American Journal of Medical Genetics 34:480–486 (1989)

Inferring Mode of Inheritance by Comparison of Lod Scores

David A. Greenberg
Mount Sinai Medical Center, Department of Psychiatry, New York

One usually must assume a mode of inheritance when using lod scores for linkage analysis. In this study, we asked the question, "If one assumed mode of inheritance in a linkage analysis gives a higher lod score than another, does that indicate that the mode of inheritance that led to the higher lod score is more 'correct' than the other?"

We simulated data under a variety of penetrances, assuming either dominant or recessive inheritance. We then analyzed those simulated data under the correct mode of inheritance, assuming a range of penetrance values, and under the incorrect model, also assuming a range of penetrance values. We found that, if there was enough information for a maximum lod score of at least 3.0, assuming the correct penetrance value or mode of inheritance in the analysis led to a higher lod score than assuming the incorrect penetrance or the incorrect mode of inheritance. These results cannot yet be generalized outside of the specific modes of inheritance and penetrance combinations that we have modeled. Also, penetrance was modeled as "random." The effect of "reduced penetrance" caused by other genetic factors has not yet been tested.

We also tested the effect of non-standard ascertainment on drawing conclusions about mode of inheritance from linkage data. Even when families were ascertained only if the family was *multiplex* (i.e., more than one affected sib), assuming the correct mode of inheritance gave a higher lod score than assuming the incorrect mode of inheritance.

This method has the promise of both simplifying and expanding the application of linkage analysis. The explorations reported here are, however, preliminary, and the method should be used with caution.

KEY WORDS: linkage analysis, segregation analysis, reduced penetrance, simulation, genetics

INTRODUCTION

Recently, we have attempted to examine what effect the origin of reduced penetrance has on linkage analysis [Greenberg and Hodge, 1989]. We asked the question: Does it make any difference to the linkage analysis whether the "reduced penetrance" is the result of random environmental factors or strictly genetic factors? We found that, for the most part, linkage analysis will give approximately correct results, in terms of estimates of θ, the recombination fraction and Z, the maximum lod score, whatever the cause of the so-called reduced penetrance. We also showed that the results of even a correct segregation analysis cannot always be unambiguously applied in a linkage analysis. For example, imagine that it is possible to show unambiguously that a trait is the result of 2 loci, one dominantly and one recessively inherited. If analysis is limited to single-locus linkage programs, which mode of inheritance should one assume for the trait in a linkage analysis? Both dominant and recessive? How can one interpret the results? In fact, linkage analysis is often done with little or no knowledge of mode of inheritance.

In this work, we extend the previous study and attempt to explore the question of how the results of linkage analysis are influenced by reduced penetrance and by mode of inheritance. These questions are usually intractable when approached by analytical methods. When analytic solutions are difficult or impossible, simulation offers a reliable way to try and obtain insight into the effect that the assumed penetrance and assumed mode of inheritance have on linkage analysis. We simulate data under a variety of penetrances and different modes of inheritance and then analyze those simulated data under different assumptions of penetrance and mode of inheritance in an attempt to observe

Received for publication September 20, 1988; revision received May 5, 1989.

Address reprint requests to David A. Greenberg, Ph.D., Department of Psychiatry, Box 1230, Mount Sinai Medical Center, One Gustave L. Levy Place, New York, NY 10029.

whether the results of linkage analysis yield clues into the actual mode of inheritance and penetrance. We come to the conclusion that comparison of lod scores can yield information about the mode of inheritance, information that appears to be robust with respect to ambiguities in ascertainment.

METHODS

The study involved the following steps:

(1) Simulate data with a given mode of inheritance (either a single-locus recessive or a single-locus dominant) and with a specified penetrance.
(2) Analyze those data under the same (correct) mode of inheritance, but use a range of assumed (analysis) penetrances. For each assumed penetrance value, determine the maximum lod score at the grid of values $\theta = .01$, $\theta = .05$, $\theta = .1$, $\theta = .2$, $\theta = .3$, and $\theta = .4$.
(3) Analyze the data under the *incorrect* mode of inheritance (i.e., *not* the mode of inheritance used to generate them), also assuming a range of penetrances.

