Heterozygote detection in phenylketonuria
Measurement of discriminatory ability and interpretation of the phenylalanine loading test by determination of the heterozygote likelihood ratio

A. WESTWOOD* and D. N. RAINÉ†

Summary. Two tests of heterozygosity for phenylketonuria, the phenylalanine/tyrosine ratio in the plasma after a fast and the response to an oral load of phenylalanine, have been compared in obligate heterozygotes for phenylketonuria and apparent normal homozygotes. The discriminatory ability of the tests is measured using both parametric and non-parametric methods, the correlation between them is determined, and the reproducibility of the loading test is examined.

Because complete separation between the genotypes cannot be achieved by these tests, a method of interpretation which avoids possibly erroneous categoric statements and which enables the family history of a subject and his test result to be combined in a quantitative and meaningful way is described.

In the last decade there has been a growing interest in the detection of heterozygous carriers of autosomal and sex-linked recessive diseases. This is required not only when genetic advice is sought before marriage or further pregnancies but it may also be important in assessing the chance of fetal abnormality when abortion is contemplated, and, if it can be perfected, it may allow a more extended control of inherited metabolic disease than is possible at present.

In normal circumstances these heterozygotes are phenotypically and clinically normal and identical with homozygotes for the normal (wild-type) gene, but they can sometimes be detected by careful biochemical examination. For autosomal diseases the most satisfactory tests are those in which the enzyme is studied directly, but when a suitable tissue is not available, as in phenylketonuria where the deficient enzyme, phenylalanine hydroxylase, only occurs in the liver, tests of heterozygosity must assess the enzyme activity indirectly as by the phenylalanine loading test (Hsia et al, 1956) or fasting plasma phenylalanine/tyrosine ratio (Perry et al, 1967b).

In the present study a standard form of phenylalanine loading test, which does not differ in any important respect from those used in a number of other centres in Britain, is proposed; the best parameters for discrimination of phenylketonuria heterozygotes from normal homozygotes are derived; and a method of interpreting the test which avoids potentially erroneous categoric statements by reporting carrier status as a probability is described.

Materials and methods

After an overnight fast a sample of capillary blood was taken, at approximately 09.00 h, into an heparinized tube. A load of 7 g L-phenylalanine was then given orally (as a slurry in strong orange or lemon juice) and capillary blood sampled hourly for the next 4 h. This dose is for adult subjects; children 3 a and over may be tested using a dose of 100 mg/kg body weight (200 mg/kg under 3 a) but as yet we have no experience of the test in children. The blood samples were centrifuged immediately and the plasma stored at −20° (at which temperature it is stable for at least 6 months) until analysed.

Plasma phenylalanine and tyrosine were determined by the method of Cooke and Raine (1970) using a Technicon Sequential Multi-sample (TSM) amino-acid

---

Received 21 October 1974.
* Biochemistry Department, Institute of Mental Subnormality, Lea Castle Hospital, Kidderminster.
† Clinical Chemistry Department, The Children’s Hospital, Birmingham.
analysers. Slight modifications, precipitation of proteins with a sulphasalicylic acid solution and electronic range expansion in the chart recorder (x3), allowed analysis of small samples of capillary blood. Estimation of the precision of the method from replicate analyses of pools of plasma preparations over a period of 6 months, as part of a quality control programme, gave, for phenylalanine and tyrosine, coefficients of variation of 7.4% and 7.8% in normal plasma, and 5.3% and 11.2% in phenylketonuric plasma.

**Results**

The tests as described above were carried out on 22 apparently normal individuals (presumed normal homozygotes) with no family history of phenylketonuria and 26 parents (obligate heterozygotes) of phenylketonuric children (Table I and Fig. I). The response of one individual in the normal group (C.G. in Table I) to the phenylalanine load was

