Haplotype Sharing Transmission/Disequilibrium Tests That Allow for Genotyping Errors

[1] *Qiuying Sha,* [1] *Jianping Dong,* [1] *Renfang Jiang,* [1] *Huann-Sheng Chen,*

[1,2,*] *Shuanglin Zhang*

[1]Department of Mathematical Sciences

Michigan Technological University, Houghton, MI 49931

[2]Department of Mathematics, Heilongjiang University

Harbin 150080, China

* Corresponding author:

Shuanglin Zhang, Ph.D.

Department of Mathematical Sciences

Michigan Technological University

1400 Townsend Drive

Houghton, MI 49931

Phone: (906) 487-2095

Fax: (906) 487-3133 (Fax to Shuanglin Zhang)

Email: shuzhang@mtu.edu

Key Words: Genotyping error, TDT test, family-based study, linkage, association, haplotype sharing, multiple markers

Running Title: HS-TDT Allowing for Genotyping Errors

ABSTRACT

The present study introduces new haplotype sharing transmission/disequilibrium tests that allow for random genotyping errors. We evaluate the type I error rate and power of the new proposed tests under a variety of scenarios and perform a power comparison among the proposed tests, the HS-TDT and the single-marker TDT. The results indicate that the HS-TDT shows a significant increase in type I error when applied to data in which either Mendelian inconsistent trios are removed or Mendelian inconsistent markers are treated as missing genotypes, and the magnitude of the type I error increases both with an increase in sample size and with an increase in genotyping error rate. The results also show that a simple strategy, that is, merging each rare haplotype to a most similar common haplotype, can control the type I error inflation for a wide range of genotyping error rates, and after merging rare haplotypes the power of the test is very similar to that of without merging the rare haplotypes. Therefore, we conclude that a simple strategy may make the HS-TDT robust to genotyping errors. This strategy may also be applicable to other haplotype-based TDT tests.

1. Introduction

The transmission/disequilibrium test (TDT) (Spielman et al. 1993) and the allied tests have become popular tools for testing genetic linkage and association between a marker and a susceptibility locus. It is now widely accepted that the study of multi-marker haplotypes will likely yield more genetic information than the study of a single marker. Therefore, several authors have proposed TDT-type tests using multi-marker haplotypes (Lazzeroni and Lange 1998; Merriman et al. 1998; Clayton and Jones 1999; Clayton 1999; Rabinowitz and Laird 2000; Zhao et al. 2000; Seltman et al. 2001; Li et al. 2001; Bourgain et al. 2000, 2001, 2002). One negative feature of the haplotype-based tests is that the number of haplotypes will increase rapidly with the number of markers. A large number of haplotypes and, thus, a large number of degrees of freedom will limit the power of the haplotype-based TDT tests. Recently, Zhang et al. (2003) proposed a Haplotype-Sharing TDT (HS-TDT) that uses the information of similarities between the haplotypes to reduce the degrees of freedom. The degrees of freedom (in a broad sense) of the HS-TDT increases linearly with the number of markers instead of with the number of haplotypes. Simulation results show that, in most cases, the HS-TDT is more powerful than the single marker TDT.

Similar to most of the other TDT-type tests, the HS-TDT assumes that the genetic data are without genotyping errors. In practice, genotyping errors can occur for many reasons, including sample swaps in the lab, pedigree errors such as non or incorrectly specified relationships, and technology failures. Much attention has recently been paid to methods for detecting genotyping errors (Lincoln and Lander 1992; Brzustowicz et al. 1993; Ott 1993; Lunetta et al. 1995; Ehm et al. 1996; Stringham and Boehnke 1996; Ghosh et al. 1997; O'Connell and Weeks 1998, 1999; Douglas et al. 2000, 2002; Ewen et al. 2000) and evaluating the effect of undetected errors on genetic analysis (Gordon and Ott 2001; Akey et al.

2001; Heath 1998; Knapp and Becker 2004; Mitchell et al. 2003). There are only a few recent papers that consider methods allowing for errors in linkage analysis (Goring and Terwilliger 2000a, 2000b, 2000c, 2000d) and LD analysis (Gordon et al. 2001).