We modified our well-tested simulation program [Greenberg, 1984] to generate linkage data from nuclear families. Each family consisted of both parents and 2 or more offspring. Sibship sizes were determined according to a well-specified distribution (mean = 2.8, $\sigma = 2.3$) [Cavalli-Sforza and Bodmer, 1971]. Ascertainment was always through the offspring and followed the classical model of Weinberg [1912] with the ascertainment probability $\pi = 0.05$. Thus, many families had only one affected child. The simulated data were such that all families were fully informative at the marker locus, that is, all mating types were AB × CD.

We confined the simulation and analysis to dominant and recessive inheritance. The computational time involved for the generalized single-locus model (one penetrance for each possible genotype) is much greater, as is the complexity in interpreting the results. We are currently investigating the more generalized case.

Data were generated under the assumption of a single-locus dominant or a single-locus recessive model with some fixed penetrance. Each genetically affected person was determined to be penetrant or not on the basis of a random number generator. Data were analyzed using the program LIPED [Ott, 1975], which was used without modification. Using LIPED as a "black box" allowed us to thoroughly check that the simulation was generating data as we specified and reduced the possibility of programming error.

For each model, we generated data on 600 families and analyzed these families in groups of 20—each group was considered a "data set." We felt that a data set size of 20 where all matings were informative was a fair representation of the amount of information in many published linkage data sets. The results for each of these data sets were then used to calculate the means and standard deviations for the maximum lod scores.

The figure of 30 data sets per simulation run represented a compromise between the computer time required to analyze the data, the disk space available, and the technical limitations of the operating system. All computations were done on a Microvax II with a floating-point unit under the ULTRIX operating system.

ANALYSES

Relationship Between *Actual* Penetrance and *Assumed* Penetrance

All recessive data were generated under a gene frequency of 0.04 and all dominant data were generated under a gene frequency of 0.01. The mating types at the marker locus were always AB × CD. For each set of parameters used to generate the data (i.e., mode of inheritance, penetrance, recombination fraction), we examined the results to see whether there was a relationship between the generating penetrance and that assumed penetrance that yielded the highest maximum lod score. For example, if we generated the data under the assumption of 60% penetrance, would a penetrance of 60% used in the *analysis* give us a higher maximum lod score than an assumed penetrance of 0.5 or 0.7? We analyzed the data 10 times, each time assuming a penetrance of either 0.1, 0.2, 0.3 . . . up to a penetrance of 1.0. The gene frequency was maintained at its correct (i.e., generating) value.

Relationship Between the Assumed Penetrance and the Mode of Inheritance

In some cases, we used the incorrect as well as the correct mode of inheritance for the analysis at each assumed penetrance. When a dominant model was assumed for the analyses, a gene frequency of 0.01 was assumed, and when the analysis model was recessive, the gene frequency was 0.04. (The actual (generating) gene frequencies were 0.04 for recessive and 0.01 for dominant.)

As discussed above, the 30 data sets were analyzed assuming different penetrance values. We plotted the mean maximum lod score values over the 30 data sets against the assumed penetrance value to produce a plot that, to avoid having to refer to the cumbersome "highest maximum lod scores," we refer to below as the LVP (maximum-lod-score vs. penetrance) curve. When we refer to the highest or maximum LVP value, we mean the point where the *highest* maximum lod score value occurred on a given LVP curve.

RESULTS

Our results can be summarized as follows:

(1) The mean maximum lod score (over the 30 data sets), as a function of the assumed penetrance, occurred at or near the value of the assumed penetrance that matched the generating penetrance, although the LVP curve was usually fairly flat near the maximum.
(2) When comparing the maximum lod score vs. assumed penetrance curves for the incorrect and correct modes of inheritance, we observed that the curve with the highest maximum lod score was always produced by the generating (correct) mode of inheritance *if* that maximum lod score was above about 3.0.

