Am. J. Hum. Genet. 69:1149–1150, 2001

Reporting of Linkage Results

To the Editor:

In a recent report in the *Journal,* Province (2001, p. 661) correctly points out that a bias occurs if Fisher's method of combining *P* values is used naively to pool the results of "many of the very popular nonparametric (i.e., model free) linkage methods (in particular, variance components, affected sib pair, and extremely discordant sib-pair linkage)." Indeed, Morton (2000, p. 9) has stated that "methods that force lods to be non-negative values create a serious bias that should be avoided as far as possible." The earliest variance-component linkage method is that formulated by Haseman and Elston (1972), and the same methods as were described in that paper to estimate the proportion of alleles that a sib pair share identical by descent can be used, in the so-called mean test, to test for linkage on the basis of data on either affected sib pairs or extremely discordant sib pairs. This has long been implemented in the computer program SIBPAL. This program automatically computes one-tailed *P* values that are uniformly distributed between 0 and 1, under the null hypothesis of no linkage, and so do not suffer the bias that occurs whenever LOD scores truncated at 0 are reported. It has been argued elsewhere (Elston1998; Nyholt 2000) that it would be preferable to summarize the results of linkage analysis in terms of *P* values rather than in terms of LOD scores, and the problem discussed by Province is a further reason for doing so. Although the bias can be obviated in the manner described by Province, a more accurate pooling of linkage results would be obtained if every study in which there is evidence against linkage is assigned the relevant *P* value between $\frac{1}{2}$ and 1, rather than assigning to all such studies an average value. Pooling of *P* values in this manner was done by Wilson and Elston (1995) in a meta-analysis of linkage results for alcoholism. All that would be required for this method of pooling to become standard, if we are willing to make the asymptotic assumption to convert LOD scores to *P* values, is for software that currently truncates LOD scores at 0 to calculate and quote negative LOD scores, corresponding to negative variance components and/or recombination fractions $> \frac{1}{2}$, when this is appropriate.

Consider, for example, the sib-pair methods based on maximizing a LOD score over allele-sharing probabilities z_i, where z_i is the probability that a sib pair shares i alleles identical by descent at a genomic location. Because $z_2 = 1 - z_0 - z_1$, the parameter space $0 \leq z_0, z_1, z_2 \leq 1$ can be represented by the triangle depicted in figure 1, in which the smaller hatched triangle represents the possible parameter values for affected sib pairs when there is simple monogenic inheritance (Holmans 1993). When a complex disease is being studied, linkage can be missed if the maximization for affected sib pairs is restricted to this smaller hatched triangular area (Dizier et al. 2000); on the other hand, maximizing the LOD score over the whole parameter space depicted in figure 1 can lead to a positive LOD score when there is evidence against linkage. The dashed line in figure 1 divides the parameter space into two parts, corresponding to the proportion of alleles shared by a pair of sibs being either $> \frac{1}{2}(z_2 > z_0)$ or $< \frac{1}{2}(z_2 < z_0)$. Maximization should be over the entire parameter space, but the resulting maximum LOD score should be made negative if the result is in the direction opposite to linkage (see fig. 1). In this particular case, because >1 df is involved, the conversion of a maximum LOD score to a *P* value is not simple, even if based on asymptotic considerations. In other situations (e.g., see Kong and Cox 1997; Whittemore and Tu 1998), when only 1 df is involved, the conversion is simple if we are prepared to make asymptotic assumptions. In any cases of doubt, a permutation test (sampling from the entire permutation distribution, if necessary, for computational feasibility) can be performed. As pointed out elsewhere (Elston et al. 1996) ~10,000 sib pairs are required if, in an affected-sib-pair study, a quoted *P* value of 10^{-4} is to be within 20% of the true value when we rely on as-

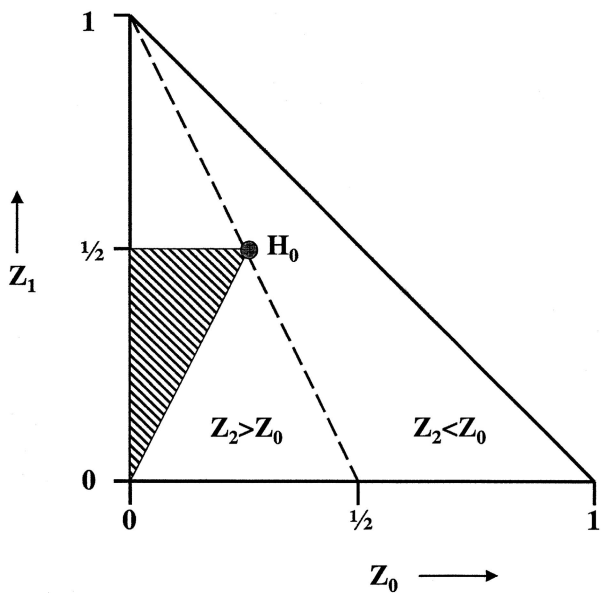

Figure 1 Parameter space for sib-pair allele sharing. H_0 is the null hypothesis (no linkage) point $z_0 = \frac{1}{4}, z_1 = \frac{1}{2}, z_2 = \frac{1}{4}$. When concordant sib pairs are studied, points to the left of the dashed line are in the direction of linkage, and points to the right are in the direction opposite to linkage; when discordant sib pairs are studied, the converse is the case. The smaller hatched triangle represents the parameter space for concordant sib pairs when there is simple monogenic inheritance.

