Genetic Epidemiology 19:18–29 (2000)

Circumventing Multiple Testing:
A Multilocus Monte Carlo Approach to Testing for Association

Lauren M. McIntyre,[1,2,4]* Eden R. Martin,[3] Katy L. Simonsen,[5] and Norman L. Kaplan[6]

[1]*Institute for Clinical and Epidemiological Research, Veterans Affairs Medical Center, Durham, North Carolina*
[2]*Division of Biometry, Duke University Medical Center, Durham, North Carolina*
[3]*Department of Human Genetics, Duke University Medical Center, Durham, North Carolina*
[4]*Department of Agronomy, Purdue University, West Lafayette, Indiana*
[5]*Department of Statistics, Purdue University, West Lafayette, Indiana*
[6]*National Institute for Environmental Health Sciences, Research Triangle Park, North Carolina*

Advances in marker technology have made a dense marker map a reality. If each marker is considered separately, and separate tests for association with a disease gene are performed, then multiple testing becomes an issue. A common solution uses a Bonferroni correction to account for multiple tests performed. However, with dense marker maps, neighboring markers are tightly linked and may have associated alleles; thus tests at nearby marker loci may not be independent. When alleles at different marker loci are associated, the Bonferroni correction may lead to a conservative test, and hence a power loss. As an alternative, for tests of association that use family data, we propose a Monte Carlo procedure that provides a global assessment of significance. We examine the case of tightly linked markers with varying amounts of association between them. Using computer simulations, we study a family-based test for association (the transmission/disequilibrium test), and compare its power when either the Bonferroni or Monte Carlo

Contract grant sponsor: NCI; Contract grant number: 5PO1CA72099; Contract grant sponsor: NIH; Contract grant numbers: GM45344, AG16996; Contract grant sponsor: NSF; Contract grant number: DB1-9808026; Contract grant sponsor: Agriculture Experiment Station, Purdue University.

*Correspondence to: Lauren M. McIntyre, Department of Agronomy, 1150 Lilly Hall of Life Sciences, Purdue University, West Lafayette, IN 47907. E-mail: mcintyre@dept.agry.purdue.edu

Received 8 January 1999; Revision accepted 3 August 1999

procedure is used to determine significance. Our results show that when the alleles at different marker loci are not associated, using either procedure results in tests with similar power. However, when alleles at linked markers are associated, the test using the Monte Carlo procedure is more powerful than the test using the Bonferroni procedure. This proposed Monte Carlo procedure can be applied whenever it is suspected that markers examined have high amounts of association, or as a general approach to ensure appropriate significance levels and optimal power. Genet. Epidemiol. 19:18–29, 2000.

Key words: Monte Carlo method; association; Bonferroni correction

INTRODUCTION

Current molecular techniques have advanced to such a level that it is possible to genotype individuals at several hundred loci spaced throughout the genome. Genome scans of this density are used to identify regions that are linked to a disease. However, often these regions are too large to make a detailed molecular examination of the entire region feasible. One strategy for narrowing a candidate region is to carry out an association study using more densely spaced markers that saturate the region. A dense map with 2,227 single nucleotide polymorphisms (SNP's) is already available [Wang et al., 1998]. It is anticipated that within 18 months a standard set of 100,000 SNP's will exist [Marshall, 1998] making association testing across the whole genome feasible. In either scenario, a dense set of markers is used to test for association between each of the marker alleles and the disease, and multiple testing occurs. The alleles at tightly linked marker loci may themselves be associated and, therefore, the tests conducted at each locus may not be independent, making a Bonferroni correction overly conservative.

Studies to identify the association between marker and disease alleles can be designed to collect either population-based case-control data or simplex family data (i.e., a mother, father, and an affected child). A population-based case-control design is prone to false positives caused by inappropriately matched cases and controls, which can occur if, for example, the population is stratified. Family-based studies avoid this problem, but parental data may be difficult to obtain, especially for late-onset diseases. Whichever design is selected, tests for association are conducted at multiple marker loci and this multiple testing must be accounted for when reporting results.

In this paper we focus on simplex family data. We do not assume that haplotypes are known, but we do assume that a full set of genotypic information is available for each family, i.e., all markers are genotyped for all members of each family. For these data it is reasonable to use the transmission/disequilibrium test (TDT) [Spielman et al., 1993] to test for association, and thus we will focus on this test. However, the methodology developed here is more widely applicable.

