the main population. We have done this for CK results from a sample of twins and have been able to arrive at acceptable models of variation and estimates of genetic and environmental effects, which could not be achieved by the use of transformations. Separate consideration of the subjects excluded by truncation allows tentative conclusions about them also.

It appears from these results that there is an additive genetic component of variance which is common to both men and women, despite the large differences between values in the two sexes; that the heritability of plasma CK levels in women is high; and that the tendency for some men to have values two to six times the modal value may have a genetic component also.

SUBJECTS AND METHODS

Subjects. A total of 206 pairs of monozygotic (MZ) and dizygotic (DZ) twins, aged between 18 and 34 years (mean 23.1) were recruited from the Australian NH & MRC Twin Registry for a study of alcohol metabolism and susceptibility to intoxication [8]. There were 42 MZ male pairs, 43 MZ female, 38 DZ male, 44 DZ female and 39 DZ pairs of opposite sex. Zygosity was determined as previously described [14], primarily by blood typing. Both members of a twin pair attended on the same day. A total of 89 twins (50 men and 39 women) attended on more than one occasion, with an interval between visits of 1 to 17 months (mean 4.5 months), and the results of these 89 are used to assess the repeatability of the measurements within an individual. Blood was taken for this and other measurements before the subjects ingested any alcohol.

Methods. Heparinised blood was centrifuged within two hours of venipuncture and plasma was stored at -20 C for up to two days before analysis. Creatine kinase activity was measured on a Centrifichem 300 (Union Carbide, Tarrytown, New York) using Calbiochem reagents (Calbiochem-Behring, La Jolla, California).

RESULTS AND DISCUSSION

Frequency Distributions

The descriptive statistics of the frequency distributions for men and for women are shown in Table 1. It will be seen that the normal frequency distribution cannot provide an acceptable description of the observed results in either sex. Log-transformation failed to convert the results to a normal distribution, as has been found previously [11].

Since it is known that exercise may increase the plasma CK, it seemed likely that the observed distributions might be due to the combination of one major frequency distribution and a smaller number of subjects whose CK values are higher because of

Table 1. The Distribution of Plasma CK Results, by Sex, in the Entire Sample; Showing the Estimated Statistics on Two Different Assumptions About the Nature of the Distribution, the Kolmogorov Statistic T, and the Probability that the Sample Data Could Have Come from Such a Distribution

	Men	Women
All subjects; single normal distribution		
N	197	213
Mean (iu/l)	192	95
SD	140	65
Skewness	2.5	6.5
Kurtosis	7.2	64.5
Goodness-of-fit[a]	$T = 0.21$ $p < 0.01$	$T = 0.15$ $p < 0.01$
Two-distribution hypothesis		
N	170 27	199 14
Mean (iu/l)	150 487	85 261
SD	60 167	28 164
Goodness-of-fit[a]	$T = 0.07$ $p > 0.1$	$T = 0.05$ $p > 0.1$

[a] Kolmogorov test for goodness-of-fit to postulated distribution or combination of distributions.

this factor. This can be generalised to cover many variables with a relationship of the type:

$$y_i = (m + r_i) + b_i x_i$$

where y_i is the observed value of y for the ith individual, m is the mean of the baseline population, r_i is the derivation from the mean m due to all factors except x. x_i is the value of some independent variable (in this case postulated to be the amount of exercise but in other situations perhaps the amount of alcohol consumed or cigarettes smoked); it may be negligible in most people. b_i is the degree of response to the variable x and, unlike the usual regression coefficient, it too may vary between people.

The data might then fit acceptably to a combination of two curves with different means and variances. Since in this case we do not know the value of x for each person, we have to rely on there being little overlap between the two postulated groups so that we can study genetic and environmental influences on the residuals (r in the equation). The cumulative percentage frequency distributions were calculated and plotted on normal probability paper [7], as shown in Figs. 1a and 1b. A single Gaussian frequency distribution gives a straight line on such a plot and it will be seen that at least two lines are required, giving support to the concept of two distributions. The next step is to arrive at estimates of the mean, variance, and numbers of subjects for each distribution in order to best account for the observed data. The lines have been drawn in a way which minimises the Kolmogorov statistic [2], which is the maximum difference between the observed and expected cumulative frequency distributions, and the means