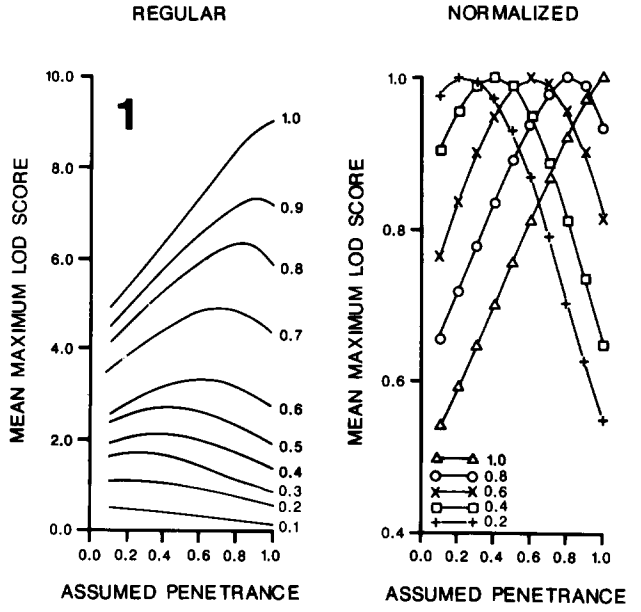

Fig. 1. Plot of the mean maximum lod score vs. the assumed penetrance for data generated under different assumed penetrances for the recessive model. The data generated under *each* generating penetrance were analyzed under 10 different values of the assumed penetrance, those being 0.1 through 1.0 in steps of 0.1. **Graph on the left:** The results plotted on a lod score scale. The number at the end of each curve is the value of the *generating* penetrance. **Graph on the right:** The same data as the left, but normalized to be on a scale from 0 to 1 to show exactly where the maxima in the plots occur. For the sake of readability, only the even-numbered penetrances are shown.

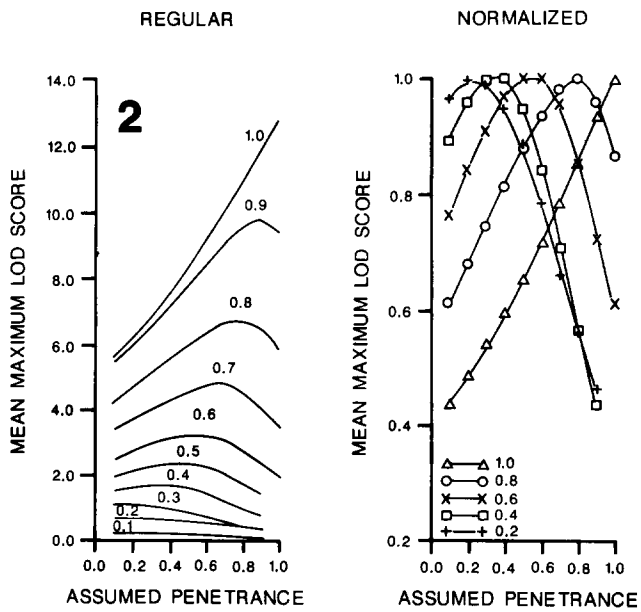

Fig. 2. Same as Figure 1, but for the dominant model.

Figures 1 and 2 show the curves for the mean maximum lod score vs. assumed penetrance for the recessive and dominant models, respectively. These data were generated at $\Theta = 0.01$. Each point on each curve represents the mean maximum lod score for 30 twenty-family data sets. The label on each curve (i.e., 0.1, 0.2, 0.3, . . . 1.0) represents the *generating* penetrance for that curve. The abscissa of each curve is the value of the penetrance assumed for the analysis. The ordinate on Figures 1 and 2, left graph, shows the mean maximum lod score resulting from the analysis of the 30 data sets at the given assumed penetrance. The left graph of Figures 1 (recessive) and 2 (dominant) show the LVP curves all plotted on the same maximum lod score scale. Since plotting all the curves on the same scale obscures the exact location of the maxima for those data generated with the lower penetrances, we normalized the data for each curve by dividing all of the data points for that curve by the highest value for that curve. Thus, the results for these normalized curves are always between zero and one. In Figures 1 and 2, the right graph curves are all plotted on this normalized scale to illustrate where the maxima occur. (Only the even-numbered penetrances have been plotted so as to make the graphs more legible.) Table I gives the means and standard deviations at the maxima for the curves shown in Figures 1 and 2, left graph. Notice that, to a good approximation, the maxima occur at the generating value of the penetrance. The analysis under the dominant model could not distinguish between the different penetrances as consistently as the recessive if the generating penetrance was less than 0.3. Notice also that the graphs for the dominant are occasionally fairly flat around the maxima. (There was also some computational difficulty in assuming a penetrance of 1.0 for the analysis when the data were generated with low penetrances. The simulation would generate impossible families under the

TABLE I. Value of the Lod Score at the Maxima of the LVP Plots for Each Generating Penetrance*

