The Cincinnati Lipid Research Clinic Family Study: Familial Determinants of Plasma Uric Acid*

D. C. Rao¹, P. M. Laskarzewski², J. A. Morrison³, P. Khoury⁴, K. Kelly⁵, and C. J. Glueck⁶

¹Division of Biostatistics, Washington University Medical School, Box 8067, 4566 Scott Avenue, St. Louis, MO 63110, USA
²The Lipid Research Clinic, The General Clinical Research Center, The CLINFO Center, University of Cincinnati, College of Medicine, 234 Goodman Street C2-3, Cincinnati, OH 45267, USA

Summary. Commingling analysis of plasma uric acid levels in a random sample of 160 nuclear families supports the hypothesis that there is a mixture of three distributions. Assuming one, two, and three components in the underlying distribution, we obtained the corresponding p-values (for power transformation) as 0.059, 1.040, and 1.643, respectively. Path analysis with p = 0.059 gives genetic (h²) and cultural (c²) heritabilities as 0.256 and 0.199, without much support for intergenerational differences, assortative mating, or maternal effects. Complex segregation analysis with p = 0.059 supports multifactorial inheritance, consistent with the findings of Gulbrandsen et al. (1979) and Morton (1979) in other populations. This study also fails to support a major locus hypothesis, contrary to earlier reports.

Introduction

Age, sex, obesity, hematocrit, and alcohol consumption were shown to have considerable effects on plasma uric acid levels (Gulbrandsen et al. 1979). After suitable covariance adjustments, Gulbrandsen et al. (1979) showed that the residual variation involved a relatively small genetic component (genetic heritability = 0.249 ± 0.041). Further, they failed to obtain consistent evidence for megaphenic effects. Even when based on families ascertained through gout, Morton (1979) failed to obtain evidence for a major locus, contrary to earlier claims of dominant inheritance (Smyth et al. 1948; Stecher et al. 1949). Our aim, in this analysis of 160 white families randomly selected in the Princeton School Lipid Research Clinic's Family Study, is two-fold: to assess familial determinants of plasma uric acid, and to seek additional evidence on the major locus hypothesis.

Materials

The Cincinnati Lipid Research Clinic (LRC) Princeton School District Family Study (1976–1978), (Morrison et al. 1982) was part of the National Heart, Lung, and Blood Institute's multicenter collaborative program designed to assess the familial aggregation of lipids and lipoprotein levels (Heiss et al. 1980). Briefly, the Princeton School District Population Study was a two-stage epidemiologic survey of lipids, lipoproteins, and other coronary heart disease risk factors in a biracial population of school children in grades 1–12 and their parents. Following the first two visits of the Prevalence Study (Morrison et al. 1978), a subgroup of probands was drawn from this larger prevalence population for the Family Study. All first degree relatives and spouses of selected probands were contacted; socio-demographic data, fasting plasma lipids and lipoproteins, uric acid, and clinical chemistry measurements were obtained. Proband for the Family Study included both randomly selected subjects and hyperlipidemic subjects (Kelly et al., to be published; Morrison et al. 1982). Here we analyze data on 160 nuclear families, all white, ascertained through the randomly selected probands. There were a few three-generation families which were split into the component nuclear families, avoiding duplications whenever possible. Relatives of the probands studied include spouses, children, parents, and sibs. Very few adopted relatives and a few half-sibs were also studied. Due to very small sample sizes, the data on adopted relatives and half-sibs are not analyzed here. More details of the population studied can be found in Laskarzewski et al. (to be published), and in Morrison et al. (1982).

Transformation of Data

We first investigated the effects of age, sex, and some social and physiologic variables. Specifically, plasma uric acid level was regressed on sex, age, age², age × sex, age² × sex, age³ × sex, contraceptive usage, hematocrit, special diet usage, and obesity as defined by the Quetelet index (weight/height²) in a stepwise fashion retaining only the significant terms. In fact, both forward and backward stepwise regression gave the same result:

\[ U = f(A, S) + g(Z) + e \]

where

\[ f(A, S) = \text{polynomial in age and sex} \]
\[ g(Z) = \text{linear function of hematocrit, special diet usage, and obesity} \]
\[ e = \text{residual error} \]