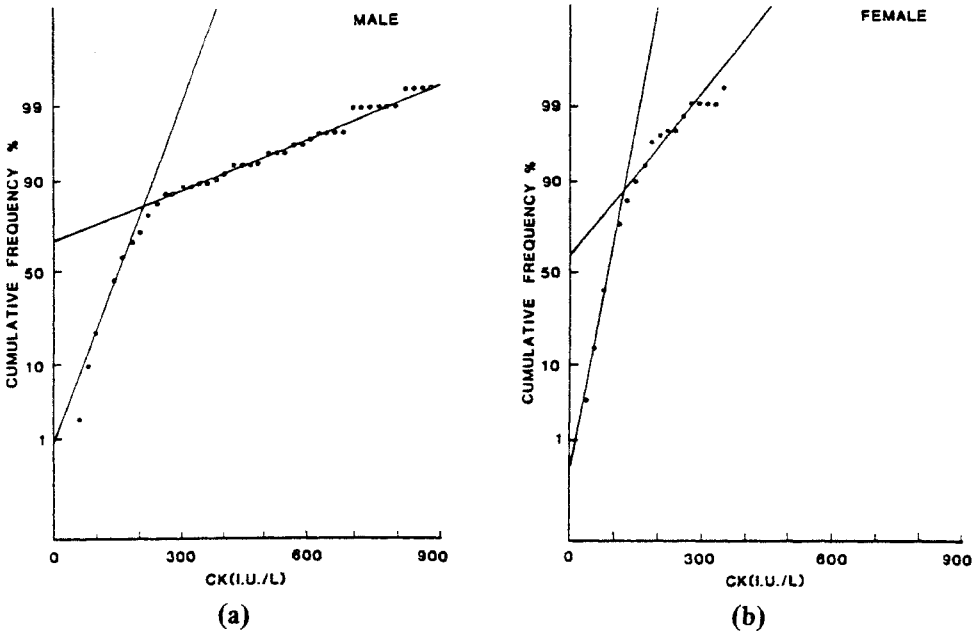


Fig. 1. Cumulative frequency distribution for plasma CK in (a) men and (b) women, showing the fit of the data to two Gaussian distributions (straight lines).

and standard deviations can be estimated from the percentiles of the estimated distributions. The results for men and women are shown in Table 1; from these results we can estimate main populations with mean and SD of 150 +/− 55 iu/1 for men and 85 +/− 28 iu/1 for women. Less than 1% of subjects from these main distributions will have values greater than 300 iu/l. We have therefore considered the consequences of truncation of both male and female values at 300 iu/l.

Repeatability

The results for those subjects who attended on two occasions were used in an analysis of variance to calculate the within-and between person components of variance and the intraclass correlation, which is the between-person variance as a proportion of the total. The results of this are shown in Table 2, separately for male and female subjects, using truncation at 300 iu/l for the reasons discussed above.

Significant differences between individuals exist in both men and women, with the intraclass correlation being about 0.4 in both. These results are similar to those of van Steirteghem et al [13] and of Williams et al [15] for men in the same age-group, but are rather different for the women. Recalculating the data of Williams et al, who did not use intraclass correlation, for 18-35 year olds, gives an intraclass correlation of 0.59 for men and 0.23 for women.

Table 2. Components of Variance (s^2_w, Within-Person; s^2_b, Between-Person) and Intraclass Correlation (R_i), Calculated from Repeat Determinations on 89 Subjects Studied on Two Occasions

	s^2_w	s^2_b	R_i
Men	1900	1516	0.45**
Women	693	455	0.40**

** $P < 0.01$.

Genetic and Environmental Effects on Variation

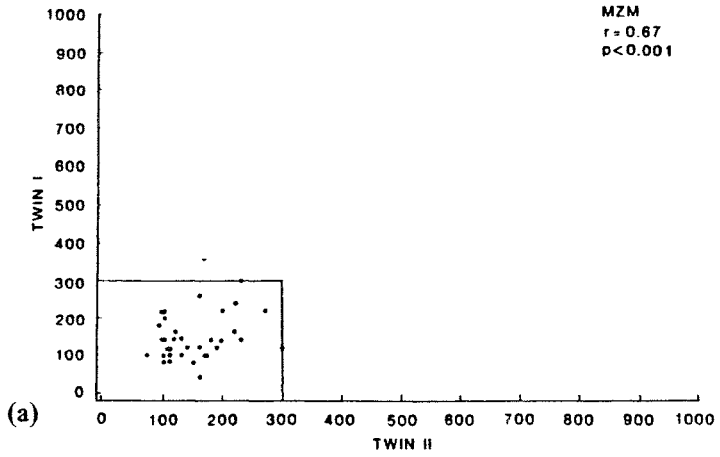
The ascription of total variance to genetic and environmental causes may be achieved by comparison of the similarity of pairs of twins of different zygosity. This may be done in various ways using either the correlations, or the mean squares within and between pairs, of the various sex and zygosity groups. The within-pair and between-pair mean squares for the twins grouped according to sex and zygosity may be used in a model-fitting procedure [3] which tests various hypotheses on the sources of variation and which results in a preferred model and estimates of the variance from additive genetic (V_A), shared environmental (E_2), or individual environmental (E_1) sources.

