Am. J. Hum. Genet. 55:834–840, 1994

Using Lod-Score Differences to Determine Mode of Inheritance: A Simple, Robust Method Even in the Presence of Heterogeneity and Reduced Penetrance

David A. Greenberg and Barbara Berger

Department of Psychiatry, Mount Sinai Medical Center, New York

Summary

Determining the mode of inheritance is often difficult under the best of circumstances, but when segregation analysis is used, the problems of ambiguous ascertainment procedures, reduced penetrance, heterogeneity, and misdiagnosis make mode-of-inheritance determinations even more unreliable. The mode of inheritance can also be determined using a linkage-based method (maximized maximum lod score or mod score) and association-based methods, which can overcome many of these problems. In this work, we determined how much information is necessary to reliably determine the mode of inheritance from linkage data when heterogeneity and reduced penetrance are present in the data set. We generated data sets under both dominant and recessive inheritance with reduced penetrance and with varying fractions of linked and unlinked families. We then analyzed those data sets, assuming reduced penetrance, both dominant and recessive inheritance, and no heterogeneity. We investigated the reliability of two methods for determining the mode of inheritance from the linkage data. The first method examined the difference (Δ) between the maximum lod scores calculated under the two mode-of-inheritance assumptions. We found that if Δ was >1.5, then the higher of the two maximum lod scores reflected the correct mode of inheritance with high reliability and that a Δ of 2.5 appeared to practically guarantee a correct mode-of-inheritance inference. Furthermore, this reliability appeared to be virtually independent of α, the fraction of linked families in the data set, although the reliability decreased slightly as α fell below .50. The second method we tested was based on choosing the higher of the two maximum lod scores calculated under the different mode-of-inheritance assumptions. This method became unreliable as α decreased. These results suggest that the mode of inheritance can be inferred from linkage data with high reliability, even in the presence of heterogeneity and reduced penetrance.

Received July 28, 1993; accepted for publication June 13, 1994.

Address for correspondence and reprints: Dr. David A. Greenberg, Box 1229, Department of Psychiatry, Mount Sinai Medical Center, New York, NY 10029

0002-9297/94/5504-0025$02.00

Introduction

Until recently, formal segregation analysis has been the only method for determining the mode of inheritance for a disease. Segregation analysis has been most successful for diseases that are caused by a single locus, with high penetrance and without significant environmental influences. It has been more problematic when applied to the common, complex diseases because the technique has had difficulty coping with the problems of heterogeneity, ambiguous diagnosis, and nonsystematic ascertainment. Even if one does overcome those problems, interpretation of the results of segregation analysis can also be difficult. These problems have led researchers to look to other methods for determining the mode of inheritance.

Both association analysis and linkage analysis can be used to help determine the mode of inheritance of a disease. Thomson and others (Greenberg et al. 1982; Thomson 1983) have shown that the mode of inheritance can be inferred for diseases that are associated with a marker allele. Association methods have the obvious disadvantage that an association must exist in order to be able to use them to determine mode of inheritance. Linkage data can also be used to determine the mode of inheritance (Elston, 1989; Hodge and Elston, in press). This is called the *maximized maximum lod score* (MMLS; Greenberg 1989) or *mod score* (Clerget-Darpoux et al. 1986) method. This linkage-based method appears to have wider application than the association-based methods because, when a linkage is proved to exist, one knows that a disease gene is in the vicinity of the marker—the distance of the disease gene from the marker most likely being θ, the recombination fraction. In contrast, when an allele at a marker locus shows a population association with a disease but does not show linkage, the relationship between the marker locus and the disease is not so straightforward (Greenberg 1993).

However, major questions remain with regard to the linkage-based method. The most important questions are (1) how *much* information (i.e., what magnitude of lod score) is necessary before one can be reasonably confident that the inferred mode of inheritance is correct and (2) how does *heterogeneity* affect inferring the mode of inheritance from linkage data?

