Evaluation of segregation ratio in Wilson's disease

TOMOHIRO SAITO

From the Departments of Environmental Health and Pediatrics, Jichi Medical School, Minamikawachi-machi, Japan.

SUMMARY Two problems relating to segregation analysis for Wilson's disease are discussed and a practical solution is presented. A problem in the ascertainment of families with Wilson's disease is illustrated by comparing segregation ratios calculated by the single selection, complete truncate, and multiple incomplete selection methods. The effect on the segregation ratio of exclusion from the analysis of those sibs who had died of other diseases at a young age is also discussed and a method of adjustment of the number of affected using the data on age at onset is proposed.

The segregation ratio by multiple incomplete selection (Weinberg proband method) after adjustment for those sibs who had died of other diseases was 0.243, consistent with the theoretical value for autosomal recessive inheritance. The segregation ratio calculated by the single selection method tended to give a lower value, while that calculated by the complete truncate method was greater than the theoretical value. Recessive inheritance is, however, supported. The actual effect of exclusion of those sibs who had died of other diseases on gene frequency estimation is shown to be very small.

Wilson's disease (hepatolenticular degeneration) is a genetic disorder of copper metabolism which presents mainly with hepatic and neurological symptoms. The disease is one of the most common genetic disorders in Japan, measured by the use of Dahlberg's formula which estimates the gene frequency in the population from the frequency of consanguinity. The incidence was estimated to be from 1 in 20,000 to 1 in 30,000 in Japan.

An autosomal recessive mode of inheritance has been suggested by segregation analysis in various small studies in which different analytical methods were used. However, the method of analysis should be determined by the manner of ascertainment of sibships and this question was not thoroughly discussed in the above studies. In Wilson's disease this question of ascertainment needs particular attention. The presence of substantial numbers of missed cases who had died, mainly of hepatic failure without proper diagnosis being made before the index cases were diagnosed, clearly indicates this problem. The knowledge of a sib's death caused by hepatic failure or with neurological symptoms would increase the probability of diagnosing Wilson's disease in a subsequent child. This suggests that the diagnosis of this disease would be made more frequently in a family with such sibs or already diagnosed cases and less frequently in a family where only one affected child had died of acute hepatic failure within a short time. Thus, families with multiple cases would be over-represented in a sample. In this sense, the ascertainment of pedigrees of Wilson's disease through medical institutions does not seem to suit complete truncate selection nor single selection. The first purpose of this report is to examine this phenomenon by comparing segregation ratios calculated by three different methods based on a large series of patients collected in Japan.

Another problem frequently encountered in segregation analysis is that of dealing with dead sibs. In Wilson's disease, sibs who have died of hepatic failure or with neurological symptoms without a definite diagnosis cannot be excluded as they are most likely cases of the disorder. The number of these is quite large and their exclusion would greatly distort the segregation ratio. Treatment of sibs in the analysis who died of other diseases is another question. The exclusion altogether of these sibs, as in some studies would falsely raise the segregation ratio. Exclusion of these pedigrees rather than the exclusion of sibs from the analysis is regrettable in such a rare disease and would considerably lower the number of available pedigrees. The second purpose of this report is to illustrate this problem and present a practical solution using available data, namely, age at onset of the disease.
Source of data

All reports published between 1965 and 1977 citing cases of Wilson’s disease were traced through Igaku-Chuo-Zasshi (Japana Centra Revuo Medicina) which records short summaries of all articles in medical journals published in Japan. Case presentations of the disease at medical meetings were also included. Fifty-three departments of university hospitals and 44 non-university hospitals reported cases. A standard questionnaire was then sent to each hospital involved in the original collection of data requesting further details of each patient and family members. For cases whose questionnaires were not returned, data were filled in by the author from the original articles and summaries of case presentations.

Such details were also requested of the remaining 154 departments of medicine, paediatrics, and psychiatry at university hospitals from which cases had not been reported in medical journals or at medical meetings. Eighty-nine departments (58%) responded to the questionnaires and 43 departments had cases.

The possibility of duplication of registration was eliminated by a thorough checking of all data. A total of 289 families, each with at least one case, was thus collected.

Methods

Among 289 families collected there were 338 cases in which the diagnosis of Wilson’s disease was confirmed. In addition, there were 62 sibs who had died either of hepatic failure or with neurological symptoms without any diagnosis being made. These 62 sibs were classified as ‘suspected’ cases and treated as cases of Wilson’s disease in the present analysis.