ymptotic assumptions to calculate it. In any case, it would seem preferable for analysts to quote one-sided *P* values rather than to have to resort to the average bias-controlling procedure suggested by Province—although this should not be taken to imply that either *P* values or LOD scores are completely satisfactory summary statistics.

Acknowledgments

This work was supported in part by grants from the U.S. Public Health Service: research grant GM-28356, from the National Institute of General Medical Sciences; research grant DK-57292, from the National Institute of Diabetes and Digestive and Kidney Diseases; and resource grant RR-03655, from the National Center for Research Resources.

ROBERT C. ELSTON
Department of Epidemiology and Biostatistics, School of Medicine
Case Western Reserve University
Cleveland

References

Dizier HM, Quesneville H, Prum B, Selinger-Leneman H, Clerget-Darpoux F (2000) The triangle test statistic (TTS): a test of genetic homogeneity using departure from the triangle constraints in IBD distribution among affected sib-pairs. Ann Hum Genet 64:433–442

Elston RC (1998) Methods of linkage and analysis and the assumptions underlying them. Am J Hum Genet 63:931–934

Elston RC, Guo X, Williams LV (1996) Two-stage global search designs for linkage analysis using pairs of affected relatives. Genet Epidemiol 13:535–558

Haseman JK, Elston RC (1972) The investigation of linkage between a quantitative trait and a marker locus. Behav Genet 2:3–19

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

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

Morton NE (2000) Unsolved problems in genetic epidemiology. Hum Hered 50:5–13

Nyholt DR (2000) All LODs are not created equal. Am J Hum Genet 67:282–288

Province MA (2001) The significance of not finding a gene. Am J Hum Genet 69:660–663

Whittemore AS, Tu IP (1998) Simple, robust linkage tests for affected sibs. Am J Hum Genet 62:1228–1242

Wilson AF, Elston RC (1995) Linkage analysis in the study of the genetics of alcoholism. In: Begleiter H, Kissin B (eds) The genetics of alcoholism. Oxford University Press, New York, pp 353–376

Address for correspondence and reprints: Dr. Robert C. Elston, Department of Epidemiology & Biostatistics, Case Western Reserve University, Cleveland, OH 44109. E-mail: rce@darwin.cwru.edu

0002-9297/2001/6905-0024$02.00

Am. J. Hum. Genet. 69:1150–1152, 2001

Examinations of Methylenetetrahydrofolate Reductase C677T and A1298C Mutations—and In Utero Viability

To the Editor:

The recently published study by Isotalo et al. (2000) analyzed the methylenetetrahydrofolate reductase (MTHFR) C677T and A1298C mutations in neonatal and fetal groups, to determine whether particular MTHFR genotype combinations are associated with decreased in utero viability. Isotalo et al. (2000) observed all possible genotype combinations in the fetal group, but combined 677CT/1298CC and 677TT/1298CC genotypes were not observed in the neonatal group. Therefore, they hypothesized that decreased viability exists among fetuses carrying the 677CT/1298CC and 677TT/1298CC genotypes, with a possible selection disadvantage in fetuses with an increased number of mutant MTHFR alleles. They also did not observe the 677CT/1298CC and 677TT/1298CC genotypes in a population consisting of healthy adult controls.

We have tested for the MTHFR C677T and A1298C mutations in a Hispanic population of Mexican descent, to determine risk for spina bifida (SB) (Volcik et al. 2000). Although we observed all possible MTHFR 677/1298 ge-

Table 1

Combined MTHFR C677T/A1298C Genotype or Allele Frequencies in a Hispanic Population of Mexican Descent, Composed of Patients with SB, Their Parents, and Controls

Genotype or Allele	Observed Frequency in			
	Patients (n = 302)	Mothers (n = 281)	Fathers (n = 143)	Controls (n = 82)
MTHFR C677T/A1298C genotype:				
CT/CC	.003	.000	.007	.012
TT/AC	.020	.028	.021	.024
TT/CC	.000	.004	.000	.000
MTHFR allele:				
677C	.493	.477	.549	.527
677T	.507	.523	.451	.473
1298A	.854	.856	.815	.811
1298C	.146	.144	.185	.189