As noted by Gordon et al. (2001) and Mitchell et al. (2003), genotyping errors may lead to an inflated type I error rate for the single-marker TDT (Spielman et al. 1993), and Gordan et al. (2003) has proposed a new TDT test (TDT_{ae}) that allows for random genotyping errors by incorporating an error model into a likelihood-based linkage disequilibrium analysis. Recently, Knapp and Becker (2004) has pointed out that genotyping errors also lead to an inflated type I error rate for the HS-TDT. The reason is that the transmitted haplotypes are partially checked for genotyping errors by Mendelian inconsistency (MI), whereas there is no such checking at all for nontransmitted haplotypes. As a result of the unbalanced checking for genotyping errors, nontransmitted haplotypes appear less similar than transmitted haplotypes, which may lead to an inflated type I error rate for the HS-TDT. This is especially true for the case in which there is only one child per nuclear family (Zhang et al. 2004). For the HS-TDT, the simulation studies in Knapp and Becker (2004) and Zhang et al. (2004) show that even quite small probabilities of genotyping errors ($\leq 0.5\%$) can lead to a dramatic inflation of the type I error. In this report, we propose several strategies to modify the HS-TDT such that it can be robust to random genotyping errors. We use simulation studies to evaluate the performance of the modified HS-TDTs and the original HS-TDT (Zhang et al. 2003). Our simulation results show that, after merging each rare haplotype to a most similar common haplotype, the HS-TDT has correct type I error rate for a wide range of genotyping error rates, and the power is very similar to that of the HS-TDT in which rare haplotypes are not merged. We conclude that using such method to modify the HS-TDT can make

it robust to genotyping errors and thus make the HS-TDT more practical.

Background of HS-TDT

Suppose n nuclear families are sampled with t_i children in the ith family and L tightly linked markers are typed both for the children and for the parents. Let y_{ik} denote the trait value of the kth child in the ith family (for qualitative trait, $y = 1$ and $y = 0$ denote affected and unaffected status, respectively). For each haplotype H, Zhang et al. (2003) began by defining a Haplotype-Sharing (HS) score at the lth marker as

$$X_H(l) = \frac{1}{4n} \sum_{i=1}^{n} \sum_{j=1}^{4} S_{H,H_{ij}}(l),$$

where $H_{i1}, \ldots, H_{i4}$ denote the four parental haplotypes in the ith family; $S_{H_1,H_2}(l)$ is the Haplotype Length (HL) similarity between haplotypes H_1 and H_2 at the lth marker, which is defined as the length of the contiguous region around the lth marker over which the two haplotypes are Identical By State (IBS). As pointed out by Zhang et al. (2003), it is expected that, around the disease locus, the HS score will be larger for a haplotype with disease mutation than that for a haplotype without disease mutation.

At the lth marker, let $X_{i1}(l), \ldots, X_{i4}(l)$ denote the HS scores of the four parental haplotypes of the ith family. For the kth child in the ith family, let $x_{ik}(l)$ denote the difference of HS scores between the transmitted and nontransmitted haplotypes. For example, among the four parental haplotypes $H_{i1}, \ldots, H_{i4}$, H_{i1} and H_{i2} are transmitted to the kth child, then $x_{ik}(l) = X_{i1}(l) + X_{i2}(l) - X_{i3}(l) - X_{i4}(l)$. Furthermore, define $U_i(l)$ as

$$U_i(l) = \frac{1}{t_i} \sum_{k=1}^{t_i} (y_{ik} - c) x_{ik}(l),$$

where $c = 0$ for the case that trait value is qualitative and only affected children and their parents are sampled, and $c = \overline{y} = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{t_i} \sum_{k=1}^{t_i} y_{ik}$ for all other cases.

The HS-TDT statistic is defined by

$$U = \max_{1 \le l \le L} |U(l)|,$$

where $U(l) = \sum_{i=1}^{n} U_i(l)$, and the p-value of the test is evaluated by the permutation procedure.

The above mentioned test procedure assumes that the multi-marker haplotypes of both parents and children are known. In the case of ambiguous haplotypes, the first step to use the HS-TDT is to estimate haplotype frequencies by the expectation-maximization algorithm incorporating family information (Chen and Zhang, 2003). Zhang et al. (2003) proposed two different methods to make use of estimated haplotype frequencies in case of ambiguous phase information in the families of the sample: (1) each ambiguous family is assigned its most likely haplotype configuration and (2) each possible haplotype configuration of an ambiguous family is weighted by its relative likelihood.

Methods to control Type I error inflation due to random genotyping errors

Merging Rare Haplotypes

When we consider tightly linked markers within a candidate gene, the LD between markers is expected and thus the total number of haplotypes across a set of tightly linked markers is not large. Consequently, one genotyping error occurred in a haplotype will most likely generate a new rare haplotype. If we merge each of the rare haplotypes to the corresponding most similar common haplotype, we probably can recover most of the typing errors. When we merge the rare haplotypes, all the rare haplotypes that are not caused by genotyping errors are also merged to the corresponding common haplotypes. However, under null hypothesis, each rare haplotype that is not caused by genotyping errors will be equally likely to be transmitted and non-transmitted, and thus merging the rare haplotypes that are not caused by genotyping errors will not affect the type

I error rate of the test. Based on this idea, we propose the following modified test procedure: (1) find all possible haplotype configurations for each family and estimate the haplotype frequencies using EM-FD (Chen and Zhang, 2003); (2) each of the rare haplotypes is merged to a common haplotype that is the most similar to the rare haplotype, and all possible haplotype configurations for each family and the haplotype frequencies will be changed accordingly; (3) follow the same steps used in the HS-TDT (Zhang et al. 2003). We denote the statistical test based on this procedure by HS-TDT$_m$.