Methods of segregation analysis used in this study were complete truncate method,10 single selection method,10 and the Weinberg proband method11 12 which is incomplete multiple selection.

Among the 289 families, there were 174 families in which information on the following particulars were complete: the number of sibs in the family, age and health of each member of the family including information on the presence or absence of the disease, and where death had occurred, cause and age except in the ‘suspected’ cases. Among them 164 families were not single sibship families and were suitable for segregation analysis. After the exclusion of pedigrees with sibs who had died of other diseases, the total number of pedigrees remaining was 135. There were 43 such sibs among the 164 families and causes of death included malnutrition, pneumonia, enterocolitis, measles, renal disease, drowning, and other accidents. These 164 pedigrees were used in the analysis.

The figure shows the probability of a sib of known age who had died of a disease other than Wilson’s disease being a recessive homozygote for Wilson’s disease. The probability was calculated in the following way. Probability (a sib who had died of another disease being a recessive homozygote for Wilson’s disease at age x) = 0.25 [1-(cumulative percentage of age at onset at age x)]. As the figure shows, all children under 4 years of age have a probability of 25%, which is the theoretical segregation ratio under autosomal recessive inheritance. This is because no patient developed the disease before 4 years of age. As the age increases, the probability decreases. The probability after 31 is zero because all patients developed the disease before that age.

Based on this figure, each of the 43 sibs who were identified as having died of a disease other than Wilson’s disease were given a probability of their being a recessive homozygote according to the age of death. Adjustment was made in this way concerning the number of the affected in the pedigrees in the segregation analysis.

The adjustment was also made in the estimation of gene frequency in the population. The gene frequency in the population q was estimated by the extended Dahlberg formula13

\[ q = \frac{c/16-kF}{(k-c)+(c/16-kF)} \]

where q is the gene frequency in the population to be estimated, c is the proportion of first cousin marriages in the population, F is the coefficient of inbreeding, and k is the proportion of recessive homozygotes from first cousin marriages to those from all marriages from the data. The genotype.
TABLE 1 Segregation ratios by different methods in different sibship groups.

<table>
<thead>
<tr>
<th>Sibship group</th>
<th>No of sibships</th>
<th>Complete truncate</th>
<th>Single selection</th>
<th>Weinberg proband</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Sibships without</td>
<td>135</td>
<td>0.358</td>
<td>0.241</td>
<td>0.258</td>
</tr>
<tr>
<td>dead sibs</td>
<td>(0.029)</td>
<td>(0.024)</td>
<td>(0.026)</td>
<td></td>
</tr>
<tr>
<td>B Sibships with dead</td>
<td>164</td>
<td>0.321</td>
<td>0.214</td>
<td>0.227</td>
</tr>
<tr>
<td>sibs as unaffected</td>
<td>(0.024)</td>
<td>(0.020)</td>
<td>(0.021)</td>
<td></td>
</tr>
<tr>
<td>C Sibships with dead</td>
<td>162</td>
<td>0.355</td>
<td>0.241</td>
<td>0.251</td>
</tr>
<tr>
<td>sibs excluded</td>
<td>(0.026)</td>
<td>(0.021)</td>
<td>(0.024)</td>
<td></td>
</tr>
<tr>
<td>D Sibships with dead</td>
<td>164</td>
<td>0.341</td>
<td>0.231</td>
<td>0.243</td>
</tr>
<tr>
<td>sibs adjusted</td>
<td>(0.024)</td>
<td>(0.020)</td>
<td>(0.021)</td>
<td></td>
</tr>
</tbody>
</table>

Standard errors are in parentheses.
*Two sibships were excluded because they became a one sib family by exclusion of a dead sib.

frequencies were estimated by $Fq + (1-F)q^2$ for recessive homozygotes and $(1-F)2pq$ for heterozygotes, where $p = 1-q$. The number of families used for this analysis was 162.

Results

Results of the segregation analysis by the three different methods are shown in table 1. The figures are segregation ratios and those in parentheses are their standard errors.

The sibs in group A were those from families in which all the sibs except the cases of Wilson's disease survived; sibships with any dead sib were excluded. The sibships in group B included sibs who were known to have died of other diseases and these sibs were treated as non-affected sibs, that is, as non-homozygotes carrying the genes. In group C sibs who had died of other diseases were excluded as they would have had they not been born, as done in previous studies.2 5 6 In group D the same dead sibs were given a probability of their being a recessive homozygote according to their age of death, as described previously.

