Letters to the Editor

Am. J. Hum. Genet. 64:1785, 1999

Combining the Sibling Disequilibrium Test and Transmission/Disequilibrium Test for Multiallelic Markers

To the Editor:

Horvath and Laird (1998) describe the SDT (sibling disequilibrium test) which, like the sibling-association test (Curtis 1997), is a test for association in addition to linkage even when applied to sibships larger than sib pairs. These tests thus differ from the sibling transmission disequilibrium test (S-TDT [Spielman and Ewens 1998]), which is a test for linkage but not for association (unless attention is restricted to sib pairs). The possible advantage that the SDT has over Curtis's test is that it uses all affected sibs in the sibship, although it does not allow for special provision to be made to detect a recessive effect by testing whether there is an excess of affected sibs homozygous for one particular allele. Horvath and Laird demonstrate how the SDT can be applied to a multiallelic marker and how, in the case of a biallelic marker, the SDT and TDT can be combined, but they do not show how the tests can be combined for a multiallelic maker. Curtis described by using logistic regression how his test could be combined with multiallelic TDT data as implemented in the extended TDT (ETDT [Sham and Curtis 1995]), and here we show, using their multivariate sign test, that it is straightforward to apply Horvath and Laird's own approach to combine the multiallelic SDT with multiallelic TDT data.

Horvath and Laird use the "component" sign test (Bickel 1965; Randles 1989) as follows. For N sibships and a marker with m alleles, let s_i^j be 1, 0, or -1, according to whether, in the ith sibship, the frequency of allele j in affected sibs is higher than, equal to, or lower than that in unaffected sibs. Then define $\mathbf{S} = (S^1, S^2, \ldots, S^{m-1})$ where $S^j = \Sigma_{i=1}^{N} s_i^j$ and a matrix $\mathbf{W}$ having elements $W_{jk} = \Sigma_{i=1}^{N} s_i^j s_i^k$. The multiallelic SDT statistic is then $T = \mathbf{S}'\mathbf{W}^{-1}\mathbf{S}$, which is asymptotically χ^2_{m-1} under the null hypothesis of no association or no linkage. In order to extend this approach to include TDT data, we note that we can apply exactly the same formula to a sample of $N/2$ trios (containing N parents) by using s^j_i to denote, instead, the transmission for the ith parent, being 1 or -1 if the parent has one copy of allele j and, respectively, does or does not transmit it to the affected subject and being 0 if the parent is uninformative for allele j (i.e., has 0 or 2 copies). Then the same statistic, $T = \mathbf{S}'\mathbf{W}^{-1}\mathbf{S}$, provides a non-parametric multiallelic TDT statistic. (This test is mathematically identical to the Stuart [1955] test presented by Sham [1997], and is in fact the score test of the Bradley-Terry model [Bradley and Terry 1952].) Of course it is obvious that we can sum both forms of s_i^j over a mixed sample of sibships and trios in a combined multiallelic SDT and TDT analysis. Formally, if we write $\mathbf{S}_{\mathrm{SDT}}$ and $\mathbf{W}_{\mathrm{SDT}}$ for the totals derived from the sibship data and $\mathbf{S}_{\mathrm{TDT}}$ and $\mathbf{W}_{\mathrm{TDT}}$ for those from the trios then $\mathbf{S}_{\mathrm{BOTH}} = \mathbf{S}_{\mathrm{SDT}} + \mathbf{S}_{\mathrm{TDT}}$, $\mathbf{W}_{\mathrm{BOTH}} = \mathbf{W}_{\mathrm{SDT}} + \mathbf{W}_{\mathrm{TDT}}$ and $T_{\mathrm{BOTH}} = \mathbf{S}'_{\mathrm{BOTH}}\mathbf{W}^{-1}_{\mathrm{BOTH}}\mathbf{S}_{\mathrm{BOTH}}$ is the combined statistic.

In order to use TDT data from families with more than one affected child, we can follow Martin et al. (1997) and can define s_i^j for the ith parent as being 1 if the parent is heterozygous for allele j and transmits this allele to more than half the affected children, -1 if the allele is transmitted to fewer than half the affected children, and 0 if the allele is transmitted to exactly half the affected children or if the parent is uninformative for this allele. (When there is only one affected child, this scoring scheme is equivalent to that given above.) TDT data can be used if only one parent is genotyped, provided that affected children homozygous for the marker are disregarded (Curtis and Sham 1995). When all these procedures are combined, the summations of the appropriate s_i^j can be performed over families consisting of discordant sibships and consisting of one or two parents having one or more affected children. The overall statistic $T_{\mathrm{ALL}} = \mathbf{S}'_{\mathrm{ALL}}\mathbf{W}^{-1}_{\mathrm{ALL}}\mathbf{S}_{\mathrm{ALL}}$ provides a test, for association with linkage, that makes appropriate use of all the available information from these different family types and that is asymptotically χ^2_{m-1}.

We propose that further efforts could proceed in three directions. First, the work of Horvath and Laird that considers the relative power of SDT and TDT could be extended in order to determine which is preferable to apply to a family suitable for either. This would depend on the transmission model of the disease and on the numbers of parents, affected siblings, and unaffected sib-

lings who were genotyped. Second, a comparison of the performance of the above test versus those of tests utilizing logistic regression would be of interest. Third, the appropriateness of the asymptotic distribution could be investigated, since, for markers having large numbers of alleles, it might be that a Monte Carlo approach to assessment of significance could be desirable.

Acknowledgments

This work was supported by Wellcome Trust Project Grant 055379.

DAVID CURTIS,[1] MICHAEL B. MILLER,[3] AND PAK C. SHAM[2]

[1]*Academic Department of Psychological Medicine, St. Bartholomew's and Royal London School of Medicine and Dentistry,* [2]*Department of Psychology, University of Missouri and* [3]*Department of Psychological Medicine, Institute of Psychiatry, London*

References

Bickel PJ (1965) On some asymptotic competitors to Hotelling's T^2. Ann Math Stat 36:160–173

Bradley RA, Terry ME (1952) Rank analysis of incomplete block designs. I. The method of paired comparisons. Biometrika 39:324–345

Curtis D (1997) Use of siblings as controls in case-control association studies. Ann Hum Genet 61:319–333

Curtis D, Sham PC (1995) A note on the application of the transmission disequilibrium test when a parent is missing. Am J Hum Genet 1995 56:811–812

Horvath S, Laird NM (1998) A discordant-sibship test for disequilibrium and linkage: no need for parental data. Am J Hum Genet 63:1886–1897

Martin ER, Kaplan NL, Weir BS (1997) Tests for linkage and association in nuclear families. Am J Hum Genet 61: 439–448

Randles RH (1989) A distribution-free multivariate sign test based on interdirections. J Am Stat Assoc 84:1045–1050

Sham P (1997) Transmission/disequilibrium tests for multiallelic loci. Am J Hum Genet 61:774–778

Sham PC, Curtis D (1995) An extended transmission/disequilibrium test (TDT) for multi-allele marker loci. Ann Hum Genet 59:323–336

Spielman RS, Ewens WJ (1998) A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test. Am J Hum Genet 62:450–458

Stuart A (1955) A test of homogeneity of the marginal distribution in a two-way classification. Biometrika 42:412–416

Address for correspondence and reprints: Dr. David Curtis, Department of Adult Psychiatry, 3d Floor, Outpatient Building, Royal London Hospital, Whitechapel, London E1 1BB, United Kingdom. E-mail: dcurtis@hgmp.mrc.ac.uk

0002-9297/99/6406-0034$02.00