In order to merge rare haplotypes to their similar haplotypes, we give another similarity measure called Allele Count (AC) similarity. The AC similarity of the two haplotypes is defined as the number of markers at which the two haplotypes have the same allele. For two haplotypes H and h covering L markers, let H_i (h_i) denote the alleles of the haplotype H (h) at marker i. Then, the AC similarity between haplotype H and h is $\sum_{l=1}^{L} I_{(H_l=h_l)}$, where indicator function $I_{(H_l=h_l)} = 1$ if $H_l = h_l$ and $I_{(H_l=h_l)} = 0$ if $H_l \neq h_l$. To merge the rare haplotypes, we first specify a cutoff value α_0, and all the haplotypes with frequencies $\leq \alpha_0$ are considered as rare haplotypes. We merge each of the rare haplotype to a common haplotype (frequency $>\alpha_0$) that is most similar to this rare haplotype based on AC similarity. When several common haplotypes have the same AC similarity with a rare haplotype, the latter is merged to the common haplotype with the highest haplotype frequency.

A Similarity measure that is more robust to genotyping errors

As discussed in Zhang et al. (2004), one reason of the type I error inflation of the HS-TDT due to genotyping errors is the HL similarity used in Zhang et al. (2003). The HL similarity is sensitive to genotyping errors. A genotyping error in one marker may substantially change the HL similarity and then the HS score of a haplotype. In Zhang et al. (2004), we proposed another haplotype

similarity called Adjust Haplotype Length (AHL) similarity that is defined as follows. For two haplotypes H and h, let H_i (h_i) denote the alleles of the haplotype H (h) at marker i. To find the AHL similarity of the two haplotypes H and h around marker i, we compare alleles of the two haplotypes in the right-hand side markers beginning from marker $i+1$, until marker $i+r$ that satisfies $H_{i+r} \neq h_{i+r}$ and either $H_{i+r+1} \neq h_{i+r+1}$ or $H_{i+r+2} \neq h_{i+r+2}$. Then, similarly, we compare alleles of the two haplotypes in the left-hand side markers beginning from marker $i-1$, until marker $i-l$ that satisfies $H_{i-l} \neq h_{i-l}$ and either $H_{i-l-1} \neq h_{i-l-1}$ or $H_{i-l-2} \neq h_{i-l-2}$. The AHL similarity measure is defined as the distance between marker $i-l$ and marker $i+r$. See Figure 1 for the difference of the HL similarity and AHL similarity. Note that a genotyping error occurs at one marker but does not occur at the nearby markers will not affect the AHL similarity and the probability that typing errors occur in several consecutive markers is very small. Thus, we expect that the HS-TDT by using AHL similarity instead of HL similarity will be robust to genotyping errors. We denote the HS-TDT based on AHL similarity by HS-TDT$_s$.

Modified HS-TDT tests and other tests compared

We consider three modified HS-TDT tests which include the HS-TDT$_m$, the HS-TDT$_s$, and the HS-TDT$_{ms}$ (merging rare haplotypes and using AHL similarity). For the cases with and without genotyping errors, we use simulation studies to evaluate the type I error and power of the original HS-TDT proposed by Zhang et al. (2003), the three adjusted HS-TDT tests and a single-marker TDT (Spielman et al. 2003). The test statistics compared in this article are summarized in Table 1. The HS-TDT and adjusted HS-TDTs test for linkage and association between a candidate gene and a trait. The p-value of the tests is evaluated by the permutation procedure described in Zhang et al. (2003). To make a fair comparison between the haplotype-based tests and the single-marker

TDT, we adjust the single-marker TDT as follows to test for linkage and association between a candidate gene and a trait. Suppose that there are L markers typed within a candidate gene. Let TDT_i denote the value of the TDT statistic for marker i. We use $TDT_{\max} = \max_{1 \le i \le L} TDT_i$ as the test statistic. The p-value of $TDT_{\max}$ is also evaluated by the permutation procedure. In this report, we use 1000 permutations to evaluate the p-values of all the tests.

Simulations

We use simulation studies to evaluate the type I error and the power of the tests. For each scenario, we generate 1000 samples. In all the simulation studies, we consider qualitative trait and nuclear families with one affected child. This is the case in which the HS-TDT has the largest inflated type I error rate (Zhang et al. 2004).