However, the model makes a number of assumptions which, if violated, will cause rejection of all models of variation. It is necessary that the data are approximately normally distributed and that the total variance is not significantly different between MZ and DZ twins of the same sex. In practice, we have found that some skewness in the data makes little difference to the conclusions [eg,14], but a severely skewed distribution may result in apparent differences in variance between groups when a few outliers happen to fall into one group rather than another. This can be seen with the CK data.

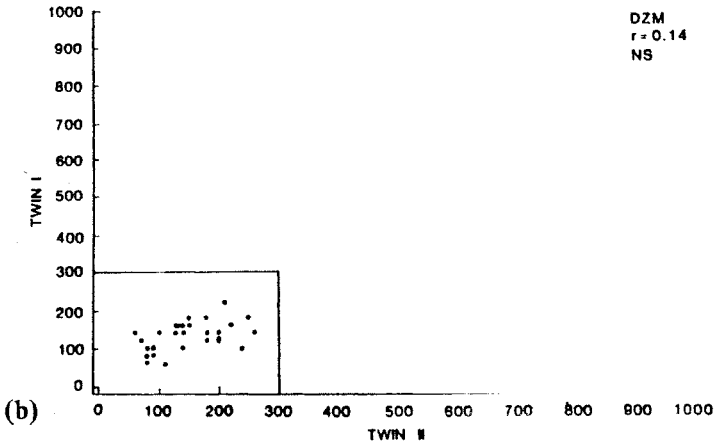
The within-pair and between-pair mean squares for each of the sex and zygosity groups, with and without truncation, are shown in Table 3. Some of che characteristics of the five groups are more readily appreciated by referring to Figs. 2a to 2e, which

Table 3. Within- and Between-Pair Mean Squares (and Their Associated Degrees of Freedom), by Sex and Zygosity

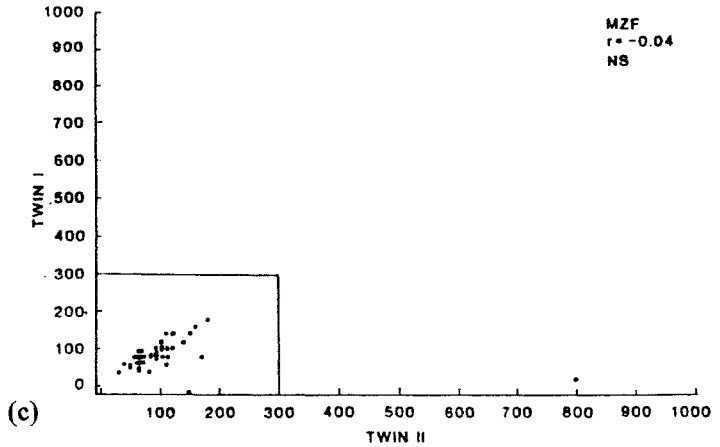
		All data	Truncated at 300 iu/l
MZ male	Between	37272 (41)	4172 (33)
	Within	7359 (42)	2231 (34)
MZ female	Between	7782 (41)	2024 (39)
	Within	8143 (42)	283 (40)
DZ male	Between	19170 (35)	3640 (28)
	Within	14405 (36)	1823 (29)
DZ female	Between	2499 (43)	2499 (43)
	Within	917 (44)	917 (44)
DZ opposite sex	Between	15075 (38)	3551 (35)
	Within	7734 (38)	1923 (35)


MZM
r = 0.67
p<0.001
TWIN I
TWIN II
0
100
200
300
400
500
600
700
800
900
1000

(a)


DZM
r = 0.14
NS
TWIN I
TWIN II
0
100
200
300
400
500
600
700
800
900
1000

(b)


