Am. J. Hum. Genet. 63:1507–1516, 1998

A Comparative Study of Sibship Tests of Linkage and/or Association

S. A. Monks,[1,2] N. L. Kaplan,[1] and B. S. Weir[2]

[1]Biostatistics Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC; and [2]Program in Statistical Genetics, Department of Statistics, North Carolina State University, Raleigh

Summary

Population-based tests of association have used data from either case-control studies or studies based on trios (affected child and parents). Case-control studies are more prone to false-positive results caused by inappropriate controls, which can occur if, for example, there is population admixture or stratification. An advantage of family-based tests is that cases and controls are well matched, but parental data may not always be available, especially for late-onset diseases. Three recent family-based tests of association and linkage utilize unaffected siblings as surrogates for untyped parents. In this paper, we propose an extension of one of these tests. We describe and compare the four tests in the context of a complex disease for both biallelic and multiallelic markers, as well as for sibships of different sizes. We also examine the consequences of having some parental data in the sample.

Introduction

Family-based tests such as the transmission/disequilibrium test (TDT) have proved to be powerful tools in the search for disease genes (Spielman and Ewens 1996). Depending on the data structure, these are tests of either linkage or linkage and association. Unlike case-control tests, the tests are not affected by population stratification, which can lead to spurious associations of a marker allele with disease susceptibility.

Family-based tests have largely required knowledge of parental marker genotypes; however, for late-onset diseases, parental data are often not available. Recently, Curtis (1997), Boehnke and Langefeld (1998), and Spielman and Ewens (1998) have each developed tests of linkage and association that use unaffected siblings as surrogates for untyped parents. For two-allele markers, the three tests are equivalent, but, for multiple-allele markers, the tests can be different. For a multiple-allele marker, Curtis proposed a likelihood ratio (LR) test similar to the extended transmission/disequilibrium test of Sham and Curtis (1995). The discordant-alleles test (DAT) of Boehnke and Langefeld is based on a Pearson homogeneity statistic for a $2 \times m$ contingency table, and a permutation procedure is used to perform the test. For the sib-TDT (S-TDT), Spielman and Ewens calculated a two-allele statistic for each marker allele; the test is based on the maximum of the absolute value of these statistics. To perform the test, they use the same permutation procedure as Boehnke and Langefeld use.

The S-TDT generalizes to sibships that contain more than single affected and single unaffected siblings; however, when using these larger sibships, the test is valid only as a test of linkage. An alternative approach for analyzing larger sibships, proposed by Curtis (1997) and resulting in a test of linkage and association, is to reduce each sibship to two siblings, by first randomly choosing an affected individual and then choosing the unaffected sibling whose marker genotype is maximally different from that of the affected sibling. This strategy does not depend on the test statistic and can also be used for the tests proposed by Boehnke and Langefeld (1998) and Spielman and Ewens (1998).

The multiallele TDT statistic, T_{mhet}, proposed by Spielman and Ewens (1996), suggests a sib statistic, T_{MSTDT}, that can be calculated for a marker with any number of alleles and for sibships of any size, with the assumption that each sibship has at least one affected and at least one unaffected sibling. The statistic generalizes the S-TDT statistic, since it reduces to the square of the statistic for a biallelic marker. Just as for S-TDT, the test based on T_{MSTDT} is a test of linkage and association if the sample contains only sibships with one affected and one unaffected individual and is a test of linkage if the sample contains larger sibships. The permutation procedure of Boehnke and Langefeld (1998) and Spielman and Ewens (1998) can be used to perform the test. Alternatively, simulations show that, under either the null hypothesis of no linkage or the null hypothesis of no linkage or no association, the distribution of T_{MSTDT} can

Received May 4, 1998; accepted for publication August 21, 1998; electronically published October 16, 1998.

Address for correspondence and reprints: S. A. Monks, Biostatistics Branch, NIEHS, P.O. Box 12233, Research Triangle Park, NC 27709. E-mail: monks@stat.ncsu.edu

0002-9297/98/6305-0029$02.00

be approximated with a χ^2 distribution with $m - 1$ df when the marker has m alleles.

To compare the powers of the four tests, a simulation study was conducted for several complex-disease models. For a multiallele marker and sibships of size 2, the powers of the DAT and the T_{MSTDT} test to detect linkage and association are comparable. The power of the S-TDT relative to those of the DAT and the T_{MSTDT} test is dependent on the nature of the association between the marker and disease alleles. If the association is concentrated on only a few of the marker alleles, then the S-TDT can be more powerful. Alternatively, if the association is spread over most of the alleles, then the S-TDT can be less powerful. The LR test was consistently less powerful than the other three tests. When one is testing for linkage, the simulations suggest that it is preferable not to reduce the size of each sibship as proposed by Curtis (1997). However, if one requires a test of linkage and association, then one has no choice but to reduce the size of each sibship, for the tests described in the present study.

Curtis (1997) and Spielman and Ewens (1998) have considered the possibility that some of the sibships in the sample have parental data. The statistic T_{MSTDT} generalizes for these mixed data, and the distribution of the resulting statistic, under either of the null hypotheses, can also be approximated with a χ^2 distribution with $m - 1$ df. The simulations suggest that it is advantageous to combine families for which there are parental data and families for which parental data are not available.