Data sets for assessing the type I error: The first set of simulations has 19 tightly-linked and equidistant biallelic marker loci. The haplotypes (29 different haplotypes occured in the population) and the corresponding frequencies are given in Table A (online-only supplemental material) of Knapp and Becker (2004). For all family trios, we generate the parents' genotypes according to haplotype frequencies. The haplotype pair in the child is obtained by randomly selecting one of the two haplotypes in each parent. Genotyping errors are introduced independently into the alleles according to the stochastic error model, that is, an allele is changed with probability ϵ. We consider the cases of $\epsilon = 0$ (no genotyping errors), $\epsilon = 0.005$, $\epsilon = 0.01$ and $\epsilon = 0.02$. If a genotyping error comes visible by leading to Mendelian inconsistency (MI), we treat the genotypes of all individuals of the family at the marker with MI as missing values and our method of estimating haplotype frequencies allows missing values. However, when we perform the single-marker TDT, the family is discarded from the analysis in the presence of MI.

In the second set of simulations, the haplotypes are obtained using a direct simulation method described in Zhang et al. (2003) and Tzeng et al. (2003). Briefly, diploid individuals are paired at random in their generation and mated. The number of children per couple is randomly drawn from a Poisson distribution with mean λ. Each population is founded by 500 individuals and the expected size remains at 500 for 50 generations (the reproduction rate $\lambda = 2$). This initialization, together with small population growth in early generations, generates random linkage disequilibrium among alleles on normal chromosomes. After 50 generations, the population grows exponentially for 100 generations to a final size of 10,000 individuals. In this period, the reproduction rate λ is determined by the exponential growth rate. One disease mutation was introduced on one chromosome in the 51st generation. To generate the chromosome in the founder population, we generate alleles at each of the markers independently according to allele frequencies. The minor allele frequency at each marker is drawn from a uniform distribution over the interval (0.1, 0.4). The simulation program produces populations from which samples of haplotypes can be drawn. The methods to generate family trios and genotyping errors are the same as those in the first set of simulation. Eleven biallelic markers were simulated, covering a 1 cM region, with a spacing of 0.1 cM between the adjacent markers. The disease locus is assumed to be in linkage equilibrium with the eleven markers. To eliminate the population effect, for each scenario in this set of simulations, we generate 100 populations and draw 10 samples from each population.

Data sets for assessing the power: To assess power of the tests, we use the direct simulation method described above to generate the population. We also consider eleven markers loci, covering a 1 cM region, with a spacing of 0.1 cM between the adjacent markers. However, we assume the 6th marker is located at the disease susceptibility locus (assuming a negligible recombination rate), but

the 6th marker is not the disease susceptibility locus itself (i.e. the marker is not the functional polymorphism). Let D and d denote the two alleles at disease locus, RR denote the relative risk of genotypes DD to dd, and q_D denote the allele frequency of allele D. For the given RR, q_D and the disease models, we generate parental genotypes at the disease susceptibility locus according to the probability of mating types under the condition that the child is affected. The parental multi-marker genotypes are generated according to the genotypes at the disease susceptibility locus. For example, if the father's genotype at the disease susceptibility locus is Dd, we randomly choose one haplotype with the disease mutation and one haplotype without the disease mutation to form the father's multi-marker genotype. Conditional on the parents' mating types, the affected child's genotype is generated by ignoring the recombination. For each scenario, we generate 1000 samples. To eliminate the population effect, the 1000 sample are drawn from 100 simulated populations with 10 samples from each population.

Results

As stated earlier, Zhang et al. (2003) proposed two different methods to deal with ambiguous phase information. The simulation results presented in this section are based on the first method that assigns ambiguous family its most likely haplotype configurations. We also did the simulation based on the second method to deal with ambiguous phase information. The results from the second method (not shown) are almost identical to the results from the first method.

Type I error rate

When data is error free, the type I error rates of all the five tests have no significant difference from the nominal levels (results not shown). When there are genotyping errors, we consider genotyping error rate $\epsilon = 0.005, 0.01$ or 0.02, number of family trios $N = 100$ or 200 and rare haplotype cutoff value $\alpha_0 = 0.01, 0.02$ or 0.03 (for tests HS-TDT$_m$ and HS-TDT$_{ms}$). For each combination