MZF
r = -0.04
NS
TWIN I
TWIN II
0
100
200
300
400
500
600
700
800
900
1000

(c)

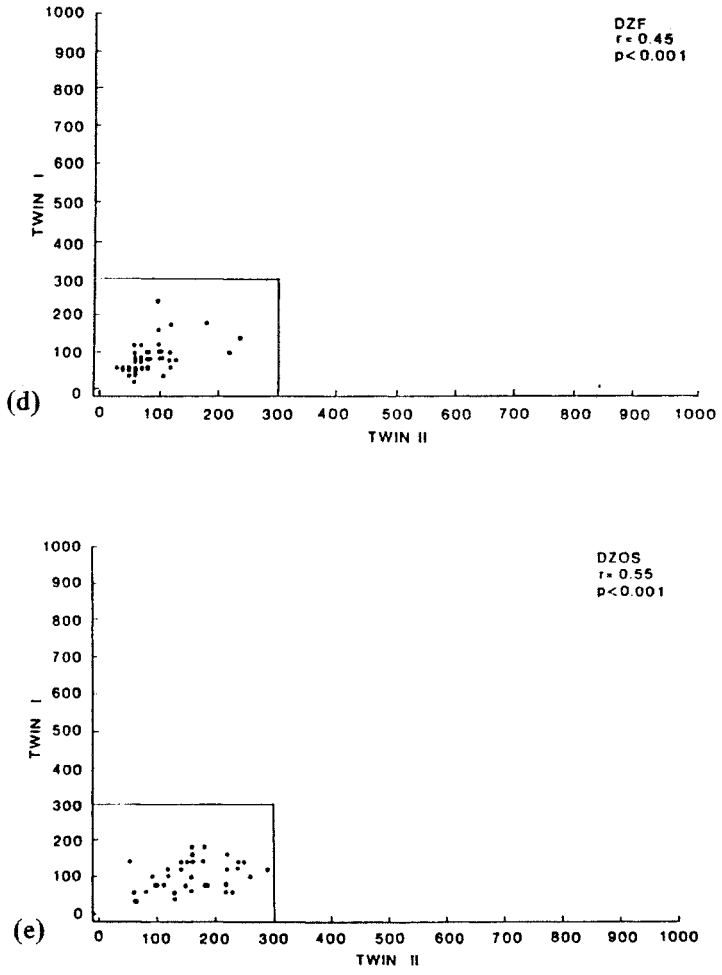


Fig. 2. Similarity of members of twin pairs, grouped by zygosity: (a) MZM, monozygotic male pairs (b) DZM, dizygotic male pairs (c) MZF, monozygotic female pairs (d) DZF, dizygotic female pairs (e) DZOS, dizygotic pairs of opposite sex.

plot the results for the two members of each twin pair against each other. A number of points are of interest. Firstly, with the untruncated data the MZ and DZ female pairs differ greatly in the total variance; this effect is greatly reduced when the presumed outliers are exluded. It can also be seen, from Figs. 2c and 2d, why the correlations found with the untruncated data are so misleading; two MZ female pairs are so discordant (one member above and one below 300) that they make the excellent agreement below 300 insignificant overall. Secondly, the variance for the male pairs drops greatly when those with high CKs are excluded, and the proportions of the within- and between-pairs mean squares alter. These results (Figs. 2a, 2b; Table 3) are the opposite of those for the women, showing significant correlation overall for the MZ pairs but poorer correlation in the < 300 range.

The results of model-fitting to the untruncated data are shown in Table 4. It will be seen that for the men, models containing only E_1, or E_1 and E_2, effects are rejected and the $E_1 V_A$ model is accepted, indicating a mixture of non-shared enrivonmental and additive genetic sources of variation, with an apparent heritability of 0.62. For women, on the other hand, no model is acceptable and reference to the mean squares in Table 3 shows that this is because the DZ pairs are more similar than the MZ.