Greenberg (1989) showed that for lod scores calculated

under both dominant and recessive inheritance, if a lod score of ≥ 3.0 is reached under either assumption, then the higher lod score would be most likely to point to the correct mode of inheritance. However, those results were obtained from data simulated without heterogeneity. In order for the MMLS method to be truly useful for the common, complex diseases, we must be able to obtain reliable results even in the presence of heterogeneity.

This paper addresses the questions of (1) how much information is needed to determine which of two possible modes of inheritance is the correct one when inferring mode of inheritance by comparing lod scores and (2) how heterogeneity affects such conclusions. We examine a statistic for testing MMLS data under conditions of heterogeneity plus reduced penetrance. This statistic, designated Δ, appears to be robust. We found that when the Δ statistic is >1.5, then the higher of the two maximum lod scores reflects the true mode of inheritance >95% of the time. If Δ is >2.5, the higher lod score almost always reflects the true mode of inheritance in our simulations. Surprisingly, this observation appeared to be virtually independent of the mode of inheritance of the linked and unlinked forms. It also appeared to be almost independent of the fraction of unlinked families in the data set, i.e., the fraction of the families that had disease not caused by the linked locus. This suggests that, if there is sufficient information in the data set, choosing between the two modes of inheritance by comparing the maximum lod scores is a robust approach, even in the presence of heterogeneity and reduced penetrance.

We also examined whether simply choosing the higher of the two lod scores calculated under the two different mode-of-inheritance assumptions would lead to the correct mode-of-inheritance inference. We found that this method, unlike the Δ method, became unreliable as α decreased.

Methods

Parameters

We simulated nuclear families on the basis of a heterogeneity model: affected individuals could have a disease caused by either of two different loci, one of which was linked to the marker (the "linked" form) and the other of which was unlinked to either the marker or the first disease locus (the "unlinked" form). The probability of families having both diseases segregating is relatively small, since nuclear families were selected (Durner et al. 1992). The linked-locus gene frequency was fixed at .05. The gene frequencies at the second locus were varied from .005 to .10. The parameter α refers to the fraction of families in which the disease is caused by the linked form. Hence, α is a function of the gene frequencies of the disease alleles at the two disease loci.

Data sets were generated under the following models: the D + D (dominant at both loci), the R + R (recessive at both loci), and the R + D (recessive at the linked locus and dominant at the unlinked locus) (Durner and Greenberg 1992). The true, or generating, recombination fraction (θ) was .01. All data were generated using a penetrance of .5. They were then all analyzed under the assumptions of both dominant and recessive inheritance, with .5 penetrance and no sporadics.

In a preliminary study, we determined the value of the assumed penetrance at which the MMLS would occur, depending on whether the correct or incorrect mode of inheritance was assumed. We found that when MMLS values were calculated under the assumption of the correct mode of inheritance, the penetrance estimated by maximizing the lod score as a function of penetrance was identical to the generating penetrance (Greenberg 1989). We also maximized the maximum lod score with respect to penetrance under the *incorrect* mode of inheritance. We found that, for the models considered here, the MMLS value was generally found at a slightly lower penetrance than the generating penetrance. For example, if we generated families with a true penetrance of .5, the MMLS value occurred at a penetrance of .5 when analyzed under the correct mode of inheritance, but at a penetrance of .4 when analyzed under the incorrect mode of inheritance. However, for any given analysis model, the lod-score values for "adjacent penetrances" were so close to each other (difference $< .004$) that we simply kept the analysis penetrance fixed at .5. This increased efficiency at almost no cost to accuracy. (By "adjacent penetrances," we mean the assumed penetrance values just below and above the correct, or generating, penetrance value, in increments of .1. In the current case, the generating penetrance was always .5, and the "adjacent" penetrances were .4 and .6.)