<table>
<thead>
<tr>
<th>Subject</th>
<th>Plasma Phenylalanine (mmol/l)</th>
<th>Plasma Tyrosine (mmol/l)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0h</td>
<td>1h</td>
</tr>
<tr>
<td>Normal homozygotes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L.B.</td>
<td>0.062</td>
<td>0.680</td>
</tr>
<tr>
<td>S.O.</td>
<td>0.060</td>
<td>1.025</td>
</tr>
<tr>
<td>J.W.</td>
<td>0.068</td>
<td>0.455</td>
</tr>
<tr>
<td>A.M.</td>
<td>0.060</td>
<td>0.503</td>
</tr>
<tr>
<td>A.Mc.</td>
<td>0.100</td>
<td>0.491</td>
</tr>
<tr>
<td>E.W.</td>
<td>0.059</td>
<td>0.890</td>
</tr>
<tr>
<td>C.G.</td>
<td>0.090</td>
<td>1.048</td>
</tr>
<tr>
<td>F.G.</td>
<td>0.075</td>
<td>1.035</td>
</tr>
<tr>
<td>P.M.</td>
<td>0.056</td>
<td>1.280</td>
</tr>
<tr>
<td>K.M.</td>
<td>0.058</td>
<td>0.800</td>
</tr>
<tr>
<td>S.P.</td>
<td>0.044</td>
<td>0.860</td>
</tr>
<tr>
<td>G.B.</td>
<td>0.066</td>
<td>0.920</td>
</tr>
<tr>
<td>A.S.</td>
<td>0.040</td>
<td>0.902</td>
</tr>
<tr>
<td>S.D.P.</td>
<td>0.120</td>
<td>1.030</td>
</tr>
<tr>
<td>J.C.</td>
<td>0.066</td>
<td>0.916</td>
</tr>
<tr>
<td>B.M.</td>
<td>0.151</td>
<td>0.715</td>
</tr>
<tr>
<td>D.T.</td>
<td>0.066</td>
<td>0.864</td>
</tr>
<tr>
<td>P.S.</td>
<td>0.072</td>
<td>0.642</td>
</tr>
<tr>
<td>A.W.</td>
<td>0.056</td>
<td>0.827</td>
</tr>
<tr>
<td>M.S.</td>
<td>0.070</td>
<td>1.020</td>
</tr>
<tr>
<td>M.J.</td>
<td>0.050</td>
<td>0.830</td>
</tr>
<tr>
<td>M.S.J.</td>
<td>0.060</td>
<td>0.900</td>
</tr>
<tr>
<td></td>
<td>0.069</td>
<td>0.829</td>
</tr>
<tr>
<td></td>
<td>0.026</td>
<td>0.208</td>
</tr>
</tbody>
</table>

*Eliminated from the normal group (see test).*
markedly different from the other normals on two separate occasions (the phenylalanine concentrations at three and four hours after the load were more than five standard deviations above the normal mean). This individual, probably an unsuspected phenylketonuria heterozygote, was therefore excluded from the normal group.

**Discriminatory ability of heterozygote tests.** Most methods of heterozygote detection measure a variable which is continuously distributed in the two genotypes, heterozygotes and normal homozygotes. There will usually be significant overlap between the genotypes, and in order to compare different methods and to select that with most discriminatory ability some means of measuring the separation between them and the size of the overlap is required. When the test has a Gaussian distribution in both genotypes a good measure of the discriminatory power is given by the generalized distance between the means (Penrose, 1951). This is a function which increases in magnitude with the discriminatory ability of the test, and from it an estimate of the percentage of each group overlapping the midpoint between them can be obtained. It is determined by dividing the difference between the two sample means by the arithmetic mean of the two standard deviations: $\bar{x}$(het) − $\bar{x}$(norm)/$\frac{1}{2}$s(het) + $\frac{1}{2}$s(norm). As Penrose (1951) points out, the mean standard deviation gives a better estimate of the within-group standard deviation than the more conventional methods used in significance tests since, in this application where the variances will usually differ markedly, these would give too much weight to the larger variance.

The overlap between the genotypes has also been estimated by a method not dependent on Gaussian distributions. In this graphical method each test result in each distribution is plotted against its estimated centile (Fig. 2). The point at which the two distributions intersect gives the test result at which there is the smallest equal percentage overlap between the two groups. Overlap determined by this

---

**TABLE II**

<table>
<thead>
<tr>
<th>Test</th>
<th>Overlap estimated from:</th>
<th>Generalized Distance</th>
<th>Centiles</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fasting (0 h) levels</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>32</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Tyrosine</td>
<td>48</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>Phe/tyr ratio</td>
<td>22</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td><strong>Loading test</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenylalanine at 1 h</td>
<td>41</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>2 h</td>
<td>11</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>3 h</td>
<td>6</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>4 h</td>
<td>5</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Tyrosine at 1 h</td>
<td>25</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>2 h</td>
<td>35</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>3 h</td>
<td>42</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>4 h</td>
<td>46</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Phe/tyr ratio at 1 h</td>
<td>30</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>2 h</td>
<td>14</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>3 h</td>
<td>16</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>4 h</td>
<td>16</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Log 4 h phenylalanine</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>
non-parametric method has been compared, in the
tests of heterozygosity for phenylketonuria, with
that estimated from the generalized distance between
the means (Table II).
This clearly shows that the most discriminating
parameters are the plasma phenylalanine concentra-
tions at three and four hours after the load, and that
the tyrosine concentrations are in general poor dis-
criminants. Combination of the phenylalanine and
tyrosine concentrations as a ratio improves the dis-
criminatory ability of both the fasting and one hour
tests. But there is no improvement with these
ratios at 2, 3, and 4 h, probably because they give
too much weight to the relatively poor tyrosine dis-
criminants. Improved discriminatory ability could
almost certainly be obtained by combining the
phenylalanine and tyrosine concentrations by dis-
criminant function analysis (Wang et al, 1961).
The best single discriminant is the plasma phenyl-
alanine concentration at four hours after the load, and
its use in the logarithmic transformation, also a good
discriminant (Table II), has an additional advantage
which will be discussed later.
To assess the reproducibility of the loading test it
was carried out, in a shortened form in which only
the 4 h phenylalanine concentration was measured,
in 22 individuals on two occasions separated by 1
month to 3a (Fig. 3). The calculated correlation
coefficient \( r \) is 0.84 and the standard deviation of
the second test from any given value of the first test
was 0.13, which is 5.4\% of the normal mean and
4.6\% of the heterozygote mean.