Model	Generating penetrance	Analysis penetrance where highest mean maximum occurred	Mean LVP ± SD
Recessive	0.1	0.1	0.52 ± 0.5
	0.2	0.2	1.11 ± 0.9
	0.3	0.3	1.72 ± 1.0
	0.4	0.4	2.15 ± 1.4
	0.5	0.5	2.74 ± 1.3
	0.6	0.6	3.36 ± 1.4
	0.7	0.7	4.93 ± 2.0
	0.8	0.8	6.33 ± 2.0
	0.9	0.9	7.36 ± 2.6
	1.0	1.0	9.07 ± 2.4
Dominant	0.1	0.2	0.22 ± 0.28
	0.2	0.2	0.70 ± 0.54
	0.3	0.2	1.14 ± 0.74
	0.4	0.4	1.72 ± 0.85
	0.5	0.5	2.37 ± 1.19
	0.6	0.5	3.25 ± 1.21
	0.7	0.7	4.77 ± 1.70
	0.8	0.8	6.74 ± 2.08
	0.9	0.9	9.80 ± 2.46
	1.0	1.0	12.73 ± 2.71

* Values reported in the last column represent the means and standard deviations over 30 data sets of the maximum lod scores. These values are the highest value of the LVP achieved after assuming a range of analysis penetrances for each value of the generating penetrance.

assumption of complete penetrance. Hence, the missing points at an assumed penetrance of 1.0 in Fig. 2.)

Plotting the means conceals the fact that there was a great deal of variation between data sets. Figure 3 shows the plots for four individual data sets where the model was recessive, the generating penetrance was 0.6, and Θ was 0.1. (These 4 data sets were not chosen as representative of most data sets, but rather to show how much variation was observed.)

Having demonstrated that there is an apparent relationship between the actual penetrance and the assumed penetrance that gives the highest LVP, we then tested what effect the assumption of the mode of inheritance has on the relative value of the lod score. Specifically, the question we asked is as follows: If we generate data under a specified mode of inheritance and a specified penetrance, is there *any* assumed (analysis) penetrance that, when used with the *incorrect* mode of inheritance, will give a higher LVP than the correct mode of inheritance?

We generated data under a recessive model with a generating $\Theta = 0.1$ and a penetrance of 0.6 and analyzed them assuming a range of penetrance values under both dominant and recessive inheritance. Figure 4 shows the mean LVP curve for 30 data sets analyzed under both the recessive and dominant models. These results show that, on the average, assuming the true mode of inheritance led to higher lod scores than the incorrect mode of inheritance at every value of the penetrance.

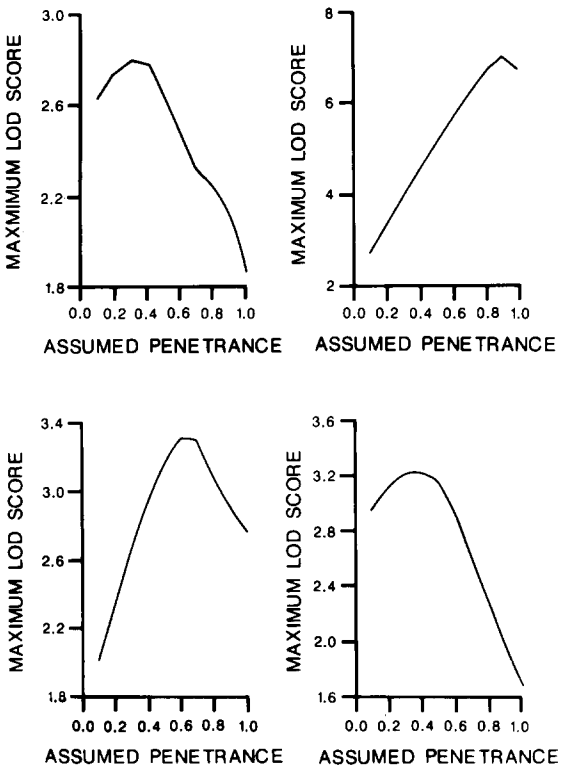

Fig. 3. Four curves illustrating the variation seen in individual data sets (20 families per data set). These examples were generated under a recessive model with 60% penetrance and $\Theta = 0.1$.