METHODS

Transmission/Disequilibrium Test

The TDT was introduced by Spielman et al. [1993] as a test for linkage, in the presence of association, between a single biallelic marker and a disease locus in

nuclear families. For families with at least one affected child, this test compares the number of times that a parent, heterozygous at the marker locus, transmits a particular allele to an affected child, to the number of times that a heterozygous parent does not transmit that allele to an affected child. When these numbers are significantly different, the null hypothesis of no linkage is rejected. If one affected child from each family is used, then the TDT is also a valid test for association. The TDT has been extended to allow for multiallelic markers [Spielman and Ewens, 1996; Sham and Curtis, 1995; Kaplan et al., 1997]; however, the application of the TDT to multiple linked marker loci has not been fully addressed.

Suppose we wish to test for association at each of l closely linked marker loci in a sample of simplex families. When testing at many loci, the null hypothesis can be defined in two alternate ways. Let H_{01} be the null hypothesis of no association between alleles at *any* of the marker loci and a disease allele. Let H_{02} be the null hypothesis of no association between any haplotype and a disease allele. For example, suppose we have two markers M, N with alleles M_1, M_2 and N_1, N_2 and disease locus D with alleles D_1, D_2. Under H_{01}, $Pr(M_i|D_r) = Pr(M_i)$ and $Pr(N_j|D_r) = Pr(N_j)$ for all i,j,r. Alternatively, under H_{02}, $Pr(M_iN_j|D_r) = Pr(M_iN_j)$ for all i,j,r. If the haplotypes are not associated with a disease allele, then the individual marker alleles are also not associated with the disease allele and so H_{02} implies H_{01}. However, there are cases where the individual marker alleles may not be associated with the disease allele but the haplotypes are associated with the disease allele; therefore, H_{01} does not imply H_{02}.

When examining a number of markers it is natural to conduct the test at each marker locus individually and to look at the smallest P value first. The type I error α_T is the experimentwise type I error. Such an error occurs if at least one of the single-marker tests is significant and the null hypothesis is true. It is this experimentwise type I error that we want to control. If the test statistics have the same asymptotic distribution under the null (which would occur when all markers are biallelic as in the case of SNP's), then the test for detecting linkage based on the minimum P value is equivalent to the test based on the maximum value of the test statistics. In this paper, we examine only biallelic markers; therefore, we study the maximum TDT statistic [Spielman et al., 1993] which we denote by TDT_{MAX}. If each of the tests are independent then the distribution of TDT_{MAX} can be defined explicitly as the maximum order statistic, and the experimentwise type I error can be easily controlled. The maximum order statistic for a series of l statistics that are independent and identically distributed has a distribution of

$$lf_x(x)[F_x(x)]^{l-1}, \tag{1}$$

where $f_x(x)$ is the probability density function and $F_x(x)$ is the cumulative distribution function [Casella and Berger, 1990]. This formulation assumes independence between the individual statistics. The assumption of independence is the issue in this paper. Therefore, we will not pursue the specific distribution of the maximum order statistic but rather will examine closely the independence assumption.

Defining the distribution of the maximum when the tests are correlated is not straightforward. If the markers are loosely linked, then Spielman and Ewens [1996] suggest that the tests are independent. In the Appendix we show that if there is no

association between marker alleles, the tests are independent under H_{02}, but will be correlated if there is association between alleles at the marker loci.

Bonferroni Procedure

It has been suggested that it is appropriate to use a standard Bonferroni correction to control the experimentwise level when applying the TDT at multiple marker loci [Spielman and Ewens, 1996; Risch and Merikangas, 1996]. A Bonferroni correction means that each single marker test is conducted at level $\alpha = \alpha_T/l$, which guarantees an experimentwise type I error of at most α_T. This correction was developed as a conservative approach to controlling experimentwise type I error and becomes increasingly conservative as tests become more correlated. This conservativeness can lead to a loss of power. For example, in the case of a two marker-two allele sytem, when the alleles at the two marker loci are completely correlated only two haplotypes exist (of a possible four haplotypes), and the test statistics for each locus will be identical. In this case, the Bonferroni correction results in an experimentwise type I error of $\alpha_T/2$, rather than the desired type I error of α_T. To illustrate the impact of over-conservativeness, we calculated the power of the test statistic using a Bonferroni correction for two, three, or ten identical tests (each with a known chi-squared distribution under the null hypothesis) for specific noncentrality parameters λ, and with $\alpha_T = 0.05$ (Table I). Comparing the power of the test at $\alpha_T = 0.05$ to the power of the test at $\alpha_T/2$, $\alpha_T/3$ and $\alpha_T/10$ clearly shows that using the Bonferroni procedure can result in a substantial loss of power. However, as the noncentrality parameter λ increases, this difference is less marked when power approaches unity. This indicates that inappropriate use of the Bonferroni procedure will have the most impact in cases where the power is moderate.