Table 4. Goodness-of-Fit of Alternative Models of Genetic and Environmental Sources of Variation, on Raw and Truncated Data

		Chi-square	df	P
Raw data (no truncation)				
Male pairs only	E_1	25.98	3	< 0.001
	$E_1 E_2$	7.97	2	0.019
	$E_1 V_A$	2.17	2	0.338
Female pairs only	E_1	38.09	3	< 0.001
	$E_1 E_2$	39.74	2	< 0.001
	$E_1 V_A$	38.38	2	< 0.001
Truncated at 300 iu/l				
Male pairs only	E_1	6.63	3	0.085
	$E_1 E_2$	0.47	2	0.792
	$E_1 V_A$	1.36	2	0.505
Female pairs only	E_1	30.86	3	< 0.001
	$E_1 E_2$	11.53	2	0.003
	$E_1 V_A$	2.06	2	0.356
All pairs	E_1	48.97	9	< 0.001
	$E_1 E_2$	32.42	8	< 0.001
	$E_1 V_A$	40.28	8	< 0.001
Sex-dependent estimates, truncated				
$E_1 V_{Am} V_{Af} V_{Amf}$		35.02	6	< 0.001
$E_{1m} E_{1f} V_A$		4.47	7	0.724
$E_{1m} E_{1f} V_{Am} V_{Af} V_{Amf}$		4.36	5	0.499

When only the results from what may be termed the main distribution (<300 iu/l) are considered, a different picture emerges. For the female pairs, the $E_1 V_A$ model is chosen, while in the men none of the models is rejected. Combination of the data from all types of twin pair produces rejection of all models based on common variance component estimates for both sexes, but the use of sex-specific values is more successful.

There is no necessary reason for the components of variance to be the same in men and women, although for many variables this is what is found. For a variable such as CK, which has very different means and variances for the two sexes, different estimates of the variance components are necessary. The full model for sex-specific variance components has been presented elsewhere [1] and submodels were fitted as previously described. Table 4 shows the outcome when different values for E_1 (E_{1M} and E_{1F}) or for V_A (V_{AM}, V_{AF} and V_{AMF}) are allowed, and it will be seen that a single estimate of additive genetic variance with different nonshared environmental components for men and women gives a good result. Addition of shared environmental effects to the model resulted in only a trivial decrease in chi-square.

It therefore seems that there is a genetic component in a person's individual plasma CK level, and that it is quantitatively the same in men and women. The proportion of the total variance – the heritability – differs greatly between men and women because of very different individual environmental effects, which are about six times greater in men than women in this sample.

There was much stronger evidence for genetic variance in men in the untruncated than in the truncated data. Indeed, the best-fitting model to male CK levels in the 'normal' (<300) range was purely environmental. The tendency to high CK levels (<300), therefore, appears to have a strong genetic component in males. This is illustrated by the fact that out of 12 MZ male twins with CK values above 300 iu/1, 8 were from four pairs and 4 had a discordant cotwin (Fig. 2a). All 7 of the male DZ twins with a CK above 300 iu/l had a cotwin with a CK lower than this limit (Fig. 2b). The possible genetic tendency to high CK levels in males could be due either to genetic variance in the amount of exercise taken, or to the response of plasma CK to exercise being affected by genetic factors. Unfortunately, the data available do not allow us to discriminate between these hypotheses.

CONCLUSIONS

Conclusions about the causes of variation in CK levels are not straightforward. The distribution of CK levels is highly skewed in both sexes and cannot be normalised by logarithmic transformation. It is known that physical exercise raises plasma CK markedly and it is likely that this factor accounts for some of the very high values seen in the males. Nevertheless, the male data, including all the high values, are consistent with a simple model of variation in which around 60% of variation is genetic in origin and the remainder is due to individual environmental factors. Exclusion of pairs with one or both twins having a CK level greater than 300 iu/1 reduces not only the environmental, but also the genetic variance, so that any of the models are consistent with the reduced data set.

In females, by contrast, no model is consistent with the full data set but exclusion

of two MZ pairs containing individuals with very high values (340 & 800 iu/l) yields mean squares which are consistent with additive genetic variance accounting for 81% of female CK variance.

The truncated (<300 iu/1) data in male, female and opposite-sex pairs are consistent with a simple model in which there is a common genetic component of variation in both sexes, but a very much larger environmental contribution to variance in males, perhaps reflecting much greater variation in physical exercise in men. However, this result (Table 5) is dependent upon our decision to truncate the data at this particular

Table 5. Estimates of Variance Components and Heritabilities for the Accepted Model of Variation, and the Significance of the Differences of the Estimates from Zero

Source	Estimate	P
E_{1m}	1997	< 0.001
E_{1f}	290	< 0.001
V_A	1217	< 0.001

$h_m^2 = 0.38 +/- 0.06$
$h_f^2 = 0.81 +/- 0.05$

point, which was based on the observed frequency distribution. This eliminated the considerable genetic variation which occurs at the top of the male range.