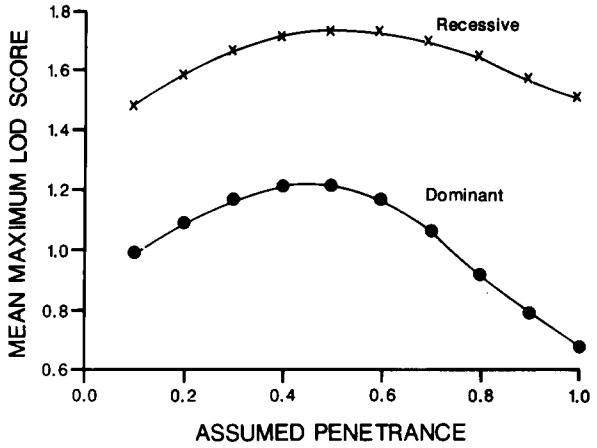

Fig. 4. Plot of data generated under a recessive model with 60% penetrance and $\Theta = 0.1$ and analyzed under the full range of assumed penetrances under both dominant and recessive models. The means are over 30 data sets of 20 families each.

Figure 5 shows the LVP curve for data generated under a dominant model, and analyzed under both dominant and recessive. The data were generated under a dominant with 40% penetrance and Θ equal to either 0.1 or 0.01. Note how much the separation between the curves for the correct and incorrect model increases when the data are generated with Θ equal to 0.01 as opposed to a Θ of 0.1—i.e., when there is more information in the data set. We repeated this analysis, but with a generating penetrance of 0.2 (not shown in the figure). The patterns of results were the same, with the correct assumed mode of inheritance and penetrance resulting in slightly higher mean lod scores.

The *mean* LVP scores for the correct mode of inheritance were higher than for the incorrect mode of inheritance. However, for some individual data sets generated

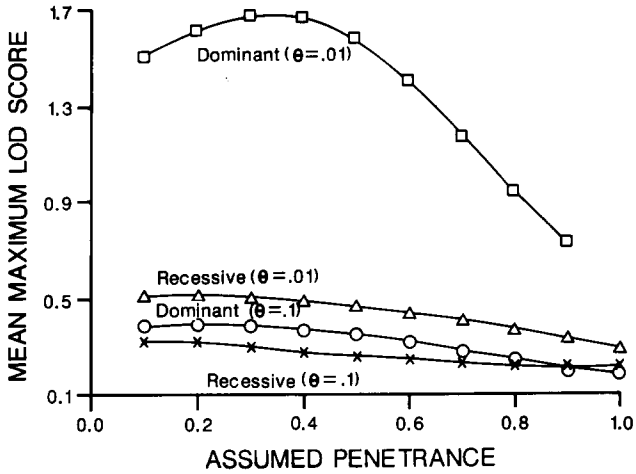

Fig. 5. Like Figure 4, except for the dominant model. Here, the generating penetrance was 40%, and the generating Θ was either 0.1 or 0.01. Note how the difference between the dominant and recessive curves increases when the generating Θ goes from 0.1 to 0.01. Note also how the dominant and recessive curves cross at the assumed penetrance of 0.9 when the generating Θ is 0.1.

under the recessive model, the maximum LVP under the assumption of dominant inheritance was higher than the highest LVP under the correct model. We observed that, out of 30 data sets generated under recessive inheritance with a penetrance of 0.6 and a recombination fraction of 0.1, 11 had a highest LVP under dominant inheritance rather than the correct recessive inheritance. Those 11 data sets had maximum LVP values of less than 3.0.

In contrast to the results under recessive inheritance, when we generated data under dominant inheritance with $\Theta = 0.1$ and analyzed it under the assumption of recessive inheritance, only one data set gave a highest LVP value for the incorrect model that was higher than for the correct dominant model. In that data set also, the highest LVP value was less than 3.0.

We turn our attention now to those individual data sets generated under recessive inheritance but that on analysis gave the highest maximum lod score value assuming dominant inheritance. As noted, the highest LVP value for those data sets was less than 3.0. Therefore, we hypothesized that in data sets with a large amount of evidence for linkage, the maximum LVP value for the correct model would *always* be greater than for the incorrect model.

We tested this idea by increasing the evidence for linkage in the data sets by 2 methods.