The fasting plasma phenylalanine tyrosine ratio
and the loading test are both assumed to reflect
phenylketonuria carrier status because they are
indirect measures of hepatic phenylalanine hydroxy-
lase, so in the present study the regression between
them, which would be expected to be highly signi-
ficant, has been tested (Fig. 4). Comparison of the
calculated regression coefficient \( b = 0.90 \) with zero
gave \( t = 3.79 \), for which \( p < 0.1\% \) \( (n = 47) \).

Heterozygote likelihood ratio (HLR). Over-
lap occurs between the phenylketonuria heterozy-
gotes and normal homozygotes even using the best
parameters of the phenylalanine loading test, and
this creates a number of difficulties in the interpre-
tation of this and other similar tests (Westwood and
Raine, 1973). It is important to consider both a
priori probability of heterozygosity and the actual
test result, and a number of quantitative methods
have been used some of which define limits between
which classification cannot be made because the
probability of error is too high (Hsia and Steinberg,
1960; Renwick et al, 1960). Alternatively a critical
value for the test, around which classification can be
made and for which the classification error is known,
can be determined (Gold et al, 1973), or the exact
probability that a given individual is heterozygous
may be determined (Wilson et al, 1965; Murphy and
Mutalik, 1969). We have preferred the last method
(Westwood and Raine, 1973), and calculate this
probability from the ratio, at the observed test re-
sult, of the two probability density functions (for
the heterozygote and normal distributions), cor-
rected for the a priori probability of heterozygosity.

![Graph](https://example.com/graph.png)

**FIG. 3.** Examination of the reproducibility of the phenylalanine
loading test, which was repeated on two separate occasions in 22 indi-
viduals. The measured linear regression (shown) of the second test
\((y)\) on the first \((x)\) is given by \(y = 0.52 + 0.79x\), and the correlation
coefficient \((r)\) is 0.84.

**FIG. 4.** Comparison of the fasting ratio and loading test in phenyl-
ketonuria heterozygotes (A) and normal homozygotes (○). The
measured linear regression (shown) of the ratio \((y)\) on the loading test
\((x)\) is given by \(y = 0.90x - 0.993\), and is highly significant \((p < 0.1\%\).
The correlation coefficient \((r)\) is 0.49.
In practice, the heterozygote test, in this case the phenylalanine concentration in the plasma 4 h after a phenylalanine load, is applied to a number of normal subjects (presumed normal homozygotes) and to as many parents of children with the disease (obligate heterozygotes) as possible. The mean and standard deviation of each group is calculated, and when the test is applied to an individual of unknown genotype the probability density function (y) for each distribution at the observed test result is found from \( y = \frac{1}{\sqrt{2\pi s^2}} \exp \left( -\frac{(x - \bar{x})^2}{2s^2} \right) \). The ratio of these two functions, \( y(\text{heterozygote})/y(\text{normal}) \), is called the heterozygote likelihood ratio (HLR). In fact, if this ratio is plotted graphically against the test result (Fig. 5) this procedure is somewhat simplified, and the advantage of a test which has approximately equal variance in each genotype (such as the log 4 h phenylalanine concentration in contrast to the simple concentration) can be seen. If the variances differ markedly this graph will not be a straight line but may be hyperbolic or parabolic with the result that individuals who fall in the lower tail of the normal distribution will appear to be more likely to be heterozygous than those near the normal mean in spite of the fact that they are further from the heterozygote distribution.

Correction of the heterozygote likelihood ratio to account for the a priori probability of heterozygosity is achieved by simple multiplication. For example, suppose subject A in Fig. 5, whose 4 h phenylalanine concentration was 0.45 mmol/l and whose likelihood ratio (from the graph) is 5, has no family history of phenylketonuria. If the incidence of heterozygotes in the normal population is 0.02, and of normal homozygotes is 0.98, then his corrected ratio becomes 0.1 (5 × 0.02/0.98), giving odds of 10 to 1 against heterozygosity. Conversely, if he has a sib with phenylketonuria the ratio becomes 10 (5 × 0.67/0.33), because two out of three clinically normal sibs of phenylketonurics will be heterozygotes, giving odds of 10 to 1 on heterozygosity. These odds can of course be expressed as a percentage probability or a decimal, and it remains to be seen in which form the individual in question will prefer to receive the information.