In all the groups, ratios by complete truncate method gave the highest values, those by single selection method gave the lowest values, and those by the Weinberg proband method gave values between the two, as is theoretically expected. All the ratios by complete truncate method were far above the theoretical value of 0.25 expected for autosomal recessive inheritance and they all showed statistically significant differences from 0.25 at the significance level of 0.05. All other ratios were not significantly different from 0.25. The Weinberg proband method gave the closest value in all the sibship groups.

The ratios in group B were the lowest among the groups and this was the result of the treatment of dead sibs as unaffected; the ratios were, therefore, underestimated. Adjustment of the number of sibs, by giving sibs who had died from other diseases a probability of their being a recessive homozygote according to their age of death in group D, raised the ratio which became closer to 0.25 by the Weinberg proband and single selection methods. The differences between groups A and D were not very large.

The criterion for acceptance of autosomal recessive inheritance proposed by Haldane10 was $p_c + 2c < 0.25 < p_c - 2s$, where $p_c$ is the ratio estimated by complete truncate method, $p_c$ is the ratio estimated by single selection, $c$ is the standard error of the ratio by complete truncate method, and $s$ is the standard error of the ratio by single selection. In all the groups these criteria were met and the values in the formula for group D were 0.390 < 0.25 < 0.101.

The effect of adjustment for those sibs who had died of other diseases on gene frequency and genotype frequency estimation is shown in table 2. When adjustment for them was not made, and if they were treated as unaffected, the term $k$ in the formula was 0.3237: if the adjustment was made it was 0.3245. As seen in table 2, the effect of the adjustment on gene and genotype frequency was very small.

Discussion

A statistical evaluation taking into account those sibs who had died of diseases other than Wilson's

Table 2 Effect of adjustment for sibs who had died of other diseases on gene frequency estimation by the extended Dahlber formula:

$$q = \frac{c(16-kF)}{(k-c)+(c(16-kF))}.$$ 

<table>
<thead>
<tr>
<th>$k$</th>
<th>$c$</th>
<th>$F$</th>
<th>$q$</th>
<th>Wilson $(1-F)^2 + Fq$</th>
<th>Wilson $F$ in</th>
<th>Carrier $(1-F)2pq$</th>
<th>Carrier $F$ in</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dead sibs as unaffected</td>
<td>0.3237</td>
<td>0.04</td>
<td>0.0031</td>
<td>5.22 x 10^{-5}</td>
<td>4.38 x 10^{-5}</td>
<td>22 900</td>
<td>1.04 x 10^{-2}</td>
</tr>
<tr>
<td>Dead sibs adjusted</td>
<td>0.3245</td>
<td>0.04</td>
<td>0.0031</td>
<td>5.22</td>
<td>4.33</td>
<td>22 300</td>
<td>1.04</td>
</tr>
</tbody>
</table>

$q = $ the gene frequency in the population.
$p = 1-q$.
$c$ = coefficient of inbreeding.
$k$ = proportion of first cousin marriages in the population.
$s$ = proportion of affected children from first cousin marriages to those from all marriages.
disease at young ages has never before been made in the genetic analysis of Wilson's disease. This question is very important where mortality, particularly infant mortality, is relatively high. Exclusion of those sibs would greatly raise the segregation ratio and exclusion of the families with such sibs would lower the number of families available for analysis. Although the method presented here is an attractive one, one point needs to be noted. The probability is given to each sib assuming that the disease is a priori autosomal recessive. This method would only be justified if autosomal recessive inheritance had been suggested by previous studies, or if the segregation ratio calculated in this way and those calculated in other ways, such as in groups A, B, and C in this study, did not differ greatly, and other modes of inheritance were not strongly suggested. If the disease were not autosomal recessive it would introduce a bias towards autosomal recessive inheritance. This bias becomes large as the number of dead sibs becomes large. In this study, the observed number of affected was 258 and the affected number was corrected to 265.4, where there were 43 such sibs among the total of 604 sibs. However, if the segregation ratio excluding those sibs is not far above 0.25 it serves to correct the ratio towards the true value. When autosomal dominant transmission is suspected this model must not be used. Instead, probabilities of twice the value, namely 0.50 at the age of 0 and so on, would be justified.

