LETTERS TO THE EDITOR

Am. J. Hum. Genet. 64:326, 1999

The Interpretation of the Parameters in the Transmission/Disequilibrium Test

To the Editor:

The transmission/disequilibrium test (TDT) proposed by Spielman et al. (1993) is a valid test for linkage in structured populations, irrespective of whether the families are simplex, multiplex, or multigenerational (Spielman and Ewens 1996). Its power to detect linkage of complex traits is potentially greater than that of allele-sharing methods (Risch and Merikangas 1996). The original TDT has been extended by a number of groups—including Sham and Curtis (1995), Schaid (1996), and Spielman and Ewens (1996)—to the study of multiallelic markers. Assuming Hardy-Weinberg equilibrium for the population under study, I present relationships between parameters in the TDT and the disease susceptibility carried by marker alleles. I hope that these relationships make the interpretation of the observed transmission disequilibrium more intuitive and that they yield further insight into the TDT.

Consider a marker H with n alleles, $H_1,\ldots H_n$, having allele frequencies $h_1,\ldots h_n$. Assume that the disease gene D has two alleles, D_1 and D_2, with allele frequencies p_1 and $p_2 = 1 - p_1$, respectively, and that the penetrance for genotype D_uD_v is f_{uv}, where $u, v = 1$ or 2. Furthermore, denote the recombination fraction between H and D as θ and measure the set of linkage-disequilibrium values, between these two loci, in terms of $\delta_{H_iD_u} = P(H_iD_u) - h_ip_u$, where $i = 1, \ldots, n$ and $u = 1$ or 2. First, consider the transmission of a marker allele from *one* parent to the affected offspring. Let $P_{ij} = P$(a parent has genotype H_iH_j and transmits H_i|offspring is affected). Sethuraman (1997) has shown that, if a sample of affected children together with their parents are ascertained at random from a population in Hardy-Weinberg equilibrium, then

$$\begin{aligned}
P_{ij} = {} & (p_1f_{11} + p_2f_{12})[h_jP(H_iD_1) \\
& -\theta(h_j\delta_{H_iD_1} - h_i\delta_{H_jD_1})]/K \\
& +(p_1f_{12} + p_2f_{22})[h_jP(H_iD_2) \\
& -\theta(h_j\delta_{H_iD_2} - h_i\delta_{H_jD_2})]/K \ , \\
P_{ji} = {} & (p_1f_{11} + p_2f_{12})[h_iP(H_jD_1) \\
& -\theta(h_i\delta_{H_jD_1} - h_j\delta_{H_iD_1})]/K \\
& +(p_1f_{12} + p_2f_{22})[h_iP(H_jD_2) \\
& -\theta(h_i\delta_{H_jD_2} - h_j\delta_{H_iD_2})]/K \ ,
\end{aligned}$$

where $K = p_1^2f_{11} + 2p_1p_2f_{12} + p_2^2f_{22}$ is the disease prevalence in the population. Sham and Curtis (1995) expressed the P_{ij} in different but equivalent forms. In the following discussion, assume that $\theta = 0$, since, between loci having linkage disequilibrium, θ is generally very close to 0. When $\theta = 0$,

$$\begin{aligned}
P_{ij} = {} & h_j[(p_1f_{11} + p_2f_{12})P(H_iD_1) \\
& +(p_1f_{12} + p_2f_{22})P(H_iD_2)]/K \ , \\
P_{ji} = {} & h_i[(p_1f_{11} + p_2f_{12})P(H_jD_1) \\
& +(p_1f_{12} + p_2f_{22})P(H_jD_2)]/K \ . \qquad (1)
\end{aligned}$$

