Wissinger B, Besch D, Baumann B, Fauser S, Christ-Adler M, Jurklies B, Zrenner E, et al (1997) Mutation analysis of the ND6 gene in patients with Lebers hereditary optic neuropathy. Biochem Biophys Res Commun 234:511–515

Address for correspondence and reprints: Dr. Neil Howell, Biology Division 0656, Department of Radiation Oncology, The University of Texas Medical Branch, Galveston TX 77555-0656. E-mail: nhowell@utmb.edu

0002-9297/98/6201-0030$02.00

Am. J. Hum. Genet. 62:202–204, 1998

Power, Mode of Inheritance, and Type I Error in Lod Scores and Affecteds-Only Methods: Reply to Kruglyak

To the Editor:

We had previously written a letter examining some of the issues involved in comparing LOD scores versus affecteds-only and other ''nonparametric'' methods (Greenberg et al. 1996). We had two motivations for that letter. The more important reason was that many of our colleagues have reported difficulties in getting linkage studies funded—or in getting linkage findings published–when LOD scores are used to analyze data. A related impetus for our letter was that there appears to be widespread ignorance of an extensive literature, some of which was cited in our letter, supporting the use of LOD scores. We believe this lack of awareness accounts for the belief of many peer reviewers, of both grant proposals and manuscripts, that LOD scores represent an analysis method inferior to or less powerful than the affecteds-only methods. We tried to address these issues in our letter, because this incorrect belief not only has the negative consequences alluded to above but also runs counter to the practice of good science. We also hoped that our letter would stimulate open discussion of the mathematical issues involved. In this respect we were glad to see a further commentary on our letter, by Kruglyak (1997; also see Farrall [1997] and our response [Greenberg et al. 1997]). However, we feel that it is necessary to focus on some of the points made by Kruglyak.

We respond to the three major points raised by Dr. Kruglyak, which concern (1) the use and meaning of the terms ''nonparametric'' and ''model free''; (2) LOD scores and power; and (3) the role of the true mode of inheritance in LOD scores and in ''model-free'' methods.

1. "*Nonparametric*" *and* "*model free.*"—In his comments, Kruglyak (1997, p. 255) gives a strict statistical definition of ''nonparametric'' or ''model-free'' tests as being those which ''are *valid* [italics his] regardless of the true (unknown) genetic parameters, in the standard sense that they give the correct false-positive rate.'' He then reiterates that this property applies to LOD-score analyses, *under the wrong model* (''wrod'' scores [Hodge and Elston 1994]), just as much as to affected-sib-pair (ASP), affected-pedigree-member, or nonparametric-linkage analyses. The fact that, regardless of whether the assumed model is correct, all of these methods, including LOD scores, satisfy the standard statistical definition of a nonparametric test is apparently not widely understood, although it was formally proved by Williamson and Amos (1990). (Of course, this guarantee of statistical validity holds only for a *single* LOD score or wrod analysis, just as it holds only for a *single* affecteds-only analysis. If an investigator wants to perform two or more linkage analyses, whether LOD score or affecteds-only, allowance must be made for multiple tests. Elsewhere, we have quantified some of this requirement [Hodge et al. 1997].) However, ''nonparametric'' is currently used by most writers to mean ''does not explicitly state a genetic model'' (but see Elston [1997]). This usage is so ingrained that, subsequently in his letter, Kruglyak himself uses ''nonparametric'' in this ''common'' way (Kruglyak 1997). Thus, this is not merely an issue of terminology. It is important because the current usage of ''nonparametric test'' hides the fact that the nominal probability of type I error is asymptotically correct in *all* of the analytic methods under discussion, including LOD scores under the wrong model.

2. *LOD scores and power.*—In his letter, Kruglyak (1997, p. 255) concludes that ''the interesting issue in the design of such [alternative linkage] methods is how to achieve a minimal loss of power while retaining robustness to a maximal range of alternatives.'' We strongly agree. However, he seems to imply—although he does not explicitly state—that, in this respect, LOD scores fare worse than other methods. He says that, when they use LOD scores, researchers who ''guess wrong'' about the genetic model can ''lose big.'' We, too, were concerned about this danger, and that concern provided the impetus for the research cited in our original letter, research that showed that this was not a danger. Kruglyak (1997, p. 255) also says that investigators can ''fish over all possible models and pay the statistical price.'' However, it is not necessary to fish over all possible models (again, the reasoning and citations are in our original letter), and our recent work has shown that comprehensive coverage of models can be had at a modest price in terms of type I error (Hodge et al. 1997).

3. *Role of the true mode of inheritance.*—Here is where the terms used in current parlance—"nonparametric" and "model free"—have proved to be somewhat misleading. Some colleagues with whom we have spoken have concluded incorrectly that all statistical properties of these methods are independent of the

true mode of inheritance of the trait being investigated. On the contrary, although *test size* (probability of type I error, discussed above) of "nonparametric" methods does not depend on the true mode of inheritance (just as it does not do so for LOD scores), their *power* does depend on the true mode of inheritance (just as it does for LOD scores). But the similarity between "model-free" tests and LOD-score analysis goes deeper than that.