Methods

Notation

Suppose that we have a marker locus with m alleles, $M_1, M_2, \ldots, M_m$, and a disease locus with two alleles, D_1 and D_2. We define D_1 as the "disease" allele and denote its frequency as "p." Penetrances, f_{rs}, are the probabilities that an individual with disease genotype D_rD_s is affected with the disease; it is assumed that $f_{11} \geqslant f_{12} \geqslant f_{22}$. We assume, unless stated otherwise, a random-mating population with no selection or mutation. The population disease prevalence is $A = p^2f_{11} + 2p(1-p)f_{12} + (1-p)^2f_{22}$, with population-attributable risk, $\text{AR} = (A - f_{22})/A$. The recombination fraction between the disease and the marker locus is θ. The frequency of marker allele M_i in the population is $\Pr(M_i)$, whereas $\Pr(M_i \mid D_r)$ is the probability that a gamete carries marker allele M_i, given that it carries allele D_r at the disease locus. Finally, suppose that we have sampled N_S sibships, each of which contains at least one affected individual and at least one unaffected individual. A sibship will not be informative if the marker genotypes are the same for all sibs. The minimal configuration for a sibship is thus one affected and one unaffected sibling with different marker genotypes. For our simulations, N_S is always the number of sampled families, and, therefore, some fraction of these families will not be informative.

If there is no linkage between the marker and the disease loci, then, for each sib, regardless of the size of the sibship, the affection status and the marker genotype are independent. Alternatively, if there is linkage but no association between alleles at the two loci, then the affection status and the marker genotype of each sib are independent only for sibships with exactly one affected and one unaffected sibling. Even in the absence of association, linkage can cause excess sharing of marker alleles among affected siblings, and this is the basis of affected-sib-pair tests (Ott 1991). Therefore, by comparing marker-allele frequencies in affected sibs with those in unaffected sibs, for a sample of sibships with the minimal configuration, we can test the null hypothesis of no linkage or no association; however, for samples containing larger sibships, we can test only the null hypothesis of no linkage.

Permutation Procedure

Boehnke and Langefeld (1998) and Spielman and Ewens (1998) used a permutation procedure to determine significance. Under either the null hypothesis of no linkage or no association, for samples of sibships with minimal configuration, or the null hypothesis of no linkage, for samples containing larger sibships, an affected individual is equally likely to have any of the marker genotypes that are contained in the sibship. Thus, randomly permuting the affection status of the sibs in each sibship leads to a new sample, which, under the null hypothesis, is statistically equivalent to the original sample. To permute a sibship with a affected sibs in a total of t sibs, a of the sibs are selected at random to be affected, and the remaining sibs are labeled "unaffected." By use of this procedure, an empirical null distribution can be constructed for any statistic, T, and significance can be calculated. Rather than calculate T for all possible permutations of the data set, as would be done for a permutation test, a Monte-Carlo approximation is used. The steps are as follows:

1. calculate T, with value T_0, for the data set;
2. for each sibship, randomly permute affection status;
3. calculate T on this pseudosample and determine whether it is more extreme than T_o;
4. repeat steps 2 and 3 N_1 times and estimate the P value as the proportion of times that the test statistic is more extreme than T_o;
5. reject the specified null hypothesis if the p-value is less than or equal to some specified α.

To perform the test on any particular data set, only one

P value is needed, and thus N_I should be large. However, for power calculations, N_O P values must be estimated by simulating N_O data sets from the population, and so, to keep computing time to a feasible level, we have used a smaller value of N_I. For our simulations, $N_I = 300$ with $N_O = 1{,}000$. Our choice of N_I and N_O gave adequate control of the bias and variability of the power estimates. Estimates of significance levels on the basis of the χ^2 distribution do not require the use of the permutation procedure, and thus a larger value of N_O can be used. We have used $N_O = 10{,}000$.

Test Statistics

Curtis (1997) and Boehnke and Langefeld (1998) were interested in a test for association and considered samples of sibships only with the minimal configuration. Curtis (1997) shrank larger sibships to the minimal configuration by selecting an affected individual at random and then choosing the unaffected sibling whose marker genotype was maximally different from that of the affected sibling. Curtis's statistics are based on the values of T_{ij}, with $i \neq j$, which are defined as follows. Each marker allele in the affected individual is compared with each marker allele in that individual's unaffected sibling. If the pair of marker alleles are the same, the comparison is ignored. If the pair of marker alleles are different, then .5 is added to T_{ij}, where M_i is the marker allele in the affected individual and M_j is the marker allele in the unaffected sibling. For a biallelic marker, Curtis (1997) defined a statistic that is asymptotically standard normal under the null hypothesis of no linkage or no association, $Z_c = [T_{12} - (N_1/2 + N_2)]/\sqrt{N_1/4 + N_2}$, where N_i is the number of sibships that increase either T_{12} or T_{21} by i. For multiallele markers, Curtis adopted a likelihood framework similar to that of Sham and Curtis (1995). The likelihood is based on the probabilities, P_{ij} with $i \neq j$, that an affected individual has marker allele M_i whereas that individual's unaffected sibling has marker allele M_j. Curtis modeled P_{ij} with parameters $B_1, B_2, \ldots, B_m$ defined by $\ln(P_{ij}/P_{ji}) = B_i - B_j$. The null hypothesis of no linkage or no association is tested by comparing the likelihood, L_0, evaluated at $B_i = 0$ for all i, to the likelihood, L_a, maximized over B_i, on the assumption that only $B_1 = 0$. Under the null hypothesis of no linkage or no association, likelihood theory would suggest that the distribution of the test statistic $\mathrm{LR} = -2\ln(L_o/L_a)$ can be approximated by a χ^2 distribution with $m - 1$ df. We show, however, that this approximation can perform poorly if only sibships with minimal configuration have been collected.