Monte Carlo Procedure

Ideally we would like the experimentwise significance level of the test to be α_T, without being overly conservative, in order to increase the power of our test. To achieve this goal we propose a Monte Carlo procedure to determine significance. This procedure guarantees the correct experimentwise level, regardless of dependencies between the individual tests, and regardless of assumptions about the asymptotic properties of the TDT at individual marker loci. The Monte Carlo procedure we describe gives the empirical distribution of TDT_{MAX} under the null hypothesis H_{02}, taking into account the fact that we have calculated l statistics and chosen the maximum. If the alleles at linked markers are not associated with each other, then the result in the Appendix shows that the tests are independent under H_{02}, and the empirical distribution of TDT_{MAX} estimated using the Monte Carlo procedure will give

TABLE I. Theoretical Power of a Chi-Squared Test Conducted at Different Significance Levels for a Various Noncentrality Parameters λ

Significance level	Power ($\lambda = 3.96$)	Power ($\lambda = 5.24$)	Power ($\lambda = 10.94$)	Power ($\lambda = 15.16$)
$\alpha_T = 0.05$	0.51	0.63	0.91	0.97
$\alpha_T/2 = 0.025$	0.40	0.52	0.86	0.95
$\alpha_T/3 = 0.017$	0.34	0.49	0.82	0.93
$\alpha_T/10 = 0.005$	0.17	0.30	0.69	0.86

results similar ot the Bonferroni procedure. While we examine this procedure for the specific case of the *TDT*, it is broadly applicable and can be used in any simplex-family-based test.

The procedure is analogous to the Monte Carlo test for the TDT at a single marker proposed by Kaplan et al. [1997]. In that procedure, the pair of alleles transmitted to an affected child and the pair not transmitted are considered. When there is no association between alleles at the marker locus and alleles at the disease locus, then either pair is equally likely to have been transmitted. This means that the labels "transmitted" and "not transmitted" can be permuted for each affected child in the sample, thereby generating the empirical sampling distribution of the data under the null hypothesis of no association between the alleles at the marker locus and the alleles at the disease locus. Similarly, for a collection of loci, there is a set of alleles across all loci that is transmitted by the parents and a set of alleles across all loci that is not transmitted by the parents. In this case, under the null hypothesis H_{02}, the labels "transmitted" and "not transmitted" can be permuted for the sets of alleles, and this is the basis of the multilocus Monte Carlo procedure we propose. This multilocus approach will take into account any correlations between marker loci. The steps for this multilocus Monte Carlo procedure are as follows:

1. Calculate TDT_{MAX}.
2. For each family, permute the "transmitted" and "not transmitted" labels randomly.
3. Calculate TDT_{MAX} for the permuted data.
4. If TDT_{MAX} for the permuted data is larger than the value of TDT_{MAX} from the original data, count 1; otherwise count 0.
5. Repeat steps 2, 3 and 4 k times.
6. Estimate the p-value by $\hat{p}$, the total count from step 4 divided by the total number of shuffles k.
7. Reject the null hypothesis if $\hat{p} \leq \alpha_T$.

In the above procedure, we permuted the labels "transmitted" and "not transmitted" for the sets of marker alleles. However, this is not the only valid permutation procedure to test the null hypothesis H_{02}. If we knew marker haplotypes in the child, then the maternal haplotype that is transmitted could be permuted with the maternal haplotype that is not transmitted, and the paternal haplotype could be separately permuted. Such a permutation procedure has been suggested by Lazzeroni and Lange [1998]. Permuting the haplotypes instead of the set of marker alleles may be more powerful in some cases, as additional information is retained by the permutation procedure. However, it depends upon knowing the phase across multiple loci. When both parents and the child are heterozygous at two loci, the phase of the parental haplotypes transmitted is unknown, and thus a permutation procedure dependent upon knowing the haplotype phase would be limiting.