For each gene frequency–mode-of-inheritance combination, 20,000 nuclear families were generated. Each family had two parents and a varying number of sibs. In order to minimize computer time, we required at least three affected offspring per family. The number of sibs was generated according to the standard negative binomial probability distribution (mean = 2.8; SD = 2.3) (Cavalli-Sforza and Bodmer 1971).

We used LIPED (Ott 1974) to calculate the lod scores. The data were analyzed under the correct mode of inheritance (i.e., the generating model) and also under the incorrect mode of inheritance. The two sets of lod scores were computed (under the two models—correct and incorrect). Data sets consisted of 20 families, and 1,000 data sets were used for each combination of parameters.

The Lod Score–Difference Test Statistic, Δ

The Δ statistic is the absolute value of the difference between the lod scores in the two analyses (correct and incorrect): $\Delta = |Zmax_{(Recessive)} - Zmax_{(Dominant)}|$. We tested the premise that if Δ were greater than some critical value,

Table 1

Probability of Incorrect Decision

Mode of Inheritance and Heterogeneity (α)	Δ Interval				
	$.0 \leq \Delta < .5$	$.5 \leq \Delta < 1.0$	$1.0 \leq \Delta < 1.5$	$1.5 \leq \Delta < 2.0$	$2.0 \leq \Delta < 2.5$
D + D:[a]					
.91	.409 (44)	.149 (27)	.044 (45)	.031 (64)	0 (75)
.83	.317 (41)	.097 (41)	.050 (60)	.042 (72)	.015 (66)
.71	.310 (153)	.075 (161)	.006 (154)	.027 (150)	.008 (126)
.55	.286 (238)	.108 (176)	.032 (157)	.028 (143)	.022 (92)
.46	.319 (385)	.111 (234)	.124 (137)	.038 (106)	0 (70)
.39	.274 (475)	.121 (224)	.088 (147)	.038 (79)	0 (36)
.34	.328 (600)	.176 (211)	.060 (83)	.059 (51)	.050 (21)
R + R:[b]					
.99	.263 (163)	.015 (192)	0 (233)	0 (174)	0 (115)
.96	.226 (181)	.035 (229)	0 (205)	0 (181)	0 (95)
.86	.265 (302)	.028 (245)	0 (191)	0 (120)	0 (69)
.61	.272 (595)	.033 (207)	.017 (115)	0 (39)	0 (24)
.41	.129 (232)	.012 (329)	.004 (255)	0 (146)	0 (80)
.28	.271 (702)	.089 (168)	.064 (78)	0 (27)	0 (15)
.20	.245 (645)	.153 (222)	.013 (75)	.074 (27)	0 (19)
R + D:[c]					
.86	.342 (295)	.086 (245)	.031 (161)	0 (131)	0 (87)
.71	.204 (426)	.038 (246)	.050 (156)	.027 (75)	.014 (51)
.61	.322 (189)	.126 (166)	.055 (182)	.037 (136)	.037 (106)
.50	.311 (331)	.183 (218)	.070 (171)	.046 (109)	0 (63)
.41	.325 (452)	.163 (252)	.014 (142)	.068 (73)	.057 (35)

NOTE.—Numbers in parentheses are the observations per cell.

[a] Dominant modes of inheritance.

[b] Recessive modes of inheritance.

[c] Recessive at the linked locus and dominant at the unlinked locus.

then the higher of the two maximum lod scores calculated under the different mode-of-inheritance assumptions would reflect the correct mode of inheritance. We plotted the distribution of Δ for each α (heterogeneity) level. We examined Δ in intervals of 0.5. For each interval, we tabulated P, defined as the observed probability of choosing the wrong model: $P = X/Y$, where X = number of data sets in a given Δ interval where the *higher* of the two lod scores resulted from the *incorrect* mode of inheritance, and Y = total number of data sets.