'Suspected' cases who had died of acute hepatic failure or with other symptoms suggestive of Wilson's disease were treated as cases of Wilson's disease in this analysis and also in other studies. Exclusion of these cases would greatly lower the segregation ratio, as there were 50 such sibs among the 164 families in the present study, 10 such sibs among 18 families in the study of Arima et al, and five such sibs among 28 families in the study of Cox et al. If they were not treated as cases, the ratio would be much lower, for example, less than 0.15 by the single selection method in this study, and the autosomal recessive mode of inheritance would not hold.

The results of the analysis supported autosomal recessive inheritance judged by the criteria proposed by Haldane. However, it was revealed that the different methods of segregation analysis gave quite different segregation ratios. The complete truncate method gave fairly high values and the single selection method gave lower values. These tendencies are naturally compatible with previous studies. The values by single selection method in the studies of Arima et al and Fukuda were 0.246 and 0.244, respectively, which were higher than the present one. This might be the result of their exclusion of sibs who had died of other diseases, hence the raising of the values. The number of pedigrees in their studies was rather small and the standard errors of the ratios were approximately twice the one in this study. This might be another reason.

The segregation ratio closest to the theoretical value of 0.25 was the one found by using the Weinberg proband method in group C, that is, the group using all sibships and excluding all sibs who had died of other diseases. Theoretically this certainly does not give the best estimate of the ratio and the result was affected by the ascertainment problem. Since ascertainment probability, \( p \), varies from 0 to 1 (when \( p = 0 \) it is single ascertainment and when \( p = 1 \) it is complete truncate ascertainment) and its effect on the segregation ratio is great, its effect surpassed the bias introduced by excluding the sibs who had died of other diseases. Therefore, in this study, the value of 0.243 by the Weinberg proband method in group D would be more pertinent than 0.251 in group C. This ascertainment problem is a major difficulty in Wilson's disease. In addition to the previous studies mentioned, in the report by Walse, on 92 families, 16 families had one or more dead children whose diagnosis was made only at necropsy or retrospectively. The possibility of over-representation of multiple case families in a sample of hospital based series needs always to be considered. The only solution to this problem would be identification of cases by mass screening.

A complex method using ascertainment probability in a more explicit way proposed by Morton and Haldane was not used in this study, the reason being the considerable difficulty in estimating the ascertainment probability in hospital cases in the light of the ascertainment problem already mentioned. A simple ratio of probands to the total number of cases would not properly estimate ascertainment probability, nor would ascertainment probability be uniform among sibships. The Weinberg proband method is thought to give a satisfactory substitute in most cases. The purpose of this paper was to illustrate the problem of ascertainment in Wilson's disease by comparing the ratios by the three conventional methods rather than to try to find out the best parameters of those methods.

There was only a small effect of adjustment on the estimation of gene frequency. This was expected because it only affects \( k \) in the formula and there should not be a substantial difference in the number of sibs who had died of other diseases in consanguineous families compared to non-consanguineous families. There were slightly more such deaths among consanguineous families in
Evaluation of segregation ratio in Wilson’s disease

this study. The numbers added to correct the number of the affected were 2.29 for 78 sibs from first cousin marriages and 4.15 for 241 other sibs, thus only raising k from 0.3237 to 0.3245. The term k is related only to first cousin marriages. For reference, the proportion of patients from all consanguineous marriages in this study was 39%. If the incidence of the disease in Japan during the study period was around 1 in 20,000,3 there must have been 70 to 90 new patients every year, which amounts to from 1000 to 1400 cases for the 15 years. In this study there were only 338 cases and 62 suspected cases. Therefore, roughly one-third were thought to be collected. The numbers used in the genetic analysis were 258 cases and 43 suspected cases. The study was not complete in terms of ascertainment of cases, but it is the biggest so far and would represent the true situation better than previous studies. The best possible way of studying the genetics of Wilson’s disease would be through mass screening.

I thank Dr K Nomiyama and Dr M Hayashi for encouragement and support, Dr J A H Waterhouse and Dr S Bundey for helpful suggestions, Miss R A Alcock for her help in the preparation of the manuscript, and most of all the attending doctors, too numerous to mention, for their co-operation in providing data for the study. Part of the study was carried out during the author’s tenure of a British Council Scholarship in the Department of Social Medicine, University of Birmingham, and part was supported by a grant from the Ministry of Education in Japan (Grant No 56770316).

References


Requests for reprints to Dr T Saito, Department of Environmental Health, Jichi Medical School, 3311-1 Yakushiji, Minamikawachi-machi, Kawauchi-gun, Tochigi-ken, Japan 329-04.