This study of the sources of variation in plasma CK has implications for the use of CK measurements. The wide range encountered in healthy men may make the definition of abnormality difficult. The fact that the E_1 estimate is greater than the within-person variation indicates that there are environmental influences, not shared within families, which remain constant for each person over time; identification of these could help to refine the reference range.

In women, especially younger women such as our subjects, interest centres more on the identification of possible carriers of the sex-linked Duchenne muscular dystrophy in selected high-risk groups. It has recently been argued [6] that the great overlap between presumably normal women and obligatory carriers rules out the use of CK determinations both for individual counselling and for the testing of hypotheses about mutation rates in a sex-linked lethal condition. However, if the heritability of plasma CK is high in women, as we have found, then comparison of the levels in suspected carriers with those in their sisters may be more useful than comparison with a reference population of normals. Such a comparison would have to take account of the chance (up to 50%, depending on the mutation rate) that the sister of a carrier will herself be a carrier. A formula for determination of the probability of carrier status, to include information on CK levels in female relatives, has been put forward [4], and the concept is strengthened by the demonstration of genetic similarity between siblings.

Acknowledgments. The Australian Twin Registry is supported by he National Health and Medical

Research Council, and the study for which the twins attended was supported by the Australian Associated Brewers. We are grateful to the twins themselves, and to the staff of the Department of Clinical Biochemistry, Royal Prince Alfred Hospital, for their cooperation.

REFERENCES

1 Clarke P, Jardine R, Martin NG, Stark AE, Walsh RJ (1980). Sex differences in the inheritance of some anthropometric characters in twins. Acta Genet Med Gemellol 29:171-192.
2. Conover WJ (1971). Practical Non-parametric Statistics. Wiley, New York, p. 295.
3. Eaves LJ, Last KA, Young PA, Martin NG (1978). Model-fitting approaches to the analysis of human behaviour. Heredity 41:249-320.
4. Emery AEH, Holloway S (1977). Use of normal daughter's and sister's creatine kinase levels in estimating heterozygosity in Duchenne muscular dystrophy. Hum Hered 27:118-126.
5. Griffiths PD (1966). Serum levels of ATP creatine phosphotransferase (creatine kinase), the normal range and effect of muscular activity. Clin Chim Acta 13:413-420.
6. Gruemer H-D, Miller WG, Chinchilli VM, et al (1984). Are reference limits for serum creatine kinase valid in detection of the carrier state for Duchenne muscular dystrophy? Clin Chem 30:724-730.
7. Hoffman RG (1963). Statistics in the practice of medicine. JAMA 185:864-873.
8. Martin NG, Perl J, Oakeshott JG, Gibson JB, Starmer GA, Wilks AV (1985). A twin study of ethanol metabolism. Behav Genet 15:93-109.
9. Meltzer HY (1971). Factors affecting serum creatine phosphokinase levels in the general population: the role of race, activity and age. Clin Chim Acta 33:165-172.
10. Meltzer HY, Dorus E, Grunhaus L, Davis JM, Belmaker R (1978). Genetic control of human plasma creatine phosphokinase activity. Clin Genet 13:321-326.
11. Miller WG, Chinchilli VM, Gruemer H-D, Nance W (1984). Sampling from a skewed population distribution as exemplified by estimation of the creatine kinase upper reference limit. Clin Chem 30:18-23.
12. Sylen JCH, Jansson E, Brandt S, Kallner A (1983). Specificity of cardiac enzymes in diagnosis of chest pain in marathon runners. Lancet ii:1505.
13. Van Steirteghem AC, Robertson EA, Young DS (1978). Variance components of serum constituents in healthy individuals. Clin Chem 24:212-222.
14. Whitfield JB, Martin NG (1983). Determinants of variation in plasma alkaline phosphatase activity: a twin study. Amer J Hum Genet 35:978-986.
15. Williams GZ, Widdowson GM, Penton J (1978). Individual character of variation in time-series studies of healthy people. II. Differences in values for clinical chemical analyses in serum among demographic groups, by age and sex. Clin Chem 24:313-320.

Correspondence: Dr. J.B. Whitfield, Dept. of Clinical Biochemistry, Royal Prince Alfred Hospital, Camperdown, NSW 2050, Australia.