of ϵ, N and α_0, we evaluate the type I error of the five tests. The results are given in Tables 2 and 3 for the two sets of the simulations, respectively. For 1000 replicated samples, the 95% confidence intervals of type I error rates are $(0.0362, 0.0638)$ and $(0.0037, 0.0163)$ for nominal levels 5% and 1%, respectively. The results in the two sets of simulations are very consistent. The results show that, though the type I error inflation of the HS-TDT$_s$ is smaller than that of the HS-TDT, both the HS-TDT and the HS-TDT$_s$ have inflated type I errors in all the cases, and the inflation increases both with an increase in the sample size and with an increase in the genotyping error rate. This results indicate that, though using AHL similarity can reduce the magnitude of the type I error inflation, the HS-TDT based on AHL still cannot control the false-positive results. The single-marker TDT also has an inflated type I error for large sample size or large genotyping error rate, and the magnitude of the inflation also increases both with an increase in sample size and with an increase in the genotyping error rate. When we use the haplotype-merging strategy, if the rare haplotype cut-off value $\alpha_0 \geq 2\%$, the type I error rates of both the HS-TDT$_m$ and the HS-TDT$_{ms}$ are less than the upper bound of the confidence intervals. When the rare haplotype cut-off value $\alpha_0 = 1\%$, the type I error rates of the HS-TDT$_m$ and the HS-TDT$_{ms}$ (for the second set of simulations) are slightly higher than the upper bound of the confidence intervals. The magnitude of the type I error inflation does increase with an increase in the genotyping error rate but not with an increase in sample size.

Power comparisons

In the first set of simulations, we compare the powers of the five tests using error-free data. The results are summarized in Figure 2. Figure 2 shows that the power of the single-marker TDT is substantially less than that of all HS-TDT tests. Among the four HS-TDT tests, the two based on HL similarity (HS-TDT

and HS-TDT$_m$) are more powerful than the two based on AHL similarity (HS-TDT$_s$ and HS-TDT$_{ms}$). After merging the rare haplotypes, the tests are slightly less powerful than the corresponding tests without merging rare haplotypes, that is, the HS-TDT$_m$ is slightly less powerful than the HS-TDT and the HS-TDT$_{ms}$ is slightly less powerful than the HS-TDT$_s$. In the other two sets of simulations, we evaluate the impact of genotyping error rates and rare haplotype cut-off values on the powers of the HS-TDT$_{ms}$and the HS-TDT$_m$. The results (given in Figures 3 and 4) show that the powers of the tests slightly decrease as the cut-off value or genotyping error rate increases.

In the simulations (both for type I error and for power) based on simulated populations, for each scenario, the 1000 samples are drawn from 100 populations. In the simulated populations, the number of distinct haplotypes varied from 35 to 75, and the common haplotypes (frequency ≥ 0.01) varied from 6 to 35. So, our simulation results cover a wide range of haplotype structures.

In summary, both the HS-TDT$_m$ and the HS-TDT$_{ms}$ can control the false-positive due to genotyping errors and HS-TDT$_m$ is more powerful than the HS-TDT$_{ms}$. The HS-TDT$_m$ is only slightly less powerful than the original HS-TDT proposed by Zhang et al. (2003) which has significant type I error inflation when there are genotyping errors in the data.

Discussion

It is widely recognized that using multi-marker haplotypes instead of considering one marker at a time can improve the statistical power of association studies. One difficulty to construct the haplotype-based test is that the large number of haplotypes will lead to a large number of degrees of freedom of the test. Zhang et al. (2003) proposed a HS-TDT, a haplotype based TDT using multiple tightly-linked markers, to test linkage and association between the disease susceptibility locus and a candidate gene. The HS-TDT is applicable to both qualitative and

quantitative traits and allows for ambiguous phase information. The degrees of freedom of the HS-TDT increases linearly with the number of markers rather than with the number of haplotypes. Simulation results show that the HS-TDT is more powerful than the single-marker TDT. One negative feature of HS-TDT, as pointed out by Knapp and Becker (2004), is that the genotyping error may lead to a type I error inflation. Our simulation results also show that the original HS-TDT (Zhang et al. 2003) is very sensitive to the genotyping errors. With only 0.5% of genotyping errors, it may lead to a significant type I error inflation. In practice, even for high quality data, genotyping errors are often unavoidable. Thus, methods that make the HS-TDT robust to genotyping errors are needed. One way to develop robust methods, as Gordon et al. (2001) did for the single-marker TDT, is to incorporate the error models in the statistical analysis. However, this method will make the statistical model much more complicated. In this article, we proposed several simple strategies to control the false-positive of the HS-TDT due to genotyping errors. From our simulation results, we found that a simple strategy, that is, merging each rare haplotype to a similar common haplotype, can control for false-positive of the HS-TDT due to genotyping errors, and the power is competitive to the original HS-TDT. When the rare haplotype cut-off value $\alpha_0 \geq 2\%$, The HS-TDT$_m$ (HS-TDT based on merged haplotypes) has a correct type I error rate (within 95% CI of the nominal level) in all the cases of our simulations. Furthermore, when the data is error-less, the power of the HS-TDT$_m$ is very similar to that of the HS-TDT and is more powerful than the HS-TDT$_s$ (HS-TDT based on AHL similarity), the HS-TDT$_{ms}$ (HS-TDT based on AHL similarity and merged haplotypes), and the single-marker TDT in all the cases we considered.