Two-Marker Simulations

For the model to be analytically tractable we adopted the assumptions of Knapp et al. [1993]: no selection or mutation and random mating (and hence Hardy-Weinberg equilibrium) at marker and disease loci. We simulated haplotype frequencies for two

biallelic markers allowing for association between alleles at the marker loci. We assumed no recombination between the markers within the nuclear family, since in a second round genome scan, with closely spaced markers, recombination between markers will be rare. In addition, it is in this scenario of no recombination between markers and association between alleles at marker loci that the difference between the power of the test using the Bonferroni procedure compared to the Monte Carlo procedure will be most noticeable. The disease locus had the same properties for all simulations: two alleles with the disease allele having frequency 0.01, recessive with penetrance 0.99, and no phenocopies. Each marker locus was completely linked to the disease. Association between alleles at each marker locus and the disease allele was defined in terms of I^* where

$$I^*_{MD} = \sum_i \frac{[Pr(M_i|\text{affected}) - Pr(M_i|\text{unaffected})]^2}{Pr(M_i|\text{affected}) + Pr(M_i|\text{unaffected})} \tag{2}$$

[Kaplan et al., 1997].

Varying degrees of association between alleles at the markers were simulated. Association between alleles at the markers was defined as the departure from frequencies expected under independence. If two markers (M, N) are independent, then $Pr(M_iN_j) = Pr(M_i)Pr(N_j)$ for all i and j. Thus departure from independence (linkage disequilibrium) can be measured by

$$C_{MN} = \sum_i \sum_j \frac{(Pr(M_i N_j) - Pr(M_i)Pr(N_j))^2}{Pr(M_i)Pr(N_j)}. \tag{3}$$

When the alleles at the two marker loci are completely associated, $C_{MN} = 1$, and when the alleles at the two marker loci are completely unassociated, $C_{MN} = 0$.

We generated 100 simplex families where parental data were simulated conditional upon having a single affected offspring. This sampling strategy mimics the sampling done in reality. To estimate the power, we simulated data under alternative hypotheses, with association between at least one marker allele and the disease allele to generate haplotype association, and then applied TDT_{MAX}. In the Bonferroni procedure, we compared the value of TDT_{MAX} to a critical value from a χ^2_1 distribution with a significance level of α_T/l. For computational convenience, in the Monte Carlo procedure the evaluation of significance was based upon the procedure described above with $k = 599$. We repeated this data simulation and evaluation 100,000 times and estimated the power of the test by the proportion of times the test statistic rejected the null hypothesis H_{02} at $\alpha_T = 0.05$.

Ten-Marker Simulation

In order to explore power differences between the Bonferroni procedure and the Monte Carlo procedure under a more realistic circumstance, we also simulated ten tightly linked marker loci and a single disease locus. We used a neutral coalescent process to simulate haplotype data for eleven loci spaced 0.1 cM apart [Simonsen and Churchill, 1997; Kaplan and Hudson, 1985; Hudson, 1983]. The coalescent model uses a Markov chain process to simulate an ancestral relationship among haplotypes

for each locus. If there is no recombination ($\theta = 0$) among linked loci, then the ancestral relationship at these loci will be identical. Recombination among linked loci will cause the ancestral relationship to vary; however, linkage ($0 < \theta < ½$) between loci will cause the ancestral relationships to be correlated as a function of recombination rate and population size. Details of this simulation procedure are given by Simonsen and Churchill [1997]. Once the ancestral relationships have been simulated, marker and disease types are assigned to the single common ancestor of all the haplotypes, and a mutational process is used to generate the present-day polymorphism among haplotypes. We constrained the mutational process to generate ten biallelic markers with frequencies between 0.25 and 0.75. The disease locus was also assumed to be biallelic, with frequency between 0.10 and 0.15, and was positioned in the center of the markers. The sample consisted of 2,000 haplotypes, of which 200–300 carried the disease allele for each simulation. The coalescent simulation was performed under the assumption of constant population size (50,000) and no selective effects.

To construct family data we simulated a recessive model where sampling of affected individuals is a criterion for sampling the family, as in Kaplan et al. [1997]. We identified all haplotypes carrying a disease allele, and, to allow for phenocopies, we replaced 5% of the disease haplotypes with non-disease haplotypes. We then sampled 200 times with replacement from this pool of haplotypes to construct 100 affected children. Each parent was then determined by pairing a single haplotype from the child with a second haplotype chosen at random from the complete set of 2,000 haplotypes. In this way, a sample of 100 families with genotypes for affected children and both parents was constructed. Once the familial data were simulated we then calculated the TDT_{MAX} statistic and determined significance at $\alpha_T = 0.05$ according to the Bonferroni procedure and the Monte Carlo procedure and estimated the power of the two procedures. To estimate power, the simulation was performed 9,600 times, including generating the haplotypes and constructing the families.

