Letters to the Editor

Am. J. Hum. Genet. 70:556–557, 2002

HLODs Remain Powerful Tools for Detection of Linkage in the Presence of Genetic Heterogeneity

To the Editor:

Recently, Whittemore and Halpern (2001) investigated whether one can meaningfully estimate the admixture parameter α (their p) in the admixture LOD score (i.e., "HLOD") when certain assumptions of HLODs are violated. They argued that such estimates of α are fundamentally problematic, a conclusion with which we agree. However, they then went on to suggest that, in such circumstances, investigators should not use HLODs to *detect* linkage in the presence of suspected heterogeneity (i.e., should not treat α as a "nuisance parameter") but should, instead, use "nonparametric" methods. Unfortunately, they do not cite any evidence to support this final recommendation.

We are writing this letter because, in fact, there is a fair body of evidence, from numerous simulation studies, *supporting* the use of HLODs to detect linkage, even when the assumed heterogeneity model is incorrect. Some of these studies have been published; others are currently in press and were not available to Whittemore and Halpern (2001) when they did their work. We summarize some of these studies.

Several already-published studies have investigated various violations of HLOD model assumptions: Goldin (1992) compared the magnitude of HLODs to the magnitude of two-locus (2L) LODs in data sets in which many families were segregating both the linked and unlinked forms of the disease (these are called "mixed" families). The assumptions made by the 2L analyses exactly matched the reality, whereas the assumptions of the HLODs did not, because the latter assume, incorrectly, that every family has either the linked form or the unlinked form but not both. Goldin found that the HLODs were almost as high as the "correct" 2L LODs. This finding, in conjunction with the work of Abreu et al. (in press; see below), implies, in turn, that the power of the HLOD is also almost as good as the power of analyses performed under the correct model. Durner et al. (1992) made similar comparisons, in data sets with varying proportions of "mixed" families, but under generating models different than those which Goldin had used, and found similar results. Vieland et al. (2001) and Huang and Vieland (2001*a*) considered yet another violation of HLOD assumptions—namely, that proportions of linked families may differ across different data sets. They showed that, for affected-sib-pair (ASP) data, a simple adaptation of the HLOD maintained higher power than did both the homogeneity LOD and certain nonparametric tests (the ASP mean test, as well as Risch's [1990] maximum LOD score under Holmans's [1993] "possible triangle" constraint) in these situations.

We would also like to bring to readers' attention several relevant just-published or not-yet-published studies: Greenberg and Abreu (in press) show that the *multipoint* HLOD has excellent power to detect linkage—and better power than that of the nonparametric NPL statistic of GENEHUNTER (Kruglyak et al. 1996). Their simulations include generating models, such as epistatic and additive models, that violate the assumptions of the HLOD. Vieland and Logue (in press) focus on another way in which the assumptions of the HLOD are commonly violated when the genetic models at the linked and unlinked loci differ. Their work shows, in agreement with that of Whittemore and Halpern (2001), that estimates of α are problematic, but simultaneously it indicates that the maximum HLOD provides a directly interpretable and powerful measure of the strength of evidence for linkage in a data set, despite the problems with α. Abreu et al. (in press) demonstrate that one does not pay much of a price in *type I error* by using HLODs to detect linkage—in most cases, considerably less than even "half" a degree of freedom to the corresponding asymptotic χ^2 statistic; also see the work of Faraway (1993) and Huang and Vieland (2001*b*).

Of necessity, our list of studies is not exhaustive, and our summaries of the papers' findings are oversimplified. For more details, interested readers can consult the papers. Also, we have mentioned only those papers that use HLODs. Another body of published work demonstrates that simple single-locus LOD scores have better power than nonparametric methods do—even *without* inclusion of the admixture parameter (e.g., see Durner and Greenberg 1992; Goldin and Weeks 1993; Abreu et al. 1999; Durner et al. 1999).