Boehnke and Langefeld (1998) represent the marker-allele data in a $2 \times m$ contingency table in which the rows represent affection status and the columns represent marker alleles. For sibships of minimal configuration, the genotypes of the affected individuals and their unaffected siblings are different, and hence there are two types of sibships that contribute counts to the contingency table. If the marker alleles in the two sibs are all different, then all four marker alleles are counted in the table. Alternatively, if an affected individual and its unaffected sibling have a marker allele in common, then these marker alleles are ignored and only the two different marker alleles are counted in the table. If n_{ij} denotes the count for the (i,j)th cell, then the statistic is the Pearson homogeneity statistic

$$\mathrm{AC}_2 = \sum_{i=1}^{m} \frac{(n_{1i} - n_{2i})^2}{n_{1i} + n_{2i}} \ .$$

Because of the correlation between sibling genotypes, the Monte Carlo permutation procedure is used to measure significance. Boehnke and Langefeld (1998) considered a number of other statistics but recommended the use of AC_2. They also considered a test based on a maximum statistic. Significance is measured using the same procedure as Spielman and Ewens (1998), described below.

For each marker allele M_i, Spielman and Ewens (1998) define the random variable, Y_i, which is the number of M_i alleles in the affected individuals from all sibships. Under the described permutation procedure, both the theoretical permutation mean, A_{ik}, and the variance, V_{ik}, can be calculated for the kth sibship's contribution to Y_i (for the formulas, see the Appendix). The normalized statistic is

$$Z_i = \frac{Y_i - \sum_{k=1}^{N_S} A_{ik}}{\sqrt{\sum_{k=1}^{N_S} V_{ik}}} \ .$$

For a biallelic marker, the test statistic is Z_1. Significance can be calculated by use of the Monte Carlo permutation procedure, or, for large samples, the P value can be estimated from the standard normal distribution. For a marker locus with m alleles, the statistic is

$$Z_{\max} = \max_{i=1\ldots m} |\, Z_i \,| \ .$$

The Monte Carlo permutation procedure is used to measure significance.

The statistic that we propose for a marker with m alleles is

$$T_{\mathrm{MSTDT}} = \frac{m-1}{m} \sum_{i=1}^{m} Z_i^2 \ .$$

Arguments similar to those of Martin et al. (1997) suggest that the distribution of T_{MSTDT}, under either the null hypothesis of no linkage or the null hypothesis of no linkage or no association, is very nearly χ^2 with $m - 1$

df. We have explored a wide variety of models by use of computer simulations (nine of which are given here) and have found this to be true. To avoid confusion, we will use "MC − T_{MSTDT}" and "$\chi^2 - T_{\text{MSTDT}}$" to denote whether the significance of the test is estimated by use of the Monte Carlo permutation procedure or by use of χ^2 critical values.

If we are testing for linkage and association by using a biallelic marker, then all four tests are equivalent. The square of Z_c and the S-TDT statistic are equal to T_{MSTDT}. Furthermore, the permutation procedure using the S-TDT statistic is equivalent to use of the DAT statistic. This can be seen from the following (for the derivation, see the Appendix):

$$\text{AC}_2 = 2\left(\frac{\sum_{k=1}^{N_S} V_{1k}}{\sum_{k=1}^{N_S} V_{1k} - N_2/2}\right) Z_1^2 \ .$$

Since the coefficient of Z_1^2 is invariant under the permutation procedure, the S-TDT and DAT are equivalent. For multiallele markers, the four tests are not equivalent, and thus our attention will focus on comparing their performance for such markers.

It may be possible to collect parental marker genotypes for some sibships. Curtis (1997) and Spielman and Ewens (1998) extended their statistics for this situation. Curtis added the number of transmissions of marker allele M_i to affected individuals from heterozygous M_iM_j parents to the T_{ij} from families without parental information. These combined counts are used in the same framework as above. Spielman and Ewens based their statistic on the sum of Y_i calculated for the N_S families that do not have parental marker genotypes and X_i, the number of transmissions of allele M_i to affected individuals from a total of N_{hi} transmissions to affected individuals from parents heterozygous for marker allele M_i. The Z_i can now be written:

$$\tilde{Z}_i = \frac{(Y_i + X_i) - (\sum_{k=1}^{N_S} A_{ik} + n_{hi}/2)}{\sqrt{\sum_{k=1}^{N_S} V_{ik} + n_{hi}/4}} \ .$$

For a biallelic marker, the statistic is $\tilde{Z}_1$; for a multiallele marker, the statistic is $\tilde{Z}_{\max} = max_{i=1\ldots m} |\tilde{Z}_i|$. The combined statistic that we propose is

$$T_{\text{mcomb}} = \frac{m-1}{m} \sum_{i=1}^{m} \tilde{Z}_i^2 \ .$$

If only one of the parental marker genotypes is known, then the recommendations given by Curtis and Sham (1995) should be used to determine whether the transmission information can be used without introduction of bias. If it cannot, then the sibship information should be used instead. This applies to both the Curtis (1997) and Spielman and Ewens (1998) statistics, as well as to T_{mcomb}.

To simplify notation, we will refer to each test in terms of the test statistic that is used to measure its significance—for example, for a multiallelic marker, the Curtis test is the LR test, the Boehnke and Langefeld test is the AC_2 test, the Spielman and Ewens test is the $Z_{\max}$ test, and our test is the T_{MSTDT} test.

Simulation Parameters

For our simulation studies, we define three complex-disease models. The disease prevalence is .05, and the disease allele has a frequency of $p = .2$. The three disease models correspond to a dominant, recessive, and additive disease allele; for each model, phenocopies were introduced to provide an attributable risk of .7; for the three models, this yielded penetrances, $\{f_{11},f_{12},f_{22}\}$, of {.112,.112,.015}, {.89,.015,.015}, and {.19,.1025,.015}, respectively. Setting disease prevalence, disease-allele frequency, the mode of inheritance, and attributable risk allowed the penetrances to be uniquely determined. We felt that an investigator would be able to assign values to these parameters. We also investigated disease models with a prevalence of .0005, a disease-allele frequency of .02, and an attributable risk of .7, as well as the models given by Boehnke and Langefeld (1998, table 2). Results for all of these models are comparable to those presented in the present study.