RESULTS

The powers of the tests for the Monte Carlo procedure and the Bonferroni procedure depend in a complex way on the amount of association between alleles at marker loci and between alleles at marker and disease loci. Even for two marker loci and a single disease locus there is no straightforward way to predict the powers of the tests for the two procedures from the values of C_{MN}, I^*_{MD}, and I^*_{ND}. Hence, to examine the difference between the power of the test for the two procedures we considered several different numerical examples. In Table II four basic patterns of values for C_{MN}, I^*_{MD}, and I^*_{ND} are described. In case 1, I^*_{MD} and I^*_{ND} are both fixed at 0.0198 and different values of C_{MN} are examined, showing that the difference in power decreases as C_{MN} decreases. Case 2 is similar to case 1, except with $I^*_{MD} = I^*_{ND} = 0.0262$, demonstrating that the pattern of results for case 1 hold for different fixed values of $I^*_{MD} = I^*_{ND}$. Case 3 represents the limiting case of complete linkage disequilibrium between markers. In case 3, $C_{MN} = 1$ and different values of $I^*_{MD} = I^*_{ND}$ are examined. Taken with the $C_{MN} = 1$ examples from cases 1 and 2, these results show that the largest power difference occurs with the maximum level of association. These powers closely correspond to the theoretical power of the two procedures predicted

TABLE II. Population Level Association Between Alleles at Two Markers *M* and *N* is C_{MN}

Case	C_{MN}	I^*_{MD}	I^*_{ND}	$\widehat{\text{power}}_B$	$\widehat{\text{power}}_{MC}$	Difference	Percent gain
1	1	0.0198	0.0198	0.40	0.50	0.10	25.0
	0.637	0.0198	0.0198	0.50	0.53	0.03	6.0
	0.355	0.198	0.0198	0.54	0.56	0.02	3.7
2	1	0.0262	0.0262	0.52	0.62	0.10	19.2
	0.896	0.0262	0.0262	0.57	0.63	0.06	10.5
	0.217	0.0262	0.0262	0.69	0.69	0.00	0.00
3	1	0.0386	0.0386	0.70	0.79	0.09	12.9
	1	0.0547	0.0547	0.87	0.92	0.05	5.7
	1	0.0758	0.0758	0.97	0.99	0.02	2.1
4	0.938	0.000113	0.000111	0.0392	0.0616	0.0224	57.1
	0.917	0.00580	0.000881	0.123	0.17	0.047	38.2
	0.870	0.00782	0.00123	0.156	0.201	0.045	28.8
	0.821	0.0039	0.000209	0.111	0.138	0.027	24.3
	0.904	0.00956	0.0132	0.273	0.336	0.063	23.1
	0.908	0.0146	0.0116	0.309	0.369	0.06	19.4
	0.911	0.0258	0.0120	0.44	0.50	0.06	13.6
	0.865	0.00775	0.0210	0.428	0.481	0.053	12.4

*The levels of association between the marker and disease alleles are I^*_{MD} and I^*_{ND}. The power for TDT_{MAX} with the Bonferroni procedure is $\widehat{\text{power}}_B$, and with the Monte Carlo procedure is $\widehat{\text{power}}_{MC}$. The next column shows the absolute difference in the power. The last column shows the percent power gained by the Monte Carlo procedure. These results are for 100 simplex families with complete genotypic information with $\alpha_T = 0.05$ and 100,000 iterations.

in Table I, where the noncentrality parameter λ given in Table I is equal to $2nI^*$, with n the number of simplex families [Kaplan et al., 1997]. Note that when $C_{MN} = 1$, I^*_{MD} and I^*_{ND} must be equal. In case 4 we examine various cases where $I^*_{MD} \neq I^*_{ND}$ for different values of C_{MN}.