One question for merging the rare haplotypes is how to choose the cut-off value α_0. Our simulations suggest that using $\alpha_0 = 2\%$ or larger can control the

false-positive for a wide range of genotyping error rates. When $\alpha_0 = 1\%$, though the HS-TDT$_m$ cannot fully control the false-positive due to genotyping errors, the magnitude of the type I error inflation of the HS-TDT$_m$ is much less than that of the HS-TDT and is substantially less than that of the single-marker TDT.

Reference

Akey JM, Zhang K, Xiong M, Doris P, Jin L (2001) The effect that genotyping errors have on the robustness of common linkage-disequilibrium measures. Am J Hum Genet 68:1447–1456.

Bourgain C, Genin E, Quesneville H, Clerget-Darpoux F (2000) Search for multifactorial genes in founder populations. Ann Hum Genet 64:255-265.

Bourgain C, Genin E, Holopainen P, Mustalahti K, Maki M, Partanen J, Clerget-Darpoux F (2001) Use of closely related affected individuals for the genetic study of complex disease in founder populations. Am J Hum Genet 68:154-159.

Bourgain C, Genin E, Ober C, Clerget-Darpoux F (2002) Missing data in haplotype analysis: a study on the MILC method. Ann Hum Genet 66:99-108.

Brzustowicz LM, Merette C, Xie X, Townsend T, Gilliam C, Ott J (1993) Molecular and statistical approaches to the detection and correction of errors in genotype databases. Am J Hum Genet 53:1137–1145.

Chen HS, Zhang SL (2003) Haplotype inference for multiple tightly linked multilocus phenotypes including nuclear family information. p.165-171: Proceeding of the International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences (ed. Valafar F, Valafar H). CSREA Press, Las Vegas.

Clayton DG (1999) A generalization of the transmission/disequilibrium test for uncertain-haplotype transmission. Am J Hum Genet 65:1170-1177.

Clayton DG, Jones H (1999) Transmission/disequilibrium tests for extended marker haplotypes. Am J Hum Genet 65:1161-1169.

Douglas JA, Boehnke M, Lange K (2000) A multipoint method for detecting genotyping errors and mutations in sibling-pair linkage data. Am J Hum Genet 66:1287–1298.

Douglas JA, Skol AD, Boehnke M (2002) Probability of detection of genotyping errors and mutations as inheritance inconsistencies in nuclear-family data. Am J Hum Genet 70:487–495.

Ehm MG, Kimmel M, Cottingham RW Jr (1996) Error detection for pedigree data, using likelihood methods. Am J Hum Genet 58:225–234.

Ewen KR, Bahlo M, Treloar SA, Levinson DF, Mowry B, Barlow JW, Foote SJ (2000) Identification and analysis of error types in high-throughput genotyping. Am J Hum Genet 67:727–736.

Ghosh S, Karanjawala ZE, Hauser ER, Ally D, Knapp JI, Rayman JB, Musick A, Tannenbaum J, Te C, Shapiro S, Eldridge W, Musick T, Martin C, Smith JR, Carpten JD, Brownstein MJ, Powell JI, Whiten R, Chines P, Nylund SJ, Magnuson VL, Boehnke M, Collins FS (1997) Methods for precise sizing, automated binning of alleles, and reduction in largescale genotyping using fluorescently labeled dinucleotide markers: FUSION (Finland-US Investigation ofNIDDM Genetics) study group. Genome Res 7:165–178.

Gordon D, Ott J (2001) Assessment and management of single nucleotide polymorphism genotype errors in genetic association analysis. Pac Symp Biocomput 2001:18–29.

Gordon D, Heath SC, Liu X, Ott J (2001) A transmission disequilibrium test that allows for genotyping errors in the analysis of single-nucleotide polymorphism data. Am J Hum Genet 69:371–380.

Goring HHH, Terwilliger JD (2000a) Linkage analysis in the presence of errors I: complex-valued recombination fractions and complex phenotypes. Am J Hum Genet 66:1095–1106.

——— (2000b) Linkage analysis in the presence of errors II: marker-locus genotyping errors modeled with hypercomplex recombination fractions. Am J Hum Genet 66:1107–1118.

——— (2000c) Linkage analysis in the presence of errors III: marker loci and their map as nuisance parameters. Am J Hum Genet 66:1298–1309.

——— (2000d) Linkage analysis in the presence of errors IV: joint pseudomarker analysis of linkage and/or linkage disequilibrium on a mixture of pedigrees and singletons when the mode of inheritance cannot be accurately specified. Am J Hum Genet 66:1310–1327.