First, we generated data using the recessive model and a penetrance of 0.6 but with a recombination fraction of 0.01. We produced the LVP plots under the assumptions of both dominant and recessive inheritance. In contrast to the 11 data sets that showed the LVP for the incorrect dominant being higher than for the correct recessive when Θ was 0.1, only 2 data sets showed such results when Θ was 0.01. Again, for both of those data sets, the highest maximum lod score was less than 3.0 (2.8 and 1.9), and the LVP for both models in those 2 data sets were approximately equal. For most data sets, the difference in the maxima between the dominant and recessive models was approximately 2 lod score units.

Second, we changed the ascertainment method for the data. For the runs thus far discussed, families were ascertained according to the classical model of Weinberg [1912], and the ascertainment probability was set at .05. Ascertainment is usually irrelevant for linkage analysis if one is estimating only recombination fraction and all other model parameters are known. However, we are trying to determine whether a simple comparison of lod scores for different models gives sufficient information to make inferences about one mode of inheritance being more likely than another. We wanted to test the effect of biasing the ascertainment in favor of multiplex families. To test this, we ascertained only those families that had 2 or more affected offspring. In addition to testing whether ascertainment has an impact on the differences in the lod score between models, this experiment also has the effect of enriching the data set for linkage analysis, so that higher lod scores result.

Figure 6 shows the LVP plot for the means over the 30 data sets under the assumptions of dominant and recessive inheritance. Data were generated under a recessive mode with 60% penetrance and $\Theta = 0.01$. The separation between the curves for the recessive and dominant models is about 6 lod score units. The results of these runs showed that *all* data sets produced LVP curves where the correct mode of inheritance (recessive) produced maxima that were higher than the incorrect (dominant) mode of inheritance. In fact, in all 30 data sets, *all* points (not only the maxima) on the LVP curve for the recessive were higher than *all* points for the dominant. The separation between the curves for the individual data sets ranged from 5 lod score units to 10 lod score units.

DISCUSSION

The conclusion one can draw from these simulation experiments is that, at least under the specific circumstances investigated here, one *can* apparently make inferences about the mode of inheritance and penetrance from a simple comparison of lod scores. Furthermore, from our limited results, it appear that those inferences about mode of inheritance may not be subject to ascertainment bias, or not to the same extent as in segregation analysis. What is not clear is how general the validity of these results is. While we observed that the maximum LVP occurred *near* the correct penetrance, it did not necessarily occur *at* the correct penetrance. Also, while we did examine the most often assumed models, that is, dominant and recessive with reduced penetrance, that does not, of course, exhaust the possible models that can be assumed. It does, however, cover those models most frequently assumed in linkage analysis. It is also noteworthy that Clerget-Darpoux et al. [1986] found in a theoretical study that misspecifying the penetrance in general leads to, " . . . a slight underestimate of the lod score." The current study shows that one can estimate the penetrance by maximizing the lod score.

Another important caveat about the above results concerns the origins of the "reduced penetrance" that we assumed. The reduced penetrance used here was *random* reduced penetrance: each person with the disease

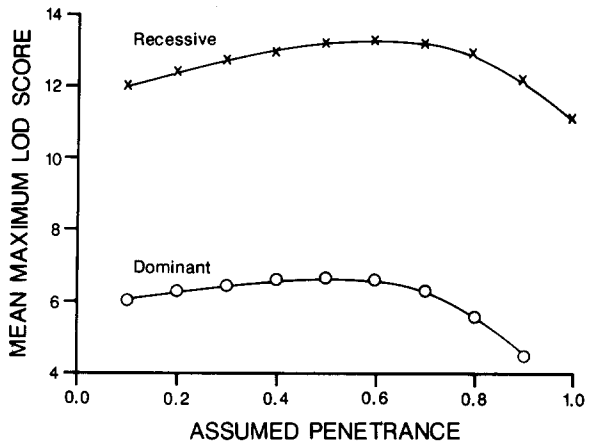

Fig. 6. Like Figure 4 except families were ascertained only if there were two or more affected family members. Data were generated under a recessive with 60% penetrance and $\Theta = 0.01$. Note that the scale on the ordinate is very different from that on Figures 4 and 5.

genotype had the same probability of being phenotypically affected. So-called reduced penetrance can also be the result of genetic factors. In such a case, the probability of a person with the disease genotype at the *linked locus* being phenotypically affected is not random but is influenced by another locus or loci. Preliminary work (Greenberg, in preparation) indicates that when data come from a trait that is the result of the epistatic interaction of two loci, then it appears that so-called "penetrance" (whatever meaning the term might have in such a case) may not be reliably estimated.