Three different six-allele markers are considered, and the population allele frequencies are given in table 1. Following Boehnke and Langefeld (1998), we allow marker allele M_1 to be positively associated with the disease allele. Association is parameterized by C, which equals the difference between the probability, $\Pr(M_1 \mid \text{affected})$, that a randomly chosen affected individual carries an M_1 allele and the probability, $\Pr(M_1 \mid \text{unaffected})$, that a randomly chosen unaffected individual carries an M_1 allele. For marker alleles $M_2,\ldots,M_m$, we assume that

$$\Pr(M_i \mid \text{affected}) - \Pr(M_i \mid \text{unaffected}) = -C\frac{\Pr(M_i)}{\sum_{i \geq 2} \Pr(M_i)} \ .$$

Table 1

Marker Frequency Distributions

MARKER (NUMBER[a])	$\Pr(M_i)$, FOR $i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$	$i = 6$
Unimodal (1)	.500	.100	.100	.100	.100	.100
Bimodal (2)	.300	.300	.100	.100	.100	.100
Uniform (3)	.167	.167	.167	.167	.167	.167

[a] For marker-frequency distributions.

Table 2

Comparison of Power—Z_{max} versus T_{MSTDT}

Conditional Marker-Allele Distribution and Penetrance[a]	N_S	Power of Z_{max}	Power of T_{MSTDT}
A:			
Incomplete	500	.80	.77
Complete	40	.86	.80
B:			
Incomplete	500	.54	.82
Complete	40	.47	.78

[a] For *A*, $\{\Pr(M_i \mid D_1)\}_{i=1}^{6} = \{.7,.0333,\ldots,.0333\}$, and $\{\Pr(M_i \mid D_2)\}_{i=1}^{6} = \{.35,.3833,.0333,\ldots,.0333\}$; for *B*, $\{\Pr(M_i \mid D_1)\}_{i=1}^{6} = \{.195,.005,\ldots,.195,.005\}$, and $\{\Pr(M_i \mid D_2)\}_{i=1}^{6} = \{.1,\ldots,.1\}$.

For given values of C and θ, these equations determine the conditional marker-allele probabilities $\Pr(M_i \mid D_1)$ and $\Pr(M_i \mid D_2)$, which can be used to calculate the distribution of parental haplotypes in a family ascertained on the basis of an affected child (Kaplan et al. 1997). Once the parental haplotypes are determined, the genotype and affection status of each of the additional siblings can be generated in the standard way. Only sibships that contain at least one unaffected sib, in addition to the affected sib on the basis of whom the family was ascertained, are included in the sample. For given values of θ and C, there are nine marker/disease models. These population models will be denoted by the marker-distribution number (table 1) and "D," "R," or "A" for the disease model (e.g., population 1A represents a unimodal marker distribution and an additive disease allele). We use a significance level of $\alpha = .05$ for all tests.

A 10-allele marker was used to compare the power of the Z_{max} and T_{MSTDT} tests. Two conditional marker distributions were selected to demonstrate how the relative power of the tests depends on how the association is distributed among the marker alleles. One of the conditional marker distributions has the association concentrated on 2 of the 10 alleles, and the other has the association spread among all of the marker alleles (table 2).

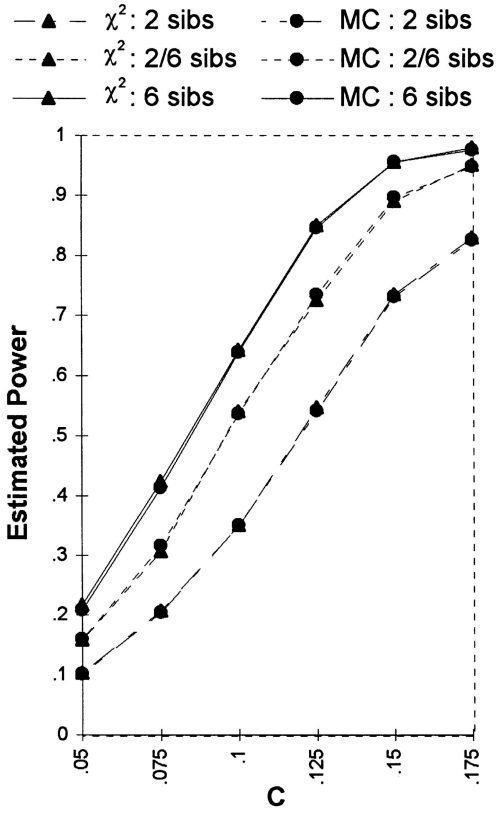

Figure 1 Estimates of power for $\chi^2 - T_{MSTDT}$ and MC $- T_{MSTDT}$, for three different sampling schemes under the 1-D genetic model. $N_S = 200$. The sampling schemes correspond to sampling 100% sibships of size 2, 50% sibships of size 2 and 50% sibships of size 6, and 100% sibships of size 6.