All the results in Table II show that the power of the test using the Monte Carlo procedure is always at least as large as the power of the test using the Bonferroni procedure. The percent gain in power depends primarily on the amount of linkage disequilibrium between alleles at the marker loci C_{MN}. If there is relatively little disequilibrium between marker alleles, the powers of the tests for the two procedures are similar, regardless of the levels of association between the marker and disease alleles I^*_{MD} I^*_{ND}. As the association between alleles at the marker loci increases, with $I^*_{MD} = I^*_{ND}$ and held constant, the difference in the powers of the tests for the two procedures increases with the maximum power difference occurring when $C_{MN} = 1$. As the $I^*_{MD} = I^*_{ND}$ value increases the overall value of the powers of the tests increases, and the difference between the powers of the tests using the two procedures decreases because the power curves will converge to 1. The power curves also converge when C_{MN} is small regardless of the values of I^*_{MD}, I^*_{ND} . For example, when $C_{MN} \leq 0.30$ the difference in the power is negligible. Case 4 illustrates that the percent gain in power can be large even when $C_{MN} < 1$. Power difference is also important when power is small. In case 4 where $C_{MN} = 0.938$, $I^*_{MD} = 0.00013$, $I^*_{ND} = 0.000111$ the Bonferroni procedure has power less than the nominal alpha level of 0.05 even though there is positive association between the markers and the disease, while the Monte Carlo procedure has power of 0.0616.

The powers of the tests for the two procedures depend upon C_{MN}, I^*_{MD} I^*_{ND} ,

and the haplotype frequencies in a complex way. Similar values of C_{MN}, I^*_{MD}, I^*_{ND} can result from very different haplotype frequencies. In Table III, we give two examples where $I^*_{MD} = I^*_{ND} = 0.0198$. In Example 1 where $C_{MN} = 0.355$, the power of the Monte Carlo procedure is 0.54, the power of the Bonferroni procedure is 0.56, and the difference is 0.02. In Example 2 $C_{MN} = 0.38$ and the power of the Monte Carlo procedure is 0.73, the power of the Bonferroni procedure is 0.75, and the difference is 0.02. While the values of the powers are quite different, the difference in the power remains similar because the difference in the power is strongly influenced by C_{MN}.

The addition of a third marker locus increases the complexity of the situation. In addition to three pairwise measures of association between the marker alleles, there is also a global measure of association between the marker alleles. In general, the results for three markers and a single disease locus are similar to the findings for two loci: namely the largest difference in power occurs when there are high levels of association between marker alleles (results not shown).

In order to see what the difference in the power might be for a more realistic situation where association between markers was not controlled but was allowed to evolve as a consequence of recombination, we constructed a multimarker simulation using coalescent theory. Using the Bonferroni procedure gave an estimated power of 0.57 (95% CI = [0.56, 0.58]), whereas the Monte Carlo procedure gave an estimated power of 0.64 (95% CI = [0.63, 0.65]). The difference in power is much less than the theoretical maximum suggested in Table I, but is nonetheless a statistically significant difference.

DISCUSSION

Using a Monte Carlo procedure to determine significance guarantees the appropriate experimentwise type I error α_T and can increase the power of the test using TDT_{MAX} compared to using the Bonferroni procedure. When the test statistics for each locus are correlated under H_{02}, which occurs when there is association between marker alleles, using the Monte Carlo procedure can have more power than using the Bonferroni procedure. When the test statistics are not correlated under H_{02}, the

TABLE III. Two Examples of Possible Sets of Haplotype Frequencies With $I^*_{MD} = I^*_{ND} = 0.0198$ and Similar Values of C_{MN}*

	Example 1: 0.355	Example 2: 0.380
$Pr(M_1N_1\|D_1)$	0.30	0.10
$Pr(M_1N_1\|D_2)$	0.40	0.20
$Pr(M_1N_2\|D_1)$	0.10	0.40
$Pr(M_1N_2\|D_2)$	0.10	0.40
$Pr(M_2N_1\|D_1)$	0.10	0.40
$Pr(M_2N_1\|D_2)$	0.10	0.40
$Pr(M_2N_2\|D_1)$	0.50	0.10
$Pr(M_2N_2\|D_2)$	0.40	0.00

*In example 1, the power of the Monte Carlo procedure is 0.54 while the power of the Bonferroni procedure is 0.56. In example 2, the power of the Monte Carlo procedure is 0.73, while the power of the Bonferroni procedure is 0.75. D_1 indicates the disease allele and D_2 indicates the non-disease allele.

powers of the tests for the two procedures is similar. Thus, using the Monte Carlo procedure may be more powerful, and will never be less powerful than the Bonferroni procedure.

The Monte Carlo procedure we describe is easily executable. Only the individual locus information is needed in order to determine the empirical distribution of TDT_{MAX}; haplotype information is not necessary. In fact, since each test is conducted at an individual marker locus, the haplotype information is ignored. Thus, the test lacks power to detect the case where haplotypes are associated with a disease allele but individual marker alleles are not associated with the disease allele.