Let P(affected$|H_i) = P$(an individual is affected | this individual receives allele H_i from one parent). This conditional probability can be regarded as the genetic risk that allele H_i carries for the disease susceptibility. The value of P(affected | H_i) can be calculated as follows:

$$
\begin{aligned}
&P(\text{an individual is affected} \mid \text{this individual receives } H_i \text{ from one parent}) \\
&= P(\text{an individual receives } H_i \text{ from one parent and is affected})/P(H_i) \\
&= \sum_{u=1}^{2}\sum_{v=1}^{2} P(\text{an individual receives } H_iD_u \text{ from one parent, receives } D_v \text{ from the other parent, and is affected})/h_i \\
&= \sum_{u=1}^{2}\sum_{v=1}^{2} P(\text{an individual receives } H_iD_u \text{ from one parent and receives } D_v \text{ from the other parent})P(\text{affected} \mid \text{genotype } D_uD_v)/h_i \\
&= \sum_{u=1}^{2}\sum_{v=1}^{2} P(H_iD_u)P(D_v)f_{uv}/h_i \ .
\end{aligned}
$$

The last equation follows from the assumption of Hardy-Weinberg equilibrium. Thus,

$$
\begin{aligned}
P(\text{affected} \mid H_i) &= [(p_1f_{11} + p_2f_{12})P(H_iD_1) \\
&\quad + (p_1f_{12} + p_2f_{22})P(H_iD_2)]/h_i \ , \\
P(\text{affected} \mid H_j) &= [(p_1f_{11} + p_2f_{12})P(H_jD_1) \\
&\quad + (p_1f_{12} + p_2f_{22})P(H_jD_2)]/h_j \ . \qquad (2)
\end{aligned}
$$

From equations (1) and (2),

$$
\begin{aligned}
P_{ij} &= P(\text{affected} \mid H_i)h_ih_j/K \ , \\
P_{ji} &= P(\text{affected} \mid H_j)h_ih_j/K \ .
\end{aligned}
$$

Therefore, the following relationship relates P_{ij} to $P(\text{affected} \mid H_i)$:

$$
\frac{P_{ij}}{P_{ji}} = \frac{P(\text{affected} \mid H_i)}{P(\text{affected} \mid H_j)} \ . \qquad (3)
$$

From equation (3), it can be seen that the transmission/disequilibrium ratio P_{ij}/P_{ji} for H_iH_j parents is *independent* of allele frequencies. For n alleles, the $n(n-1)/2$ transmission/disequilibrium ratios P_{ij}/P_{ji} are determined by $n-1$ independent parameters $P(\text{affected} \mid H_i)$.

Let n_{ij} denote the number of H_iH_j parents who transmit H_i to the affected offspring. Then, conditional on $n_{ij} + n_{ji}$ $(i \neq j)$, n_{ij} follows the binomial distribution

$$
B\left[n_{ij} + n_{ji}, \frac{P(\text{affected} \mid H_i)}{P(\text{affected} \mid H_i) + P(\text{affected} \mid H_j)}\right] .
$$

This naturally leads to the logistic regression proposed by Sham and Curtis (1995), who did not interpret the parameters as the genetic risks carried by different marker alleles.

Schaid (1996) studied the case when the joint transmission of two parents is considered simultaneously. Let

$$
\begin{aligned}
P_{ik,jl} = P(&\text{one parent has genotype } H_iH_j \\
&\text{and transmits } H_i \text{ and the other} \\
&\text{parent has genotype } H_kH_l \text{ and transmits} \\
&H_k \mid \text{offspring is affected}) \ .
\end{aligned}
$$

Without the assumption of Hardy-Weinberg equilibrium, it can be shown that

$$
\frac{P_{ik,jl}}{P_{jl,ik}} = \frac{P(\text{affected} \mid H_iH_k)}{P(\text{affected} \mid H_jH_l)} \ , \qquad (4)
$$

where $P(\text{affected} \mid H_iH_k)$ is the conditional probability that a person having genotype H_iH_k is affected—that is, the penetrance for marker genotype H_iH_k (Schaid 1996). Therefore, when two parents are considered jointly, $P_{ik,jl}/P_{jl,ik}$ is determined by the ratio of the penetrances for genotypes H_iH_k and H_jH_l. For n alleles, there are $n(n+1)/2$ possible genotypes; so, a total of $n(n+1)/2 - 1$ parameters are needed to quantify $P_{ik,jl}/P_{jl,ik}$. Schaid (1996) discussed several ways to code the genotypes in the transmission/disequilibrium test.