Results

χ^2 Approximation for T_{MSTDT} and T_{mcomb}

We provide evidence that the $\chi^2 - T_{MSTDT}$ test is a valid test both for linkage and association in sibships with minimal configuration and for linkage in larger sibships. Table 3 contains estimates of significance values for the nine disease/marker models corresponding to $C = .15$ and $\theta = .5$. Estimates are shown for sample sizes of 25, 50, and 100 sibships, as well as for sibship sizes of 2, 4, and 6. For small data sets, the $\chi^2 - T_{MSTDT}$ test was conservative, and the permutation procedure should be used. This is a result of sparseness in the $2 \times m$ contingency table of marker-allele counts in affected and unaffected individuals. We found similar results for sibships with minimal configuration for $C = 0$ and $\theta = 0$ (data not shown). We also examined estimates of significance for the T_{mcomb} test, by using χ^2 critical values, and obtained results analogous to those for the T_{MSTDT} test. Estimates were computed for different values of q, the proportion of the sibships having parental information. As for the T_{MSTDT} test, all estimates of significance estimated by use of the χ^2 distribution were reasonably accurate, with the exception of $N_S = 25$. With such a small sample, the permutation procedure should be used (data not shown).

The simulations support the validity of the $\chi^2 - T_{MSTDT}$ test as both a test of linkage and association and a test of linkage. Thus, powers for the $\chi^2 - T_{MSTDT}$ and MC $- T_{MSTDT}$ tests will also be close. To demonstrate this and to explore how power changes for different size sibships and for different values of C, we compared the powers of the $\chi^2 - T_{MSTDT}$ and MC $- T_{MSTDT}$ tests for the

Table 3

Estimates of Significance Level for the $\chi^2 - T_{MSTDT}$ Test

	Significance for[a]								
	$N_S = 100$			$N_S = 50$			$N_S = 25$		
Model	$t = 2$	$t = 4$	$t = 6$	$t = 2$	$t = 4$	$t = 6$	$t = 2$	$t = 4$	$t = 6$
Unimodal:									
Dominant	.050	.051	.047	.039	.051	.048	.027	.045	.051
Recessive	.042	.048	.054	.041	.050	.048	.026	.046	.050
Additive	.049	.050	.052	.044	.049	.050	.028	.045	.046
Bimodal:									
Dominant	.048	.050	.051	.041	.044	.056	.028	.044	.047
Recessive	.044	.047	.049	.041	.049	.049	.030	.046	.049
Additive	.049	.047	.055	.039	.048	.050	.027	.045	.049
Uniform:									
Dominant	.047	.046	.051	.042	.050	.047	.029	.047	.045
Recessive	.047	.051	.049	.040	.049	.050	.035	.046	.046
Additive	.045	.050	.050	.045	.047	.048	.029	.046	.044

[a] N_S is no of sibships sampled, of size t.

following example. We use the model with a dominant disease allele and unimodal marker distribution with $\theta = .02$ and $N_S = 200$. Figure 1 contains power estimates for three different sampling schemes. The sampling schemes differ on the basis of the size of the sibships sampled. Either all sibships of size 2 or 6 were sampled or a mixture of 50% sibships of size 2 and 50% sibships of size 6 were sampled. In all three cases, the powers estimated by use of the χ^2 critical values are close to the Monte Carlo estimates of power. We have found this to be true for other values of α as well (data not shown).

As expected, the power of the test increased monotonically with C. Furthermore, the power of the test increased with the size of the sibship. One reason for this is that, for larger sibships, more of the sibships are informative. If p_i denotes the probability that a randomly chosen sibship of size i is informative, then, for the model in figure 1 with $C = .15$, p_2 and p_6 are estimated, by simulation, as being .54 and .86, respectively. It follows that, for samples of 200 sibships of size 2, there are, on average, 108 informative sibships, versus 172 informative sibships for sibships of size 6; the respective estimated powers of the tests are .74 and .96. When the number of sampled sibships of size 2 was increased to 319 (200 × .86/.54), the power estimate of the test increased to .91. Therefore, if the sample with sibships of size 2 is enlarged so that there are, on average, the same number of informative sibships as are present in the sample of sibships of size 6, then much of the difference, in estimated power, between the tests is explained.

The distribution of T_{MSTDT}, under either the null hypothesis of no linkage (sibships of arbitrary size) or the null hypothesis of no linkage or no association (sibships of only minimal configuration), does not converge to a χ^2 distribution as the number of sibships increases; however, our results in table 3 suggest that, for a random-mating population, significance level and power can be approximated by use of χ^2 critical values. We also examined the χ^2 approximation for a population consisting of two different strata, each with the unimodal marker-frequency distribution given in table 1. We assume that marker allele M_1 is the most frequent in stratum 1 and that marker allele M_2 is the most frequent in stratum 2. Our samples were chosen so that one-half the sibships came from each stratum. For a dominant disease allele with $C = .15$ and $\theta = .5$, estimates of significance levels were again very close to the nominal level, except for the case of $N_S = 25$, for which the test was too conservative. The power estimates determined by use of the χ^2 approximation also were close to the power of the Monte Carlo test ($\theta = .02$). For example, the sampling scheme with 50% sibships of size 2 and 50% sibships of size 6 yielded estimates of .72 and .73 for the $\chi^2 - T_{MSTDT}$ and MC $- T_{MSTDT}$ tests, respectively. Similar results were found when 90% of the sibships came from stratum 1 and 10% of the sibships came from stratum 2 (data not shown).