One more point: Whittemore and Halpern (2001) also advocate looking for heterogeneity by using "subgroup

analyses"—that is, subdividing the families on the basis of known attributes that might be genetically relevant (e.g., age at onset, ethnic background, etc.) We agree wholeheartedly that this is a valuable approach when one has some idea of *how* to subdivide families. This approach has been dubbed the "predivided-sample," as opposed to the "admixture" approach (Hodge et al. 1983; Ott 1983). But in a situation in which one suspects genetic heterogeneity but in which this heterogeneity does *not* appear to fall along ethnic and other lines, the admixture approach provides an essential tool. Whittemore and Halpern's suggestion of waiting until we have identified the gene before concerning ourselves with heterogeneity does not take into account the effect of heterogeneity on identifying the gene's location in the first place.

In conclusion, this letter should not be misconstrued as attacking Whittemore and Halpern's (2001) findings concerning estimation of the admixture parameter. We agree with Whittemore and Halpern that interpretation of $\hat{\alpha}$ is problematic in many circumstances. However, as we have summarized here, despite the well-known and -recognized problems with parameter estimation, a good deal of evidence indicates that the HLOD can provide a robust and powerful tool for *detection* of linkage in the presence of heterogeneity, even when the assumptions of the HLOD are violated. Whittemore and Halpern (2001, p. 457) stated that "nevertheless, we do not recommend the use of parametric heterogeneity models in linkage analysis, even merely as a tool for increasing the statistical power to detect linkage. … because the assumptions required by these models cannot be verified, and their violation could actually decrease power." Violation of assumptions "could" decrease power, but so far all the evidence is that the power is still greater than that of nonparametric methods. There may well be situations in which such a power loss could be serious, but so far we have not seen these situations, in contrast to numerous demonstrations that simple LOD scores, especially when they incorporate admixture, have good power to detect linkage in the presence of heterogeneity, including situations in which many different assumptions of the model are violated. Whittemore and Halpern's recommendation not to use HLODs to detect linkage in these circumstances may have seemed reasonable in the light of the problems in interpretation of α, but that recommendation is not, in fact, supported by the evidence available so far. Certainly, this topic of robustness of HLODs could use more research. However, given the difficulties in dealing with complex diseases, as well as our need for as many good analytic tools as we can find, it would be a shame if readers of Whittemore and Halpern avoided this particular useful tool, the HLOD, unnecessarily.

Acknowledgments

The authors have received support from National Institutes of Health grants DK-31813, DK-31775, MH-48858, NS-27941, MH-52841, and K02-01432.

SUSAN E. HODGE,[1] VERONICA J. VIELAND,[3] AND DAVID A. GREENBERG[1,2]

[1]*Departments of Psychiatry and Biostatistics (Division of Statistical Genetics), Columbia University, and Division of Clinical-Genetic Epidemiology, New York State Psychiatric Institute, and* [2]*Genome Center, Columbia University, New York; and* [3]*Department of Biostatistics, Division of Statistical Genetics, College of Public Health, and Department of Psychiatry, College of Medicine, University of Iowa, Iowa City*

References

Abreu PC, Greenberg DA, Hodge SE (1999) Direct power comparisons between simple LOD scores and NPL scores for linkage analysis in complex diseases. Am J Hum Genet 65:847–857

Abreu PC, Hodge SE, Greenberg DA. Quantification of type I error probabilities for heterogeneity lod scores. Genet Epidemiol (in press)

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

Durner M, Vieland VJ, Greenberg DA (1999) Further evidence for the increased power of LOD scores compared with nonparametric methods. Am J Hum Genet 64:281–289

Faraway JJ (1993) Distribution of the admixture test for the detection of linkage under heterogeneity. Genet Epidemiol 10:75–83

Goldin LR (1992) Detection of linkage under heterogeneity: comparison of the two-locus vs. admixture models. Genet Epidemiol 9:61–66

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, Abreu P (2001) Determining trait locus position from multipoint analysis: accuracy and power of three different statistics. Genet Epidemiol 21:299–314