The Bonferroni procedure relies on the convergence of each single marker test to a chi-squared distribution. If for any reason the single marker test does not converge to a chi-square distribution, then the principles behind the Bonferroni procedure will not compensate for this error. On the other hand, the Monte Carlo procedure determines an empirical distribution of the TDT_{MAX} statistic, and does not rely on asymptotics. Thus, the failure of any single marker test to converge may invalidate the Bonferroni procedure, whereas the Monte Carlo procedure has no such limitation. In order to avoid relying on asymptotics, single marker significance levels may be determined using a Monte Carlo procedure as described in Kaplan et al. [1997]. The results of these single marker tests can then be combined with a Bonferroni procedure. However, if the Monte Carlo procedure is performed at each locus, then using the multilocus Monte Carlo procedure to determine overall significance requires very little additional computation.

The Monte Carlo approach described in this paper can be applied to other statistics measuring association, e.g., the HHRR (or AFBAC) [Terwilliger and Ott, 1992; Thomson, 1995], the Sib-TDT [Spielman and Ewens, 1998], or the population based case-control test. In addition, the Monte Carlo procedure can be applied to other summary statistics such as the sum of all individual tests. The powers of these and other tests under different genetic models remains a topic for future study. Although we focused on testing for association in simplex families, the Monte Carlo procedure described could be used in conjunction with tests for association that use multiple affected siblings, for example, the test proposed by Martin et al. [1997].

Another benefit of the procedure described is that it allows easily accessible sequential testing. Often the scientist wishes to know not only if the maximum is significant but whether the next largest statistic is also significant. The procedure proposed will generate a vector of P values for all the order statistics. However, we would caution that if multiple order statistics are examined, the 0.05 criteria for the maximum will no longer be appropriate.

The Monte Carlo procedure described in this paper can also be applied to whole genome scans if a significant amount of association between marker alleles is suspected as with the old order Amish or Finnish populations. For loosely linked maps, e.g., a density of 5–10 cM, the association between marker alleles is expected to be minimal, and the difference between the two procedures would be negligible. In this case, the recommendation of Spielman and Ewens [1996] to use the Bonferroni procedure is appropriate [Martin et al., 1997]. For genetic maps with as small a density as suggested by Risch and Merikangis [1996], and Wang et al. [1998] it is likely that there will be association between alleles at neighboring marker loci and using the Monte Carlo procedure will result in a more powerful test.

ACKNOWLEDGMENTS

This work was partially supported by NCI grant 5PO1CA72099 (L.M.M.), NIH grant GM45344 (K.L.S.), NIH grant AG16996 (L.M.M.), NSF grant DBI-9808026 (L.M.M., K.L.S.), and the Agriculture Experiment Station, Purdue University. We thank the two anonymous reviewers for their contributions to the manuscript.

APPENDIX

We show that under H_{02}, the TDT statistics at two tightly linked markers will be independent if the marker alleles are not associated. To simplify the discussion, we assume that the marker and disease loci are biallelic and that the marker loci are completely linked.

Consider two markers M and N, with alleles indexed M_i and $N_{i'}$. Let $P_{M_iM_j}$ be the probability that allele M_i was transmitted to an affected child and M_j was not. The TDT statistic at marker M depends on the transmission probability conditional on the parental genotypes, which we can write as

$$Q_{M_iM_j} = \frac{P_{M_iM_j}}{P_{M_iM_j} + P_{M_jM_i}} .$$

If there is no association between alleles at the marker locus and the disease locus, then whether a marker allele is transmitted or not transmitted does not depend on the marker allele; a parent will transmit either of its alleles with equal probability. Therefore, $P_{MiMj} = P_{MjMi}$, and so $P_{MiMj} = ½$. The same is true for marker N.