Tests of Linkage and Association

We examined the powers of the four tests of the null hypothesis of no linkage or no association; that is, our data consisted only of sibships with minimal configuration. For such samples, the χ^2 approximation for the distribution of LR under the null hypothesis of no linkage or no association can be poor. We computed estimates of the significance level (by setting $\theta = .5$) for the LR test for models 1D, 1R, and 1A for samples consisting of 200 and 600 sibships. Regardless of the number of sibships sampled, the estimates based on sibships with minimal configuration are conservative, whereas

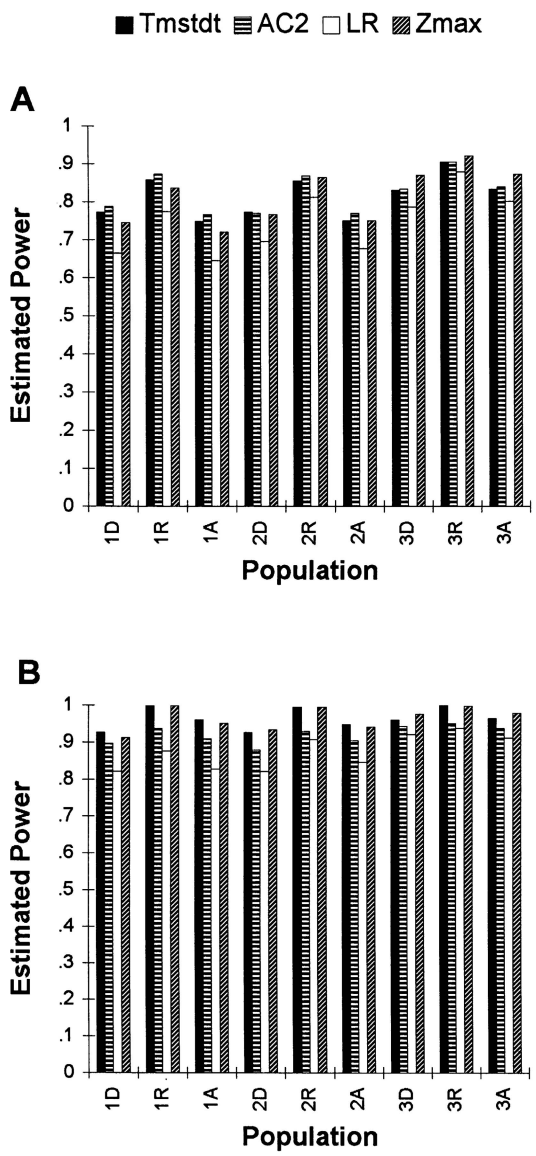

Figure 2 *A*, Power estimates for tests of linkage and association (i.e., sibships containing exactly one affected and one unaffected individual). *B*, Power estimates for 200 sibships of size 5. Z_{max} and T_{MSTDT} use all five sibs, whereas the LR and AC_2 use a random affected individual and the unaffected individual whose marker genotype is maximally different from that of the affected individual. All estimates were computed by use of the Monte-Carlo permutation procedure.

the estimates improve when the maximally different pair of sibs from sibships of size 5 are used. For example, the estimates of significance for the dominant disease model were .031 and .048 for $N_S = 200$, for sibships of size 2 and size 5, respectively. Because of the poor χ^2 approximation for the LR for samples with minimal configuration, we estimated the powers of the four tests (fig. 2*A*) by using the Monte Carlo permutation procedure. The results given are for $C = .15$ and $\theta = 0$. For all nine disease/marker models, the powers of the AC_2, Z_{max}, and T_{MSTDT} tests were comparable, whereas the LR test was consistently less powerful. We also compared the powers of the four tests for samples of size 5, using Curtis's strategy to reduce the sibships to minimal configuration, and found analogous results (data not shown).

The powers of the Z_{max} and T_{MSTDT} tests were compared for two special cases for a marker with 10 alleles. The cases differ depending on whether the association between the marker and disease locus is concentrated on only a few of the marker alleles or is spread over all of the marker alleles. The dominant disease model was used. Table 2 describes each of the models and gives the power for each case. Values of N_S were chosen so that the powers of the tests were ~.8. The tests were also compared by use of a completely penetrant dominant disease allele. If the association is spread over the marker alleles, then the Z_{max} test performs poorly compared with the T_{MSTDT} test; however, if the association is concentrated on only a few of the alleles, then the T_{MSTDT} test is only somewhat less powerful than the Z_{max} test.

Tests of Linkage

The powers of the Z_{max} and T_{MSTDT} tests as tests of linkage were estimated for samples of sibships of size 5 and were compared with estimates of the power of the AC_2 and LR tests, by use of the reduced sibships produced by the Curtis (1997) strategy (fig. 2*B*). For these simulations, $\theta = .02$ and $C = .15$. This value of C was chosen so that the tests would have high power for a sample of 200 sibships. The Z_{max} and T_{MSTDT} tests had comparable power, but the AC_2 and LR tests were consistently less powerful, with the AC_2 test always more powerful than the LR test. The drop in power with use of only two of the five sibs was not great for the AC_2 test. For example, for the model with a dominant disease allele and a unimodal marker distribution, the Z_{max} and T_{MSTDT} tests had powers .91 and .93, respectively, and the AC_2 and LR tests had powers .90 and .82, respectively. We expect a minimal loss in power when the size of the sibships is reduced when the penetrances are low, since unaffected sibs could carry a disease allele. Alternatively, if the penetrances are high, then unaffected sibs are less likely to carry a disease allele, and reducing the sibship should result in a much larger loss of power.

The Combined Test Statistic: T_{mcomb}

For late-onset diseases, the typical data set consists of sibships without parents. A question that one might ask is, How much (if any) power is lost by not having parental data, or, alternatively, how much power is gained by having some parental data? To investigate these questions, we simulated samples of 200 families, each containing one affected and one unaffected sib. For a frac-

tion, q, of the families, we recorded the genotypes of both of the parents, as well as those of all of the sibs; for the remaining fraction of the families, we recorded the genotypes of only the sibs. First, we used the complete data set and calculated T_{mcomb}. We then ignored all of the parental genotypes and computed T_{MSTDT}. Last, we ignored the unaffected sib and calculated a TDT statistic for those families with parental information. Since we used a multiple-allele marker, we used the statistic T_{mhet}, described by Spielman and Ewens (1996). We give results for the unimodal marker distribution with $C = .15$ and $\theta = .02$, for all of the disease models (fig. 3). We compared the powers of the three tests as the proportion, q, of the sibships with parental information changes. It is interesting to note that the T_{mcomb} test is always more powerful than the other two tests, regardless of the composition of the data set, suggesting that it is always possible to have a more powerful test when parental data are collected. Furthermore, if the fraction of the sample for which the data set has parental information is large enough, then ignoring the unaffected sib and using the TDT for the families with parental data will be more powerful. However, the value of q at which the T_{mhet} test becomes more powerful than the T_{MSTDT} test will depend on the disease/marker model.