Heath SC (1998) A bias in TDT due to undetected genotyping errors. Am J Hum Genet Suppl 63:A292.

Knapp M, Becker T (2004) Impact of genotyping error on type I error rate of the haplotype-sharing transmission /disequilibrium test (HS-TDT). Am J Hum Genet 74:589–591.

Lazzeroni LC, Lange K (1998) A conditional inference frame-work for extending the transmission/disequilibrium test. Hum Hered 48:67-81.

Li J, Wang D, Dong J, Jiang R, Zhang K, Zhang S, Zhao H, Sun F (2001) The power of transmission disequilibrium tests for quantitative traits. Genetic

Epidemiology. 18 (Supp 1):S632-S637.

Lincoln SE, Lander ES (1992) Systematic detection of errors in genetic linkage data. Genomics 14:604–610.

Lunetta KL, Boehnke M, Lange K, Cox DR (1995) Experimental design and error detection for polyploid radiation hybrid mapping. Genome Res 5:151–163.

Merriman TR, Eaves IA, Twells RC, Merriman ME, Danoy PA, Muxworthy CE, Hunter KM, Cox RD, Cucca F, Mc-Kinney PA, Shield JP, Baum JD, Tuomilehto J, Tuomilehto-Wolf E, Ionesco-Tirgoviste C, Joner G, Thorsby E, Undlien DE, Pociot F, Nerup J, Ronningen KS, Bain SC, Todd JA (1998) Transmission of haplotypes of microsatellite markers rather than single marker alleles in the mapping of a putative type 1 diabetes susceptibility gene (IDDM6). Hum Mol Genet 7:517-524.

Mitchell AA, Cutler DJ, ChakravartiAm A (2003) Undetected Genotyping Errors Cause Apparent Overtransmission of Common Alleles in the Transmission/Disequilibrium Test. Am J Hum Genet 72:598-610.

O'Connell JR, Weeks DE (1998) PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 63:259–266.

——— (1999) An optimal algorithm for automatic genotype elimination. Am J Hum Genet 65:1733–1740.

Ott J (1993) Detecting marker inconsistencies in human gene mapping. Hum Hered 43:25–30.

Rabinowitz D, Laird N (2000) A unified approach to adjusting association tests for population admixture with arbitrary pedigree structure and arbitrary missing marker information. Hum Hered 50:211-223.

Seltman H, Roeder K, Devlin B (2001) Transmission/Disequilibrium Test Meets Measured Haplotype Analysis: Family-Based Association Analysis Guided by Evolution of Haplotypes. Am J Hum Genet 68:1250-1263.

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

Stringham HM, Boehnke M (1996) Identifying marker typing incompatibilities in linkage analysis. Am J Hum Genet 59:946–950.

Tzeng JY, Devlin B, Wasserman L, Roeder K (2003) On the identification of disease mutations by the analysis of haplotype similarity and goodness of fit. Am J Hum Genet. 72:891-902.

Zhang S, Sha Q, Chen HS, Dong J, Jiang R (2003) Transmission/disequilibrium test based on haplotype sharing for tightly linked markers. Am J Hum Genet 73:566–579.

Zhang S, Sha Q, Chen HS, Dong J, Jiang R (2004) Impact of genotyping error on type I error rate of the haplotype-sharing transmission /disequilibrium test (HS-TDT): reply to Knapp and Becker. Am J Hum Genet 74: 591-593.

Zhao H, Zhang S, Merikangas KR, Trixler M, Wildenauer DB, Sun F, Kidd KK (2000) Transmission/Disequilibrium test using multiple tightly linked markers. Am Hum Genet 67:936-346.

Table 1. Test Statistics Compared

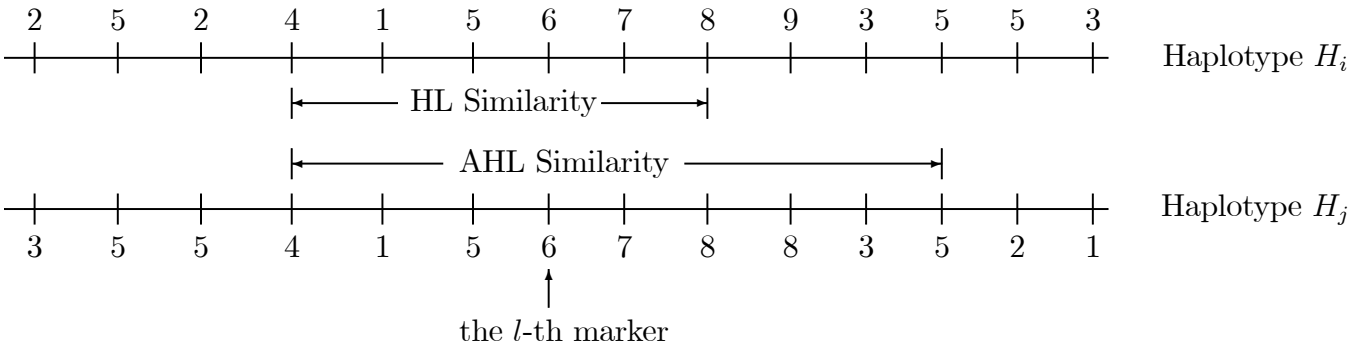
HS-TDT	using HL similarity without merging rare haplotypes (Zhang et al. 2003)
HS-TDT$_s$	using AHL similarity without merging rare haplotypes
HS-TDT$_m$	using HL similarity with merging rare haplotypes
HS-TDT$_{ms}$	using AHL similarity with merging rare haplotypes
TDT	the single-marker TDT proposed by Spielman et al. (1993)