Results

We tabulated how many times one would select the wrong mode of inheritance as a function of Δ (table 1). We found that as Δ increased, the wrong-choice probability decreased, and, in our numerous simulations, a $\Delta > 2.5$ was almost never associated with an incorrect mode-of-inheritance choice. Figures 1 and 2 show the distribution of correct and incorrect mode-of-inheritance determinations for the recessive at $\alpha = .4$ and for the dominant at $\alpha = .46$. The upper part of the bars shows the number of data sets that gave the incorrect mode of inheritance for a given Δ interval, and the lower part shows the number of data sets that gave the correct answer. The numbers in the boxes show how many data sets are represented in each part of the bar. Note how quickly the numbers of wrong mode-of-inheritance inferences fall as Δ increases.

We were surprised to observe that the percent of incorrect mode-of-inheritance inferences appeared to be virtually independent of α (table 1). The probability of an incorrect decision based on the Δ statistic was almost constant as α increased, although one can note that at the lower values of α, the percent of wrong inferences did rise slightly. Furthermore, the results were approximately the same for the D + D and R + D models. The results for the R + R model were the best, with no wrong inferences about mode of inheritance occurring if the Δ values were >2.0.

In order to get a Δ value that would reflect ~5% probability of choosing the wrong mode of inheritance, we plotted the mean P versus Δ, for each lod score interval, the mean being calculated from the P's at all α levels and all modes of inheritance (fig. 3). The Δ value where the mean probability of a wrong choice is .05 is ~1.5. Another way to examine how Δ reflects the wrong-choice probability is to ask what percentage of data sets yield wrong mode-of-inheritance inferences when the Δ statistic is above a cer-

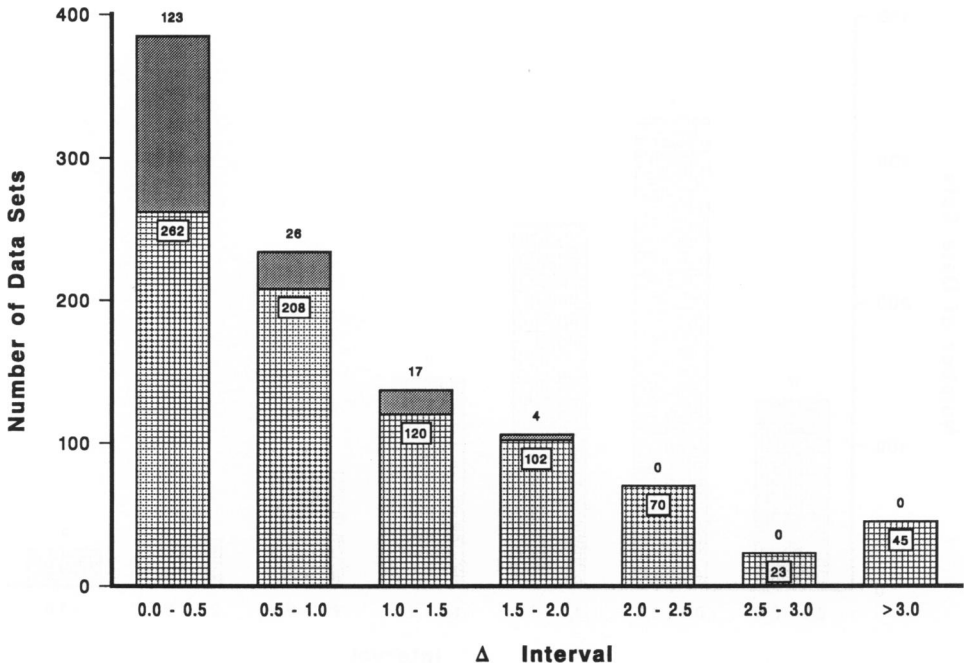

Figure 1 Distribution of correct and incorrect mode-of-inheritance determinations: D + D model with $\alpha = .4$. The height of each bar shows the total number of data sets that had a given Δ value. The upper part of each bar (and the number above it) shows how many data sets yielded the incorrect mode of inheritance when Δ fell within that interval. The lower part of each bar (and the number in the box) shows how many mode-of-inheritance determinations were correct