Hodge SE, Anderson CE, Neiswanger K, Sparkes RS, Rimoin DL (1983) The search for heterogeneity in insulin-dependent diabetes mellitus (IDDM): linkage studies, two-locus models, and genetic heterogeneity. Am J Hum Genet 35:1139–1155

Holmans P (1993) Asymptotic properties of affected-sib-pair linkage analysis. Am J Hum Genet 52:362–374

Huang J, Vieland VJ (2001*a*) Comparison of "model-free" and "model-based" linkage statistics in the presence of locus heterogeneity: single data set and multiple data set appli-

cations. Hum Hered 51:217–225
——— (2001*b*) The null distribution of the heterogeneity lod scores does depend on the assumed genetic model for the trait. Hum Hered 52:217–222
Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 58:1347–1363
Ott J (1983) Linkage analysis and family classification under heterogeneity. Ann Hum Genet 47:311–320
Risch N (1990) Linkage strategies for genetically complex traits. III. The effect of marker polymorphism on analysis of affected relative pairs. Am J Hum Genet 46:242–253
Vieland VJ, Logue M. HLODs, trait models, and ascertainment: implications of admixture for parameter estimation and linkage detection. Hum Hered (in press)
Vieland VJ, Wang K, Huang J (2001) Power to detect linkage based on multiple sets of data in the presence of locus heterogeneity: comparative evaluation of model-based linkage methods for ASP data. Hum Hered 51:199–208
Whittemore AS, Halpern J (2001) Problems in the definition, interpretation, and evaluation of genetic heterogeneity. Am J Hum Genet 68:457–465

Address for correspondence and reprints: Dr. Susan E. Hodge, Psychiatry and Biostatistics, Columbia University, NYSPI, Unit 24, 1051 Riverside Drive, New York, NY 10032. E-mail: seh2@columbia.edu

0002-9297/2002/7002-0029$15.00

Am. J. Hum. Genet. 70:558–559, 2002

Reply to Hodge et al.

To the Editor:

We thank Hodge et al. (2001 [in this issue])for the chance to discuss an important question: how should we analyze linkage data when we suspect that our families may be segregating more than one disease-susceptibility gene and that disease in some of them may be nonhereditary (i.e., due to nonhereditary factors or to chance)? In our original article (Whittemore and Halpern 2001), we argued against the admixture model of Smith (1963), which is often used to address this problem. This model assumes that all families are hereditary and that a fraction (α) of them segregate deleterious alleles of a gene in a region of interest, while the remaining families segregate alleles of other genes. Testing for linkage at a marker in the region involves specifying the frequencies and penetrances of genotypes of the gene of interest and then maximizing Smith's admixture likelihood with respect to both the fraction α and the recombination fraction (θ) between marker and trait locus. The problem is that, when the likelihood-ratio statistic from this model (i.e., the HLOD) is large, the maximum-likelihood estimate of α is often reported, with its implication that $100\alpha\%$ of disease X is due to a gene in the linked region.

In our original article, we showed that defining, interpreting, and estimating the parameter α is fraught with fundamental logical problems and major statistical pitfalls. We argued that α is not meaningful except under strong and unrealistic assumptions about the data. Moreover, even in the unlikely event that the data meet all these assumptions, estimates of α are quite sensitive to misspecification of the unknown phenocopy rate. Finally, even if the data meet all the necessary assumptions and the investigator specifies the phenocopy rate correctly, the estimates of α that are produced by standard linkage programs are calculated incorrectly and therefore are biased in the presence of phenocopies. We showed how to fix the last problem by correcting the software estimates, but, nevertheless, we recommended against using the HLOD, even as a tool for detection of linkage.

In their letter, Hodge et al. agree with us about the difficulties with α, but they take issue with our recommendation. They cite simulation studies, which suggest (*a*) only slight power loss for the HLOD test compared with the test based on the correct model and (*b*) superior power for the HLOD test compared with NPL (i.e., nonparametric linkage) tests. They also note the need for additional power comparisons between HLOD tests and NPL tests.