It would be of enormous help in genetic analysis if, once a linkage is found, we could easily and simply use that linkage information to get information about the mode of inheritance. The results of the correct work strongly suggest that, at least in cases where the situation is not clouded by questions of heterogeneity, one can draw inferences about the mode of inheritance from the lod scores.

One of our most interesting results concerns the effect of ascertainment on inferring mode of inheritance from linkage results. We tested the case where only those families were ascertained that had 2 or more affected offspring and found that the correct mode of inheritance (in this case, recessive) gave a much higher lod score than the incorrect dominant. If one tried to carry out a standard segregation analysis on such a data set without taking into account the fact that only multiplex families were ascertained, the results would be meaningless. Even more problematic would be a situation in which ascertainment was in some way quadratic [Ewens and Shute, 1986], that is, where the probability of ascertainment increased as the square of the number of affected sibs. In that case, the bias would be even worse. Since ascertainment can be an extremely complex and subtle issue in many cases, unconscious biases in sampling can easily occur [Greenberg, 1986]. It would be extremely useful if we could use linkage analysis to gain information about the mode of inheritance without having to worry about ascertainment.

It has become increasingly apparent that the question of ascertainment bias correction is a major problem in segregation analysis [Greenberg, 1986; Ewens and Shute, 1986]. The problem is especially apparent where common diseases and diseases with ambiguous presentation are concerned or where a medical study seeks out or ascertains patients in a way that is difficult or impossible to express analytically in a likelihood equation [Greenberg, 1986].

Our results also point out the practical consideration that an incorrect segregation analysis could have significant consequences for the linkage analysis. For example, if families in which a disease is segregating are ascertained in a way that is not corrected for properly in the segregation analysis, the results of the segregation analysis could indicate the incorrect mode of inheritance. Our current as well as previous results [Greenberg and Hodge, 1989] show that, under such circumstances, a linkage analysis under the incorrect mode of inheritance may mask the existence of linkage, whereas an analysis under the correct model could give a significant maximum lod score. It is probably true that with enough data, linkage will be confirmed even under the incorrect model. But a great deal more investment of time and resources would be needed for that confirmation. At the same time, one must bear in mind that our results indicate that no conclusions should be drawn about the mode of inheritance until the maximum lod score is *at least* above 3.

It appears that one can also gain some idea of the penetrance, provided that the evidence for linkage is unambiguous. Again, it must be borne in mind that all the simulations done in this work used *random* penetrance to generate the data. In Greenberg and Hodge [1989], data were generated using both random and nonrandom (i.e., genetic) "reduced penetrance."

Of course, it is valid to estimate the penetrance by maximizing the *likelihood* as a function of the penetrance [see, for example, Ott, 1985], and this has been implemented in several computer programs (PAP [Hasstedt et al., 1979], MENDEL [Lange et al., 1987]). However, we maximized the *lod score*. The difference between the lod score and the likelihood is that the log likelihood at $\Theta = 0.5$ is subtracted from the log likelihood at the test Θ to produce the lod score.

We must urge caution in using this method with actual data. As mentioned above, there is at least one specialized circumstance in which one apparently cannot reliably estimate penetrance. There may be other such circumstances, especially as there are a plethora of ways of trying to collect data. Another caveat is that these calculations were done using only *nuclear* families. What the results will be when large pedigrees are used is still unknown. Therefore, investigators must exercise great caution in trying to draw conclusions using this method when using real data. Only the kinds of models and conditions that have been explored here should be used, and even then, results should be interpreted cautiously.

It must be emphasized that we worked with lod scores, not likelihoods. We worked with lod scores rather than likelihoods because of our work on the effect of the origins of reduced penetrance on linkage analysis [Greenberg and Hodge, 1989]. That work first caused us to think that differences in lod scores between different assumed modes of inheritance could indicate the more favorable model.

As noted in the "Methods" section, when we did the analyses, we kept the gene frequency fixed at the correct value and varied the penetrance. Therefore, for all penetrances except the "correct" (or generating) penetrance, the prevalence was incorrect in the analyses. We tested the effect of this on the analyses by fixing the gene frequency at each assumed penetrance used in the analysis such that the *prevalence* was at the correct value for all analyses. We tested this for the recessive model at a penetrance of 0.6. The highest LVP still occurred at the correct penetrance.