tain value. Table 2 shows the percent of data sets that yielded a Δ statistic >1.5, >2.0, and >2.5. At most levels of heterogeneity, data sets leading to a $\Delta > 1.5$, or even to $\Delta > 2.5$, are reasonably frequent. The table also shows the percent of wrong mode-of-inheritance inferences that occurred for data sets yielding a Δ above those values. The total probability of getting a wrong mode-of-inheritance inference if Δ is >1.5 is usually $<2\%$, even when most of the families in the data sets are affected with the unlinked form of the disease ($\alpha < .5$).

Discussion

We found that the Δ statistic is robust with respect to α and, so far as we have tested, with respect to the modes of inheritance of the linked and unlinked forms. Furthermore, a Δ of 1.5–2.5 is an attainable goal. Many of the Δ values in the 1.5–2.5 range were calculated from lod scores for the correctly and incorrectly assumed modes of inheritance on the order of 4.0 and 1.0, respectively.

We also tested whether choosing the *higher* of the two maximum lod scores would lead to the correct mode-of-inheritance inference, rather than examining the difference. We hypothesized that, at some critical value, the higher of the two maximum lod scores would reliably point to the correct mode of inheritance. Greenberg (1989) had used this comparison. Instead, we found that the maximum lod score necessary for a correct mode-of-inheritance inference increased as the fraction of unlinked families in the data set increased. When α was $<.6$, the maximum lod score necessary to be 95% certain of choosing the correct mode of inheritance was >3.0 and was 5.0 at $\alpha = .4$. When the linked form was recessive and the unlinked form was dominant, the situation was worse, with the critical lod score value being 6.0 at $\alpha = .6$ and 10.0 at $\alpha = .4$. This finding alone would make comparison of the value of the maximum lod scores a poor choice for determining mode of inheritance. But since the fraction of unlinked families in the data set is usually unknown, one would have to assume the worst and look to lod scores >5 or >6 for any assurance that the correct mode of inheritance yielded the higher lod score.

One interesting aspect of our findings was that, for the Δ statistic, the percent of wrong answers was virtually independent of α and, to a large extent, independent of the modes of inheritance of the different forms of disease. The independence of Δ with respect to α results from the fact that unlinked families contribute no information about the mode of inheritance. Rather, such families appear merely to "dilute" the information in a data set.

In table 2, the percent of data sets yielding the incorrect mode of inheritance tends to rise slightly as heterogeneity increases. We know that the presence of unlinked families does not bias the results of mode-of-inheritance inferences (Durner and Greenberg 1992), so we suspected that since we held the data set size constant, it was the presence of