Kruglyak (1997, p. 255) points out that, although a "nonparametric" test is equivalent to a LOD-score analysis, "this *does not* mean that a nonparametric test *assumes* [italics his] this choice of parameters." This is, of course, correct, but the fact remains that the two equivalent tests, although not making the same assumptions, can be seen to be statistically identical. In Knapp et al.'s (1994) example, the "mean test" uses ASP data and does not assume a mode of inheritance, whereas a recessive LOD-score analysis of the same data explicitly assumes a fully penetrant recessive model. The mean test has a rejection region *identical* to that of the recessive LOD-score analysis. This means that, if the two tests are set to the same significance levels, they will necessarily also achieve *exactly* the same power, whatever the underlying true model. If the true model is dominant, for example, both analysis methods will suffer the same loss of power, compared with what would have occurred if the true model had been recessive (for some concrete examples of the magnitude of this effect for ASPs, see the report by Hodge [in press]). The important point is that it is misleading to imply that the mean test, or any "nonparametric" test, is somehow "purer" than the corresponding LOD-score analysis, as though its power did not depend on the underlying true model. One can argue about whether the mean test *implicitly* assumes a recessive model or not, but it seems to us that this is a matter of semantics or philosophy.

Whittemore (1996) has shown how this example from Knapp et al. can be generalized. Other "nonparametric" approaches also have statistical properties identical to those of the maximum-likelihood method under specific assumptions. Again, although these methods do not assume a mode of inheritance, one could, in theory, determine which set or sets of mode-of-inheritance assumptions each "nonparametric" analysis method corresponds to and then use one of those sets of mode-of-inheritance assumptions in a LOD-score analysis, to achieve statistically identical results. Thus, using a model-free test can be statistically equivalent to using a LOD-score analysis assuming a wrong genetic model, except that, in the "nonparametric" case, the corresponding genetic models are unknown and are not amenable to adjustment by the investigator.

The existing evidence suggests that the range of genetic models *at a single locus* is robustly spanned by the dominant and recessive models. Extensive work on two-locus multiplicative models (e.g., see Vieland et al. 1992, 1993; Goldin and Weeks 1993; Dizier et al. 1996; Durner et al. 1997) has shown that it is not the mode of inheritance of the disease as a "whole" that needs to be specified in a linkage search; rather, the mode of inheritance at the disease locus *linked to the marker* is critical. (Also, we have now investigated the same issue for "intermediate" and two-locus additive models [Abreu et al. 1997].) If a disease allele exists at all, it must act either alone or together with its sister allele on the homologous chromosome. With relatively little loss of power, effects at other loci can probably be subsumed in other parameters, just as the effect of a second unlinked, epistatic locus can be taken into account of by assuming that there is simple reduced penetrance (Vieland et al. 1992, 1993). If that is indeed the case, then assuming that there is one dominant model and one recessive model will actually cover most possibilities rather well. In this case, "complex" genetic disease can be viewed as being determined by a series of genetic contributions, each of which may be independently detected, and the main issue becomes the relative contribution of a given locus to both the disease and its detectability, not the fact that the disease is "complex."

As we said in our original letter, which approach is "best" will depend on the data available and on other factors particular to the trait and population being studied. We think that it is wrong to condemn a study because the investigators have chosen one analysis method over another, if that method had a reasonable chance of success. What sometimes escapes notice in discussions of analysis methods is that the greater difficulty in studying the genetic contribution to human disease lies in the problems of collecting data good enough for the methods to yield anything at all. Heterogeneity, misdiagnosis, and poorly defined phenotype represent more-serious obstacles to finding disease genes than do the statistical methods available for analysis.

Acknowledgments

We thank Dr. Alice Whittemore for helpful comments. This work was supported in part by National Institutes of Health grants DK31775 and NS27941 (to D.A.G.); MH36197, MH48858, DK31813, MH28274, and MH43878 (to S.E.H.); and MH00884-K21 and MH52841 (to V.J.V.).