It may be difficult to obtain sufficient obligate heterozygotes to characterize the chosen test accurately, especially if the disease is rare or only recently described. It is of value therefore to assess the smallest number of heterozygotes required to give a good estimate of the mean and standard deviation. This has been examined by taking the obligate heterozygotes for phenylketonuria as they accumulated in our laboratory and re-calculating the mean and standard deviation as each new subject, after the first two, was added to the series (Fig. 6). This shows that although at first these parameters fluctuate, their value after only the 10th or 12th subject has settled down almost to a constant value. The opportunity to improve the estimates by testing more heterozygotes as they become known should always be taken, but it is useful to know that the parents of only five or six patients with a given disease may provide sufficient information to allow the heterozygote likelihood ratio to be used.

**Discussion**

Phenylalanine hydroxylase, the enzyme deficient in phenylketonuria, has not been found in a readily accessible tissue, and indirect methods of assaying the hepatic enzyme have to be used. Hsia et al.
regression et al, 1961), and the authors later reported that the separation was further improved when phenylalanine and tyrosine were determined simultaneously on an amino-acid analyser (Perry et al, 1967a).

In the present study the discriminatory ability of these two tests of heterozygosity for phenylketonuria has been compared using both the discriminatory index suggested by Penrose (1951), which assumes that the distributions being compared are Gaussian, and a graphical method using centiles (see Fig. 2). The main limitation of the latter method is the small sample size, particularly when the overlap is small since fewer results will then contribute to the determination of the overlap. Nevertheless, the agreement between the two methods was good (Table II), and they clearly showed that the most discriminating test was the loading test if the plasma phenylalanine concentrations at 3 and 4 h after the load are used as discriminants. Like Renwick et al (1960) we find the logarithmic transformation of the 4 h concentration to be particularly useful since it is not only a good discriminant (Table II) but also simplifies interpretation of the test. Although the simple combination of phenylalanine and tyrosine concentrations as ratios significantly improved the fasted and 1 h parameters the discriminatory ability of the other phenylalanine concentrations was reduced, presumably because too much weight is given to the relatively poorer tyrosine discriminants. Undoubtedly, better discriminants could be obtained by discriminant function analysis to give optimal weights to the various phenylalanine and tyrosine concentrations (Wang et al, 1961), but the increase in discriminatory ability might be small compared with the five or 10 times increased workload in amino-acid analyses this would require.

Since to some degree, both the loading test and fasting ratio reflect hepatic phenylalanine hydroxylase activity it is not surprising that the measured regression between them is highly significant (Fig. 4). However, the correlation coefficient of less than 0.5 illustrates the dependence of both tests on factors other than this enzyme activity. These are not difficult to envisage. There are known to be diurnal variations in plasma phenylalanine (Gültler et al, 1969) and the tyrosine is lowered in pregnant women or those taking oral contraceptives, sufficiently to raise the fasting ratio into the heterozygote range (Jackson et al, 1971; Rose and Cramp, 1972; Yakymysyn et al, 1972). The loading test will be influenced by variables such as the rate of absorption, movement into extracellular spaces and metabolism of phenylalanine by pathways in addition to that involving hydroxylation. In addition, the enzyme itself only indirectly reflects the presence or absence of a gene and will be subject to metabolic processes not related to the gene.

Significant overlap between the genotypes will almost certainly remain while such methods are used to test for heterozygosity. The test results must therefore be assessed by proper statistical methods if errors are to be avoided or at least kept to the absolute minimum. A number of methods have been proposed, including the calculation of limits between which classifications are not made because the probability of error is too high (Hsaia and Steinberg, 1960; Renwick et al, 1960), use of a 'counselling curve' or classification line about which errors are minimized and are known (Gold et al, 1973), or determination of the probability of heterozygosity for a given test result and family history (Wilson et al, 1965; Murphy and Mutalik, 1969). In the present study we have shown how the odds or probability of heterozygosity can be determined from the phenylalanine loading test and how these are influenced by the presence and proximity of an affected relative. This method has the advantage that potentially erroneous categoric statements are avoided, and it is unnecessary to calculate limits based on financial or social burdens to the individual or to society (Gold et al, 1973). It seems likely that individuals in different circumstances, such as parents already with or without normal children, will be influenced by very different probabilities of heterozygosity in the decision to have further or any children.

This work was supported by grants from the Research Committee of the Birmingham Regional Hospital Board and the United Birmingham Hospitals Endowment Fund.

REFERENCES
Heterozygote detection in phenylketonuria