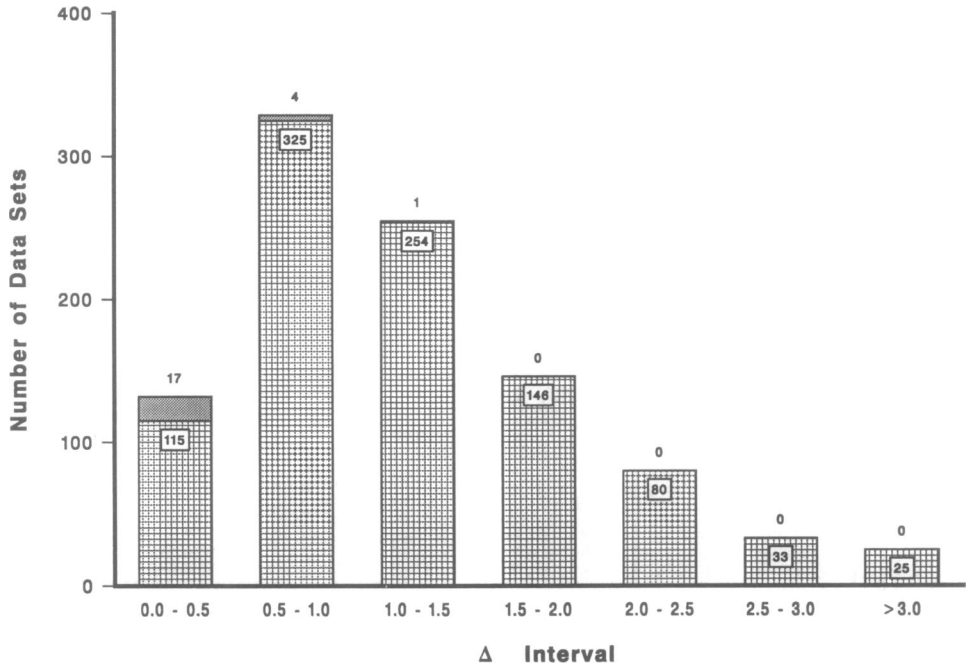

Figure 2 As in figure 1, except that the R + R model with $\alpha = .46$ is shown

less information per data set that was causing the slight increase in wrong inferences.

In order to test this hypothesis, we lowered the information content in a data set by changing the ascertainment scheme. The new scheme required only one affected offspring in order to be ascertained, but maintained the number of families per data set at 20. The generating mode of inheritance was dominant for both linked and unlinked forms; α was .5, and, as in all other experiments, penetrance was .5. We found that ~10% of the data sets had $\Delta > 1.5$, and only 4% of those yielded the incorrect mode-of-inheritance inference. Again, there were no incorrect mode-of-inheritance inferences when $\Delta > 2.5$. These results are comparable to those seen with a much lower α but when families with a higher information content are used, i.e., families ascertained through at least three affected offspring (see table 2). This suggests, as we saw in table 1, that as the amount of information for linkage in a data set decreases, the reliability of Δ decreases slightly. This result also means that reliability will increase as the amount of information in the data set increases, i.e., as more data are collected.

The above results also show that the ascertainment requirements do not appear to affect the results. The families ascertained using a more conventional ascertainment scheme (i.e., families ascertained through at least one affected family member) yielded results quite similar to those for the more stringent ascertainment scheme.

In this work, we have explored only three possible mode-of-inheritance combinations for the linked and unlinked loci. For example, we did not look at the D + R model (dominant at the linked locus and recessive at the unlinked locus). However, on the basis of the results that we have, we would expect that the D + R model would probably perform no worse than did the D + D model, in light of the results for the R + D.

We also looked at the effect of sporadic loci on the results and found that data sets with either R + S (recessive at the linked locus and sporadic at the "unlinked locus") or D + S (dominant at the linked locus and sporadic at the "unlinked locus") appear to give results that were better than for the D + D, R + R, or D + R models. We analyzed these results, assuming a dominant mode of inheritance and assuming a recessive mode of inheritance. The "alternative" mode of inheritance in these cases was not the "correct" one, because the sporadic model is not a mode of inheritance. The fact that a sporadic form of the disease appears to have less impact on mode-of-inheritance determinations than when the unlinked form is dominant or recessive was also noted by Durner and Greenberg (1992). That work showed that if two forms of disease were present in a data set when the unlinked form was sporadic, the effect on the estimate of the recombination fraction θ due to the unlinked form was much less than if the unlinked form were dominant or recessive.