Other investigators have suggested doing combined linkage and segregation analysis [Risch, 1984; Hasstedt et al., 1979]. These approaches are theoretically better than the approach that we are exploring here, but they require that we know how the families were ascertained. The problem is that the difficulties and ambiguities in

segregation analysis would not disappear. Wrong assumptions in the segregation analysis could contaminate the linkage analysis. One must also consider the increased number of parameters that must be introduced in such an analysis. Ewens and Shute [1986] have also described an assumption-free method of correcting ascertainment bias. This method is most effective when one has established a linkage. From our results, it would appear that, if one can draw inferences directly from the linkage analysis, there may be no need to be concerned about the ascertainment at all, although this conjecture requires more research.

In this work, we started with the idea of simply exploring the relationship between the actual penetrance and the assumed penetrance in a linkage analysis and were led to test whether one could draw inferences about the mode of inheritance *directly* from comparisons of the maximum lod scores for different models. The method that we are exploring appears crude, yet its simplicity is compelling. Our armamentarium of methods in genetic analysis is not extensive, and both the robustness and limitations of those methods are relatively unexplored. We believe that the apparent relationship between the mode of inheritance and the results of linkage analysis may give us more power to draw inferences about inheritance. It is possible that by first comparing the relative lod scores under different assumed modes of inheritance, one could improve segregation analysis by "feeding back" the linkage analysis results in such a way that the ambiguities of the segregation analysis—arising from possible heterogeneity, questionable ascertainment, or unknown penetrance—could be examined.

It remains to be determined how useful comparing LVPs may be for more complex modes of inheritance, such as an intermediate one. This could be a genetic model where the homozygote is fully penetrant but the heterozygote has a penetrance of only 50%. However, our results are consistent with the expectation that such a model, if it were the correct one, would yield a higher maximum LVP than either dominant or recessive, with the position of the dominant and recessive curve relative to each other depending on the actual value of the penetrance for the heterozygote. We are in the process of examining more complex modes of inheritance. The computational task of testing the intermediate model is much more time-consuming than testing the simple Mendelian models.

ACKNOWLEDGMENTS

My thanks to Drs. Susan Hodge, Warren Ewens, and Michael Boehnke for reading and ruthlessly criticizing the manuscript.

This work supported in part by grants DK31775 and NS21908 from the National Institutes of Health and by the Veterans Administration.

REFERENCES

Cavalli-Sforza LL, Bodmer WF (1971): "The Genetics of Human Populations." San Francisco: W. H. Freeman, pp 310–313.

Clerget-Darpoux F, Bonaiti-Pellie C, Hochez J (1986): Effects of misspecifying genetic parameters in lod score analysis. Biometrics 42:393–399.

Ewens WJ, Shute NCE (1986): A resolution of the ascertainment sampling problem. I. Theory. Theor Popul Biol 30:388–412.

Greenberg DA (1984): Simulation studies of segregation analysis: Application to two-locus models. Am J Hum Genet 36:167–176.

Greenberg DA (1986): The effect of proband designation on segregation analysis. Am J Hum Genet 39:329–339.

Greenberg DA, Hodge SE (1989): Linkage analysis under "random" and "genetic" reduced penetrance. Genet Epidemiol 6:259–264.

Hasstedt SJ, Cartwright PE, Skolnick MH, Bishop DT (1979): PAP: A FORTRAN program for pedigree analysis. Am J Hum Genet 31:135A.

Lange K, Boehnke M, Weeks DE (1987): "Programs for Pedigree Analysis." Los Angeles: Department of Biomathematics, University of California, Los Angeles.

Ott J (1975): A computer program for linkage analysis of general human pedigrees. Am J Hum Genet 28:528–529.

Ott J (1985): "The Analysis of Human Genetic Linkage." Baltimore: Johns Hopkins University Press, pp 132–140.

Risch N (1984): Segregation analysis incorporating linkage markers. I. Single-locus models with application to type I diabetes. Am J Hum Genet 36:363–387.

Weinberg W (1912): Methoden und Fehlerquellen der Untersuchung auf mendilische Zahlen beim Menschen. Arch Rassenbiol 9:165–174.