We agree with Hodge et al. that the relative power of HLOD and NPL tests needs more work. We also agree that, in some situations, the HLOD test may have greater power than does an NPL test. But the published evidence that they cite does not convince us that such power advantage holds more generally, when the data arise from mechanisms that differ from the rather special models used to generate the simulated data. For example, the models used in several of the papers cited by Hodge et al. assume that all cases of the disease are hereditary, which limits the generalizability of their findings. Furthermore, in the analysis of the simulated data, the correct penetrances of the relevant genotypes are sometimes assumed to be known, which is unrealistically favorable to the HLOD test.

Any power comparison among tests must begin by equating their performance under the null hypothesis—that is, when there is no gene to detect. However, the distribution of the HLOD test statistic under this null hypothesis is complex. Faraway (1993) studied it in the simple, idealized case when the outcome (recombinant vs. nonrecombinant gamete) is known for all informative meioses in all families. Even in this simple case, he found the distribution to be complicated, and he suggested using an approximation to it. In practice, the recombinant statuses of all meioses are seldom known,

and probability distributions must be assigned to them. It is not clear whether Faraway's results extend to this situation.

Moreover, Faraway did not evaluate agreement between his approximate distribution and the true distribution in the extreme tails of the latter. Lander and Kruglyak (1995) have argued that pointwise linkage-test statistics must achieve a nominal significance level of $\sim 10^{-5}$, in order to provide an overall significance level of .05 in a genomewide scan. We know little about the performance of the HLOD test statistic (*a*) in the extreme tails of its null distribution and (*b*) when the recombinant statuses of informative meioses must be inferred. In contrast, the null distribution of the NPL test proposed by Kong and Cox (1997) has been shown to conform well to the theoretical distribution on which its *P* values are based, even in its extreme tails (Nicolae et al. 1998). This issue is important, because even a small inflation of the pointwise type I–error rate could yield an overall false-positive rate that unacceptably exceeds the nominal 5% level.

In conclusion, we thank Hodge et al. for supporting our warnings that estimates of α can be misleading. And we agree with them on the need for further research on the relative power of HLOD and NPL tests to detect linkage. This research should examine test sizes in the tails of the null distributions. The models used to generate the simulated data should include nonhereditary disease, at least two disease-causing genes whose variants have different penetrances, and genes whose variants are common enough so that some families segregate more than one of them. The models used to analyze the simulated data should not be based on the correct values of either the phenocopy rate, the penetrances, or the deleterious-allele frequencies. Meanwhile, whatever may be the possible virtues of the HLOD test, we believe that its use for detection of linkage presents unresolved difficulties.

Acknowledgments

The authors received support from National Institutes of Health grant R35-CA47448

ALICE S. WHITTEMORE AND JERRY HALPERN
Department of Health Research and Policy
Stanford University School of Medicine
Stanford, CA

References

Faraway JJ (1993) Distribution of the admixture test for the detection of linkage under heterogeneity. Genet Epidemiol 10:75–83

Hodge SE, Vieland VJ, Greenberg DA (2001) HLODs remain powerful tools for detection of linkage in the presence of genetic heterogeneity. Am J Hum Genet 70:556–558 (in this issue)

Kong A, Cox NJ (1997) Allele-sharing models: LOD scores and accurate linkage tests. Am J Hum Genet 61:1179–1188

Lander E, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 11:241–246

Nicolae DL, Frigge ML, Cox NJ, Kong A (1998) Discussion [of article "Multipoint linkage analysis using affected relative pairs and partially informative markers"]. Biometrics 54:1271–1274

Smith CAB (1963) Testing for heterogeneity of recombination fraction values in human genetics. Ann Hum Genet 27: 175–182

Whittemore AS, Halpern J (2001) Problems in the definition, interpretation, and evaluation of genetic heterogeneity. Am J Hum Genet 68:457–465

Address for correspondence and reprints: Dr. Alice S. Whittemore, Stanford University School of Medicine, HRP Redwood Building, Stanford University School of Medicine, Stanford CA 94305-5405. E-mail: alicesw@stanford.edu

0002-9297/2002/7002-0030$15.00