We looked only at the case of dominant and recessive inheritance, because those are the models that investigators generally assume when doing linkage analyses of diseases where the inheritance is unknown. In each of our simulations, one of the two models was the correct, or

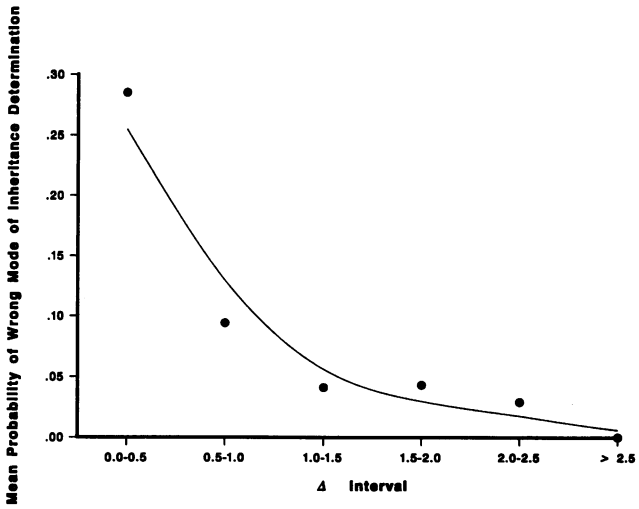

Figure 3 Approximate curve of the mean probability of an incorrect mode-of-inheritance determination over all modes of inheritance and α vs. Δ. The ordinate represents the frequency of incorrect mode-of-inheritance determinations in a given Δ interval.

generating, model. Both the dominant and the recessive are specific examples of a more general genetic mechanism, the generalized single-locus model. Since they are examples of the same underlying model, determining the mode of inheritance by comparing lod scores is theoretically justified (Elston 1989). However, it is unclear whether one could get reliable and consistent results by comparing lod scores from different genetic mechanisms. At least one experiment suggests that it is possible. Greenberg (1990) estimated the penetrance by assuming a single-locus model when the actual generating model was two locus. The penetrance estimates were slightly biased but still reasonably close to what would be theoretically predicted. While it is unclear whether one can compare different genetic mechanisms, it would be valid to compare other examples of the generalized single-locus model. For example, one could compare an intermediate model with a dominant or recessive model or with some other intermediate model. In that case, however, the significance levels determined here might be inaccurate.

Table 2

Percent of Data Sets with Δ above Specific Values

Mode of Inheritance and Heterogeneity (α)	Δ ≥ 1.5	Δ ≥ 2.0	Δ ≥ 2.5
D + D:			
.91	88.4 (.2)	82.0 (.0)	74.5 (.0)
.83	85.8 (.5)	78.6 (.1)	72.0 (.0)
.71	53.2 (.9)	38.2 (.3)	25.6 (.0)
.55	42.9 (1.4)	28.6 (.7)	19.4 (.0)
.39	15.4 (1.9)	7.5 (.0)	3.9 (.0)
.34	10.1 (4.0)	5.1 (2.0)	3.0 (.0)
R + R:			
.99	41.2 (.0)	23.8 (.0)	12.3 (.0)
.96	38.5 (.0)	21.4 (.0)	10.9 (.0)
.86	26.2 (.0)	14.2 (.0)	7.3 (.0)
.61	8.3 (.0)	4.4 (.0)	2.0 (.0)
.28	5.2 (.0)	2.5 (.0)	1.0 (.0)
.20	5.8 (3.4)	3.1 (.0)	1.2 (.0)
R + D:			
.86	30.0 (.0)	16.8 (.0)	8.1 (.0)
.71	16.7 (3.0)	9.6 (1.0)	4.6 (.0)
.61	45.9 (2.0)	32.3 (1.2)	21.9 (.9)
.50	27.5 (1.8)	17.1 (.0)	10.8 (.0)
.41	14.7 (4.8)	7.9 (2.5)	4.6 (.0)

NOTE.—Percents based on 1,000 data sets. Numbers in parentheses are the percent of data sets that gave the wrong mode-of-inheritance inferences.

[a] Abbreviations are as in table 1.

The simulated data sets we used usually had at least three affected offspring per family in data sets of 20 families each. As mentioned above, this sampling strategy was chosen to minimize compute time. The information content of our simulated data sets would be equivalent to actual data sets of 30–40 families of the type usually ascertained in a linkage study (ascertained through one to three affected family members), which is a realistic data-collection goal. The experiment we ran using families ascertained through at least one affected member rather than three (discussed above) shows that the results appear to hold even when families are ascertained with different ascertainment schemes.