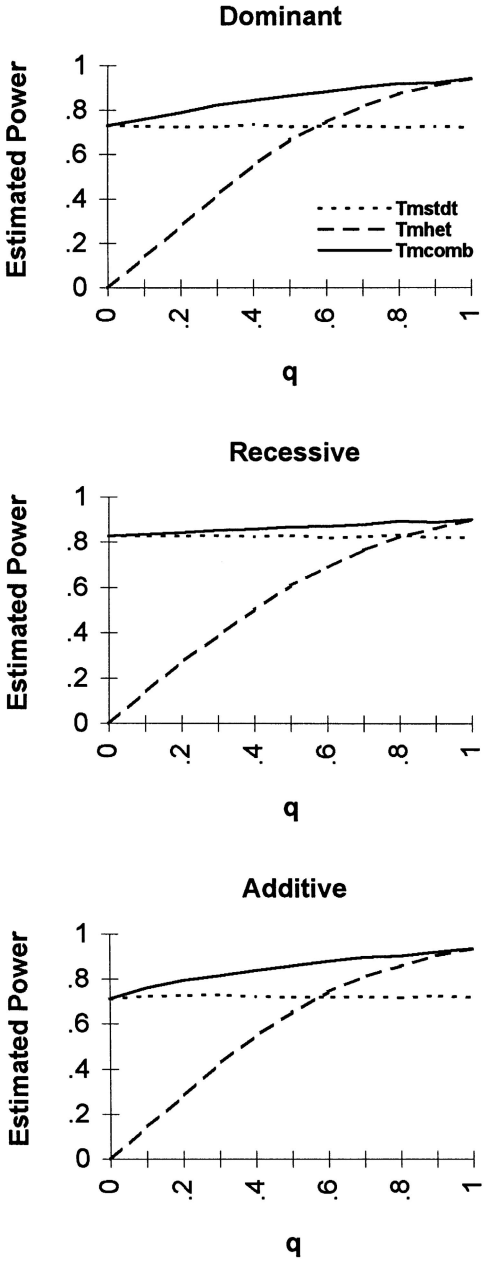

Figure 3 Comparison of T_{mhet}, T_{MSTDT}, and T_{mcomb}, for the unimodal marker distribution; q is the proportion of sibships that have parental information. The χ^2 approximation was used with $N_S = 200$.

Discussion

Although family-based tests have proved to be powerful tools in the search for genes involved in complex diseases, they have, until recently, required knowledge of parental marker genotypes. The introduction of sibship tests by Curtis (1997), Spielman and Ewens (1998), and Boehnke and Langefeld (1998) has brought about a promising new group of tests for both linkage and association and linkage. Like the TDT, these tests maintain validity even in a stratified population and yet do not require parental marker genotypes, which, for late-onset diseases, can be impractical to obtain. As is true for the introduction of any group of statistics, a number of questions arise. We have presented simulations that help answer some of these questions in the context of a complex disease.

For biallelic markers and samples consisting of sibships with minimal configuration, the Z_c, AC_2, Z_1, and T_{MSTDT} tests are equivalent. However, for a marker with more than two alleles, the LR test as a test of linkage and association was less powerful than the other tests, and therefore we would not recommend its use. The other three tests were comparable in power even if the sample of sibships of minimal configuration was reduced from larger sibships by use of Curtis's strategy. The Monte Carlo permutation procedure can be used to perform any of the three tests, and, for small samples, we would recommend its use. Our results suggest that χ^2 critical values can be used to determine significance for the T_{MSTDT} test. This is an attractive feature, especially if a large number of markers are to be tested; however, a conservative strategy would be to also perform the Monte Carlo permutation procedure on all significant findings, to guard against false-positive results.

For tests of linkage, our results suggest that it is best to use all sibs rather than to reduce the size of the sibships

by use of Curtis's strategy. Only the Z_{max} and T_{MSTDT} tests use all siblings and result in a valid test of linkage. These tests differ in how they detect linkage, with Z_{max} using a maximum allelic deviation and with T_{MSTDT} summing deviations over all alleles for a global test. Hence, the pattern of association between marker alleles and the disease alleles will impact the relative performance of the tests. If the association is concentrated on a small number of the marker alleles, then the Z_{max} test can be more powerful; alternatively, if the association is spread among many marker alleles, then the T_{MSTDT} test can be more powerful. The powers of the tests depend on the number of informative sibships sampled. If larger sibships are sampled, then a larger proportion of the sibships will be informative; however, depending on the disease under study, it may be more economical to sample more smaller sibships than to sample fewer large sibships.

It is expected that tests utilizing available parental marker data—that is, tests such as the TDT—will be more powerful than tests based only on sibships (Spielman and Ewens 1998). We find that the extent of this difference will depend on the disease/marker model. The idea proposed by Curtis (1997) and Spielman and Ewens (1998)—that is, to combine these two types of data—allows maximal use of a data set, and our results support this claim. The combined tests reduce to a test utilizing parental marker genotypes if all sibships have parental information, and they reduce to a sibship test if no parental information is available. Realistically, some of the families will have genotypic information for only one of the parents. In this case, the recommendations of Curtis and Sham (1995) should be used to determine whether the transmission information can be used without introduction of bias; if it cannot, then the sibship information can be used instead. Alternatively, the missing parental genotypes could be inferred; however, this would require estimation of marker-allele frequencies and so would be subject to stratification problems.