Here sex was coded as male = 1 and female = 2; special diet usage was simply recorded as yes = 2 and no = 1; age was taken in years;
hematocrit was expressed in percentage. Weight and height were recorded in pounds and inches respectively. Fitting of the regression model yielded the estimated contributions:
\[
\hat{f}(A, S) = -2.485 + 1.551 \text{ sex } + 0.376 \text{ age } \\
-0.00886 \text{ age}^2 + 0.000610 \text{ age}^3 \\
-0.211 \text{ age}^2 \text{ sex } + 0.00483 \text{ age}^2 \text{ sex} \\
-0.0000318 \text{ age}^3 \text{ sex}
\]
and
\[
\hat{g}(Z) = 0.350 \text{ diet } + 0.046 \text{ hematocrit } + 0.625 \text{ obesity}.
\]
Therefore, \( U - \hat{f}(A, S) \) defines the uric acid adjusted for age and sex effects (but not the effects of the index variables). We then investigated commingling in the distribution of the age-sex adjusted uric acid, using the methodology of Maclean et al. (1976). Skewness in the distribution of the residuals can simulate a major locus unless correctly transformed prior to complex segregation analysis. We, therefore, applied the general theory of Maclean et al. (1976) through the transformation
\[
Y = \left\{ \frac{X}{\theta^2} + 1 \right\} - 1
\]
where \( X \) is the residual defined above. It is postulated that \( Y \) follows either a single normal distribution, or a mixture of two or three normal distributions, each with the same variance. By maximum likelihood estimation, we then find a value of \( \theta \) that transforms the distribution of \( X \) as closely as possible to that of a mixture of up to three normal distributions. Different values of \( \theta \) can be obtained by assuming 1, 2, or 3 components in the distribution of \( Y \). We fitted a mixture of one, two and three distributions. The \( \theta \)-values and the corresponding values of \( -2\ln L + c \) (\( \ln L \) is log-likelihood, and \( c \) is a constant) are presented in Table 1. It is interesting that on mere inspection Morton (1979) preferred \( \theta = 0 \), consistent with our estimate for one distribution (\( \theta = 0.459 \)). Two distributions fit significantly better than one (\( \chi^2 = 14.29, P < 0.001 \)), and three distributions fit significantly better than two (\( \chi^2 = 7.45, P < 0.006 \)), consistent with the analysis of Japanese-Americans (Gulbrandsen et al. 1977). The evidence, therefore, supports the hypothesis that there is a mixture of three distributions, all of which would be unskewed for \( \theta = 1.643 \). A tentative genetic interpretation of this phenotypic analysis is that a codominant major locus contributes to the inheritance of plasma uric acid levels. However, within the realm of statistical inference, only complex segregation analysis under a mixed model can resolve major genes judiciously (presented later). Ideally, these data transformations should be based on random samples of unrelated individuals. Lacking such data, we performed the investigation on the family data (random sample).

**Path Analysis**

In this section we will investigate the multifactorial basis of uric acid levels. For this purpose, we transformed age-sex adjusted uric acid using \( \theta = 0.059 \) corresponding to one distribution, and called it the phenotype \( (P) \). Since family environment is not directly measured, following Rao et al. (1979b) we defined an environmental index as \( I = \hat{g}(Z) \), based on special diet usage, hematocrit, and obesity. The index \( (I) \) is taken as an estimate of the family environment \( (C) \), as shown in our model presented in Fig. 1. The present model is a special case of a more general model (Rao et al. 1979a). The ten parameters of this model are

![Path diagram showing cultural and biologic inheritance of plasma uric acid in nuclear families.](image)

**Table 1.** Commingling analysis of age-sex-adjusted uric acid levels

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>No. parameters estimated</th>
<th>(-2\ln L + c)</th>
<th>(P)</th>
<th>Skewness (\beta_1)</th>
<th>Kurtosis (\beta_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age-sex-adjusted uric acid levels</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One distribution</td>
<td>3</td>
<td>1548.22</td>
<td>0.059</td>
<td>0.000</td>
<td>3.811</td>
</tr>
<tr>
<td>Two distributions</td>
<td>5</td>
<td>1533.93</td>
<td>1.040</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Three distributions</td>
<td>1526.48</td>
<td>1.643</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

![Observed vs. expected correlations](image)

Table 2 presents the number of correlations with distinct expectations. Intergenerational differences are retained: whereas \( h^2 \) and \( c^2 \) are the genetic and cultural heritabilities in children, they are \( h^2c^2 \) and \( c^2 \) in adults. Maternal effects are also included by distinguishing the effects of paternal \((f_p)\) and maternal \((f_m)\) environments on that of their child. The model also incorporates separate indexed environments and indices for each child, in addition to a nontransmitted common sibship environment. There are 15 correlations between pairs of the following six variables: \( f_P = \) phenotype of father, \( f_I = \) index of father, \( f_M = \) phenotype of mother, \( f_I = \) index of mother, \( P_C = \) phenotype of child, and \( I_C = \) index of a child. However, since the expected correlations are identical for \((P, I_F)\) and \((P_M, I_M)\), and also for \((P, I_M)\) and \((P_M, I_P)\), the number of correlations with distinct expectations reduces to 13. Sibships generate three more correlations: between phenotypes of sibs, between indices of sibs, and phenotype of a child with sib's index. This makes the total number of correlations 16. All the 16 expected correlations are presented in Table 3. The corresponding observed correlations \((r)\), estimated by the method of maximum likelihood (Morton et al., to be published), and their sample sizes \((n)\) are also presented in Table 3.