Considering markers M and N together, let $P_{MiMj,Ni'Nj'}$ be the joint probability that marker alleles M_i and $N_{i'}$ are transmitted and M_j and $N_{j'}$ are not transmitted. The joint distribution of the TDT statistics at both markers depends on the joint transmission probability conditional on the parental genotypes, which we write as

$$Q_{M_iM_j,N_{i'}N_{j'}} = \frac{P_{M_iM_j,N_{i'}N_{j'}}}{P_{M_iM_j,N_{i'}N_{j'}} + P_{M_jM_i,N_{i'}N_{j'}} + P_{M_iM_j,N_{j'}N_{i'}} + P_{M_jM_i,N_{j'}N_{i'}}} . \quad (4)$$

The parental genotypes for two markers are comprised of two haplotypes with a particular (but unknown) phase. When H_{02} is true, then the transmission probability of a haplotype does not depend on the allele types. Therefore,

$$P_{M_iM_j,N_{i'}N_{j'}} = P_{M_jM_i,N_{j'}N_{i'}} \text{ and } P_{M_jM_i,N_{i'}N_{j'}} = P_{M_iM_j,N_{j'}N_{i'}} . \quad (5)$$

Let $Pr(M_iN_{i'},M_jN_{j'})$ represent the population frequency of the genotype comprised of those two haplotypes, with phase as shown. If there is no association between marker alleles, then each phase is equally likely. That is

$$Pr(M_iN_{i'}, M_jN_{j'}) = Pr(M_iN_{j'}, M_jN_{i'}) \quad (6)$$

for all i, j, i', j'. Furthermore, under H_{02}, the transmission probabilities only depend on the population genotype frequencies. Hence for all i, j, i', j'

$$P_{M_iM_j,N_{i'}N_{j'}} = \frac{Pr(M_iN_{i'}, M_jN_{j'})}{2} \tag{7}$$

Combining (1)-(4) shows that $Q_{M_iM_j,N_i'N_j'} = ¼$, and so $Q_{M_iM_j,N_i'N_j'} = Q_{M_iM_j}Q_{N_i'N_j'}$, showing that the tests are independent when H_{02} is true and the marker alleles are not associated. Finally, if the tests are independent, i.e., $Q_{M_iM_j,N_i'N_j'} = ¼$, then it follows from (3), (4), and (6) that (5) must be true, showing that for linked markers the tests are independent under H_{02} if the alleles at the two markers are not associated.

REFERENCES

Casella G, Berger R. 1990. Statistical inference. Pacific Grove, CA: Wadsworth and Brooks.

Hudson RR. 1983. Properties of a neutral allele model with intragenic recombination. Theoret Pop Biol 23:183–201.

Kaplan NL, Hudson RR. 1985. The use of sample geneaologies for studying a selectively neutral *m*-loci model with recombination. Theoret Pop Biol 28:382–96.

Kaplan NL, Martin ER, Weir BS. 1997. Power studies for the transmission/disequilibrium tests with multiple alleles. American Journal of Human Genetics, 60:691–702.

Knapp M, Seuchter SA, Bauer MP. 1993. The haplotype-relative-risk method for analysis of association in nuclear families. Am J Hum Genet 52:1085–93.

Lazzeroni L, Lange K. 1998. A conditional inference framework for extending the transmission/disequilibrium test. Hum Hered 48:67–81.

Marshall E. 1998. Private help for public databases? Science 280:667–68.

Martin ER, Kaplan NL, Weir BS. 1997. Tests for linkage and association in nuclear families. Am J Hum Genet 61:439–48.

Risch N, Merikangas K. 1990. The future of genetic studies of complex human diseases. Science 273:1516–7.

Sham PC, Curtis D. 1995. An extended transmission/disequilibrium test for multiallele marker loci. Ann Hum Genet 59:323–36.

Simonsen KL, Churchill GA. 1997. A Markov chain model of coalescence with recombination. Theoret Pop Biol 52:43–59.

Spielman RS, Ewens W. 1996. The TDT and other family based tests for linkage disequilibrium and association. Am J Hum Genet 59:983–9.

Spielman RS, Ewens W. 1998. A sibship test for linkage in the presence of association: the sib transmission/disequilibrium. Am J Hum Genet 61:450–8.

Spielman RS, McGinnis RE, Ewens W. 1993. Transmission test for linkage disequilibrium: the insulin gene region and insulin dependent diabetes mellitus (IDDM). Am J Hum Genet 52:506–16.

Terwilliger JD, Ott J. 1992. A haplotype-based 'Haplotype Relative Risk' approach to detecting allelic associations. Hum Hered 42:337–46.

Thomson G. 1995. Mapping disease genes: family based association studies. Am J Hum Genet 61:487–98.

Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J, Kruglyak L, Stein L, Hsie L, Topaloglou T, Hubbel E, Robinson E, Mittman M, Morris MS, Shen N, Kilburn D, Rioux J, Nusbaum C, Rozen S, Hudson T, Lipshutz R, Chee M, Lander ES. 1998. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 280:1077–82.