The results presented in the present study are meant to both summarize and compare the sibship tests that are currently available. These tests provide a promising new area of family-based tests, particularly for late-onset diseases. Our results have been presented for samples of unrelated sibships. If any of the sibships are related, then the tests are valid only as tests of linkage (Spielman and Ewens 1998).

Acknowledgments

This work was supported in part by National Institutes of Health grant GM45344. This material is based on work supported under a National Science Foundation graduate fellowship to S.A.M. We are grateful to Dr. Eden Martin, Duke University Center for Human Genetics, for numerous discussions.

Appendix

Relationship of AC_2 and Z_1, for a Biallelic Marker, for Sibships with Minimal Configuration

For a biallelic marker, there are six sibship types with minimal configuration (table A1). The total number of each type in the data set will be denoted "N_{ij}," where i represents the number of M_1 marker alleles in the affected individual and where j represents the number of M_1 marker alleles in the unaffected sibling. Boehnke and Langefeld's statistic, AC_2, can be written in terms of N_{ij}, because

$$n_{11} = n_{22} = N_{21} + 2N_{20} + N_{10} \tag{A1}$$

and

$$n_{12} = n_{21} = N_{12} + 2N_{02} + N_{01} \ . \tag{A2}$$

Spielman and Ewens' statistic, Z_1, requires calculation of the theoretical permutation mean and variance for each sibship's contribution to Y_1. Given sibship i with t_i sibs, of which a_i are affected and u_i are unaffected, let r_i be the number of sibs with genotype M_1M_1 and let s_i be the number of sibs with genotype M_1M_2; then

$$A_{1i} = (2r_i + s_i)\frac{a_i}{t_i} \tag{A3}$$

and

$$V_{1i} = a_i u_i \frac{4r_i(t_i - r_i - s_i) + s_i(t_i - s_i)}{t_i^2(t_i - 1)} \ . \tag{A4}$$

For sibships with minimal configuration, $a_i = u_i = 1$ and $t_i = 2$, so that, on the basis of equations (A3) and (A4),

Table A1

Types of Sibships with Minimal Configuration, for a Biallelic Marker

Sibship Type	Genotype of Sib		No. in Data Set
	Affected	Unaffected	
1	M_1M_1	M_1M_2	N_{21}
2	M_1M_1	M_2M_2	N_{20}
3	M_1M_2	M_1M_1	N_{12}
4	M_1M_2	M_2M_2	N_{10}
5	M_2M_2	M_1M_1	N_{02}
6	M_2M_2	M_1M_2	N_{01}

$$\sum_i A_{1i} = (N_{21} + N_{12})\left(\frac{3}{2}\right) + (N_{20} + N_{02}) \times(1) + (N_{10} + N_{01})\left(\frac{1}{2}\right)$$

and

$$\sum_i V_{1i} = (N_{21} + N_{12})\left(\frac{1}{4}\right) + (N_{20} + N_{02}) \times(1) + (N_{10} + N_{01})\left(\frac{1}{4}\right) .$$

Given $Y_1 = 2(N_{21} + N_{20}) + N_{12} + N_{10}$ and the relationships from equations (A1) and (A2),

$$Y_1 - \sum_i A_{1i} = \frac{n_{11}}{2} - \frac{n_{21}}{2}$$

and

$$\sum_i V_{1i} = \frac{n_{11}}{4} + \frac{n_{21}}{4} + \frac{N_2}{2} ,$$

where N_2 is defined, by Curtis, as the number of sibships of type 2 or type 5. Now AC_2 and Z_1 can be related:

$$\begin{aligned} AC_2 &= \sum_{i=1}^{2} \frac{(n_{1i} - n_{2i})^2}{n_{1i} + n_{2i}} \\ &= 2\left[\frac{(n_{11} - n_{21})^2}{n_{11} + n_{21}}\right] \qquad \text{(from [A1] and [A2])} \\ &= 2\left\{\frac{[2(Y_1 - \Sigma_i A_{1i})]^2}{n_{11} + n_{21}}\right\} \qquad \text{(from [A5])} \\ &= 2 \times 4\left[\frac{(Y_1 - \Sigma_i A_{1i})^2}{4\Sigma\, V_{1i} - 2N_2}\right] \qquad \text{(from [A6])} \\ &= 2\left[\frac{(Y_1 - \Sigma_i A_{1i})^2}{\Sigma_i V_{1i} - N_2/2}\right] \\ &= 2\left(\frac{\Sigma_i V_{1i}}{\Sigma_i V_{1i} - N_2/2}\right) z_1^2 . \end{aligned}$$

Hence, AC_2 and Z_1 are related by

$$AC_2 = 2\left(\frac{\Sigma_i V_{1i}}{\Sigma_i V_{1i} - N_2/2}\right) Z_1^2 .$$

References

Boehnke M, Langefeld CD (1998) Genetic association mapping based on discordant sib pairs: the discordant-alleles test. Am J Hum Genet 62:950–961

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 56:811–812

Kaplan NL, Martin ER, Weir BS (1997) Power studies for the transmission/disequilibrium tests with multiple alleles. Am J Hum Genet 60:691–702

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

Ott J (1991) Analysis of human genetic linkage, rev ed. Johns Hopkins University, Baltimore

Sham PC, Curtis D (1995) An extended transmission/disequilibrium test (TDT) for multiallele 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

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