David A. Greenberg,[1] Susan E. Hodge,[2] Veronica J. Vieland,[3] and M. Anne Spence[4]

[1]*Departments of Psychiatry and Biomathematics, Mount Sinai Medical Center, and* [2]*New York State Psychiatric Institute and Columbia University, New York;* [3]*Departments of Preventive Medicine & Environmental Health and Psychiatry, University of Iowa College of Medicine, Iowa City; and* [4]*Department of Pediatrics, University of California, Irvine Medical Center, Orange*

References

Abreu PC, Greenberg DA, Hodge SE (1997) Power to detect linkage in complex diseases: maximizing the maximum lod scores (MMLS) is robust in additive and intermediate models. Am J Hum Genet Suppl 61:A264

Dizier M-H, Babron M-C, Clerget-Darpoux F (1996) Conclusions of LOD-score analysis for family data generated under two-locus models. Am J Hum Genet 58:1338-1346

Durner M, Vieland VJ, Greenberg DA (1997) Increased power of lod scores over ASP methods. Am J Hum Genet Suppl 61:A274

Elston RC (1997) Algorithms and inferences: the challenge of multifactorial diseases. Am J Hum Genet 60:255–262

Farrall M (1997) LOD wars: the affected-sib-pair paradigm strikes back! Am J Hum Genet 60:735–737

Goldin LR, Weeks DE (1993) Two-locus models of disease: comparison of likelihood and nonparametric linkage methods. Am J Hum Genet 53:908–915

Greenberg DA, Hodge SE, Vieland VJ, Spence MA (1996) Affecteds-only linkage methods are not a panacea. Am J Hum Genet 58:892–895

——— (1997) Reply to Farrall. Am J Hum Genet 60:738

Hodge SE. Exact ELODs and exact power for affected sib pairs analyzed for linkage under simple right and wrong models. Am J Med Genet (Neuropsychiatr Genet) (in press)

Hodge SE, Abreu PC, Greenberg DA (1997) Magnitude of type I error when single-locus linkage analysis is maximized over models: a simulation study. Am J Hum Genet 60:217-227

Hodge SE, Elston RC (1994) Lods, wrods, and mods: the interpretation of lod scores calculated under different models. Genet Epidemiol 11:329–342

Knapp M, Seuchter SA, Baur MP (1994) Linkage analysis in nuclear families. II. Relationship between affected sib-pair tests and lod score analysis. Hum Hered 44:44–51

Kruglyak L (1997) Nonparametric linkage tests are model free. Am J Hum Genet 61:254–255

Vieland VJ, Greenberg DA, Hodge SE (1993) Adequacy of single-locus approximations for linkage analyses of oligogenic traits: extension to multigenerational pedigree structures. Hum Hered 43:329–336

Vieland VJ, Hodge SE, Greenberg DA (1992) Adequacy of single-locus approximations for linkage analyses of oligogenic traits. Genet Epidemiol 9:45–59

Whittemore AS (1996) Genome scanning for linkage: an overview. Am J Hum Genet 59:704–716

Williamson JA, Amos CI (1990) On the asymptotic behavior of the estimate of the recombination fraction under the null hypothesis of no linkage when the model is misspecified. Genet Epidemiol 7:309–318

Address for correspondence and reprints: Dr. David A. Greenberg, Box 1229, Mount Sinai Medical Center, New York, NY 10029. E-mail: dag@shallot.salad.mssm.edu

0002-9297/98/6201-0031$02.00

Am. J. Hum. Genet. 62:204–205, 1998

Efficient Strategies for Genome Scanning with Affected Sib Pairs

To the Editor:

Holmans and Craddock (1997) present the results of their investigations into the performance of different approaches designed to reduce the number of genotypings required to detect linkage, using a sample of affected sib pairs and their parents; but their method of evaluation is fundamentally flawed, and hence their results do not provide useful information. Two techniques are applied—sample splitting and grid tightening—to produce a two-stage test. In the first stage a reduced number of subjects and/or markers are genotyped, and only regions reaching a certain LOD-score criterion in stage 1 are followed up, in stage 2, by genotyping of all subjects at all markers. Holmans and Craddock present their results in terms of both the power of the procedure to detect linkage and the number of genotypings required. However, in many cases, increasing the number of subjects genotyped in the first stage actually reduces power. Intuitively, it is clear that the detection of linkage rests on being able to identify regions likely to contain linked markers in the first stage and then to follow them up adequately in the second stage. Yet, Holmans and Craddock's results seem to show that a more thorough search in the first stage leads to a decrease in the probability that linkage will be detected, sometimes to a substantial degree (e.g., from .62 to .52 or from .61 to .48). Furthermore, in 3 of their 18 scenarios they recommend a threshold that is higher for the first stage than for the second. This means that one could find a LOD score >3 in the first stage, which one would have to discard and not follow up, even though, if the same LOD score were to be found in the second stage, it would be taken to imply linkage.

The explanation for these paradoxical findings lies with the test strategy that Holmans and Craddock have used. What they propose as a two-stage test for linkage is to choose in advance a LOD score that must be achieved in stage 2 and then to choose as the stage 1 criterion that LOD score that will produce an overall type I error rate of .05. For example, the stage 2 criterion may be set to 3, and then simulations are performed, with the specified data set and scanning procedure, to discover that LOD score that, if used for the stage 1 criterion, will produce a genomewide probability of .05 for an unlinked locus to get through to stage 2 and produce a LOD of 3. As a test for linkage, this is perhaps valid in a narrow sense, but even intuitively it might be expected to perform badly, since it lacks any intrinsic