The question arises as to what maximum lod-score value, under either model, one should have before applying the Δ statistic. Given that linkage *does* exist, then a Δ of 1.5 appears to be of equal significance whether the original two scores were, say, 1.5 and 0 or 5.0 and 3.5. Certainly, if the latter were the case, then the existence of linkage would be far more certain than if the higher lod score were 1.5, and Δ would have more meaning. In all of our simulations, linkage did exist. We did not investigate the distribution of Δ when there is no linkage.

These data sets were generated under different values of α. We tested how frequently we would be able to detect heterogeneity, by using the admixture test in the 20 family data sets that we simulated. We found that only in ~2%–10% of the data sets, depending on α, would we be able to detect heterogeneity when the correct mode of inheritance was assumed. Presumably, if it were possible to simultaneously detect linkage, heterogeneity, *and* test mode of inheritance, then simply choosing the model that gave the highest lod score might be reliable. However, heterogeneity is difficult to detect. It is important to be able to determine the mode of inheritance, in order to take the next step and determine whether heterogeneity exists.

In summary, the Δ statistic appears to be a robust indicator of the true mode of inheritance, even in the presence of reduced penetrance *and* heterogeneity. While we only

looked at Mendelian models in this work, given the consistency and robustness across different levels of heterogeneity and modes of inheritance, we would expect that similar results would be found when other modes of inheritance are used, at the very least, if the different modes of inheritance being compared were part of the same genetic mechanism. The Δ value of 1.5 gives reasonable assurance of the correct mode of inheritance, and a value of 2.5 appears to practically ensure a correct mode-of-inheritance choice, at least in these simulations. This can be interpreted as the maximum price one pays for not knowing the mode of inheritance ahead of time. However, in light of the difficulty of determining the mode of inheritance for the common diseases and the problem of heterogeneity, to be able to determine the mode of inheritance at all is a major advantage.

Acknowledgments

This work was supported in part by grants DK31775, NS27941, MH45212, and MH48858 and by MacArthur Foundation grant PR1040/92.

References

Cavalli-Sforza LL, Bodmer WF (1971) The genetics of human populations. WH Freeman, San Francisco

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

Durner M, Greenberg DA (1992) Effect of heterogeneity and assumed mode of inheritance on lod scores. Am J Med Genet 42:271–275

Durner M, Greenberg DA, Hodge SE (1992) Inter- and intrafamilial heterogeneity: effective sampling strategies and comparison of analysis methods. Am J Hum Genet 51:859–870

Elston RC (1989) Man bites dog? the validity of maximizing lod scores to determine mode of inheritance. Am J Med Genet 34: 487–488

Greenberg DA (1989) Inferring mode of inheritance by comparison of lod scores. Am J Med Genet 34:480–486

——— (1990) Linkage analysis assuming a single-locus mode of inheritance for traits determined by two loci: inferring mode of inheritance and estimating penetrance. Genet Epidemiol 7: 467–479

——— (1993) Linkage analysis of "necessary" disease loci versus "susceptibility" loci. Am J Hum Genet 52:135–143

Greenberg DA, Hodge SE, Rotter JI (1982) Evidence for recessive and against dominant inheritance at the HLA-"linked" locus in coeliac disease. Am J Hum Genet 34:263–277

Hodge SE, Elston RC. Lods, wrods, and mods: the interpretation of lod scores calculated under different models. Genet Epidemiol (in press)

Ott J (1974) Estimation of the recombination fraction in human pedigrees: efficient computation of the likelihood for human linkage studies. Am J Hum Genet 26:588–597

Thomson G (1983) Investigation of the mode of inheritance of the HLA associated diseases by the method of antigen genotype frequencies among diseased individuals. Tissue Antigens 21:81–104