When each parent is examined separately, the contributions from the two parents are implicitly assumed to be independent. This is true if

$$
\frac{P(\text{affected} \mid H_iH_k)}{P(\text{affected} \mid H_jH_l)} = \frac{P(\text{affected} \mid H_i)}{P(\text{affected} \mid H_j)}\frac{P(\text{affected} \mid H_k)}{P(\text{affected} \mid H_l)} \ ,
$$

which holds if and only if $f_{12}^2 = f_{11}f_{22}$ (Knapp et al. 1993).

The relationships in equations (3) and (4) have been used to develop transmission/disequilibrium tests for multiple tightly linked markers (H. Zhao, K. R. Merikangas, and K. K. Kidd, unpublished results). They are also useful in the study of gene-environment interactions. For simplicity, assume that an environmental exposure R—for example, smoking—is classified as being either present $(R = 1)$ or absent $(R = 0)$. Let $P_{ik,jl}^R = P(\text{one parent has genotype } H_iH_j \text{ and transmits } H_i$, and the other parent has genotype H_kH_l and transmits $H_k \mid$ offspring is affected and environmental exposure is R). Denote the penetrance for genotype D_uD_v under

environmental exposure R by means of f^R_{uv}. If genotype D_uD_v and environmental exposure R have multiplicative effects on the disease susceptibility—that is, $f^1_{uv} = \lambda f^0_{uv}$—then it can be shown that

$$\frac{P^1_{ik,jl}}{P^1_{jl,ik}} = \frac{P^0_{ik,jl}}{P^0_{jl,ik}} .$$

Therefore, $P^R_{ik,jl}/P^R_{jl,ik}$ is independent of the environmental exposure variable R when the genotypes and the environmental exposure have multiplicative effects on the disease susceptibility. To test nonmultiplicative gene-environment interactions, standard statistical tests may be performed, to determine whether the transmission/disequilibrium ratios among families having the exposure—that is, $P^1_{ik,jl}/P^1_{jl,ik}$—are the same as the transmission/disequilibrium ratios among families without the exposure—that is, $P^0_{ik,jl}/P^0_{jl,ik}$. However, for nonmultiplicative gene-environmental interactions—for example, additive gene-environment interactions with $f^1_{uv} = \delta + f^0_{uv}$—there is no simple relationship between $P^1_{ik,jl}/P^1_{jl,ik}$ and $P^0_{ik,jl}/P^0_{jl,ik}$.

Acknowledgments

I thank two referees for their thoughtful and constructive comments. This work was supported in part by National Institutes of Health grants GM59507, GM57672, and DA09055.

HONGYU ZHAO
Department of Epidemiology and Public Health, Yale University School of Medicine, New Haven

References

Knapp M, Seuchter SA, Baur MP (1993) The haplotype-relative-risk (HRR) method for analysis of association in nuclear families. Am J Hum Genet 52:1085–1093

Risch N, Merikangas K (1996) The future of genetic studies of complex human diseases. Science 273:1516–1517

Schaid DJ (1996) General score tests for associations of genetic markers with disease using cases and their parents. Genet Epidemiol 13:423–449

Sethuraman B (1997) Topics in statistical genetics. PhD diss, University of California at Berkeley, Berkeley

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 (1996) The TDT and other family-based tests for linkage disequilibrium and association. Am J Hum Genet 59:983–989

Spielman RS, McGinnis RE, Ewens WJ (1993) Transmission test for linkage disequilibrium: the insulin gene and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet 52: 506–516

Address for correspondence and reprints: Dr. Hongyu Zhao, Department of Epidemiology and Public Health, 60 College Street, Yale University School of Medicine, New Haven, CT 06520-8034. E-mail: hongyu.zhao@yale.edu

0002-9297/99/6401-0046$02.00