Table 2. Type I error rates of the tests. The simulations are based on haplotype frequencies in Table A of Knapp and Becker (2004).

α	N	ϵ	TDT	HS-TDT	HS-TDT$_s$	cutoff $\alpha_0 = 1\%$		cutoff $\alpha_0 = 2\%$		cutoff $\alpha_0 = 3\%$	
						HS-TDT$_m$	HS-TDT$_{ms}$	HS-TDT$_m$	HS-TDT$_{ms}$	HS-TDT$_m$	HS-TDT$_{ms}$
1%	100	0.005	0.005	0.086	0.020	0.013	0.011	0.007	0.011	0.008	0.009
		0.01	0.016	0.304	0.045	0.016	0.014	0.009	0.009	0.005	0.004
		0.02	0.062	0.854	0.177	0.024	0.004	0.006	0.007	0.003	0.003
	200	0.005	0.012	0.170	0.036	0.010	0.011	0.011	0.008	0.007	0.011
		0.01	0.035	0.640	0.069	0.009	0.006	0.012	0.006	0.003	0.013
		0.02	0.203	0.900	0.360	0.018	0.009	0.008	0.005	0.005	0.007
5%	100	0.005	0.047	0.226	0.094	0.053	0.052	0.046	0.052	0.048	0.047
		0.01	0.059	0.559	0.135	0.061	0.052	0.043	0.039	0.029	0.022
		0.02	0.180	0.960	0.419	0.102	0.044	0.037	0.035	0.025	0.022
	200	0.005	0.060	0.361	0.116	0.048	0.055	0.052	0.048	0.051	0.045
		0.01	0.134	0.860	0.204	0.050	0.043	0.042	0.038	0.027	0.043
		0.02	0.414	0.994	0.612	0.078	0.042	0.032	0.033	0.034	0.048

Table 3. Type I error rates of the tests. The simulations are based on simulated populations by direct simulation method.

α	N	ϵ	TDT	HS-TDT	HS-TDT$_s$	cutoff 1%		cutoff 2%		cutoff 3%	
						HS-TDT$_m$	HS-TDT$_{ms}$	HS-TDT$_m$	HS-TDT$_{ms}$	HS-TDT$_m$	HS-TDT$_{ms}$
1%	100	0.005	0.008	0.074	0.027	0.014	0.013	0.011	0.014	0.014	0.009
		0.01	0.011	0.190	0.053	0.021	0.020	0.016	0.016	0.015	0.006
		0.02	0.091	0.587	0.188	0.068	0.024	0.013	0.011	0.011	0.006
	200	0.005	0.012	0.099	0.017	0.013	0.010	0.011	0.013	0.011	0.009
		0.01	0.080	0.350	0.072	0.017	0.013	0.015	0.009	0.013	0.011
		0.02	0.351	0.836	0.316	0.043	0.014	0.016	0.013	0.015	0.010
5%	100	0.005	0.050	0.184	0.104	0.060	0.053	0.057	0.048	0.057	0.046
		0.01	0.072	0.430	0.155	0.076	0.054	0.060	0.051	0.055	0.045
		0.02	0.278	0.786	0.367	0.144	0.078	0.061	0.050	0.044	0.038
	200	0.005	0.068	0.269	0.097	0.062	0.041	0.053	0.045	0.052	0.050
		0.01	0.222	0.601	0.220	0.068	0.050	0.053	0.046	0.061	0.051
		0.02	0.474	0.933	0.525	0.127	0.060	0.063	0.043	0.057	0.058


2 5 2 4 1 5 6 7 8 9 3 5 5 3
Haplotype H_i
HL Similarity
AHL Similarity
Haplotype H_j
3 5 5 4 1 5 6 7 8 8 3 5 2 1
the l-th marker

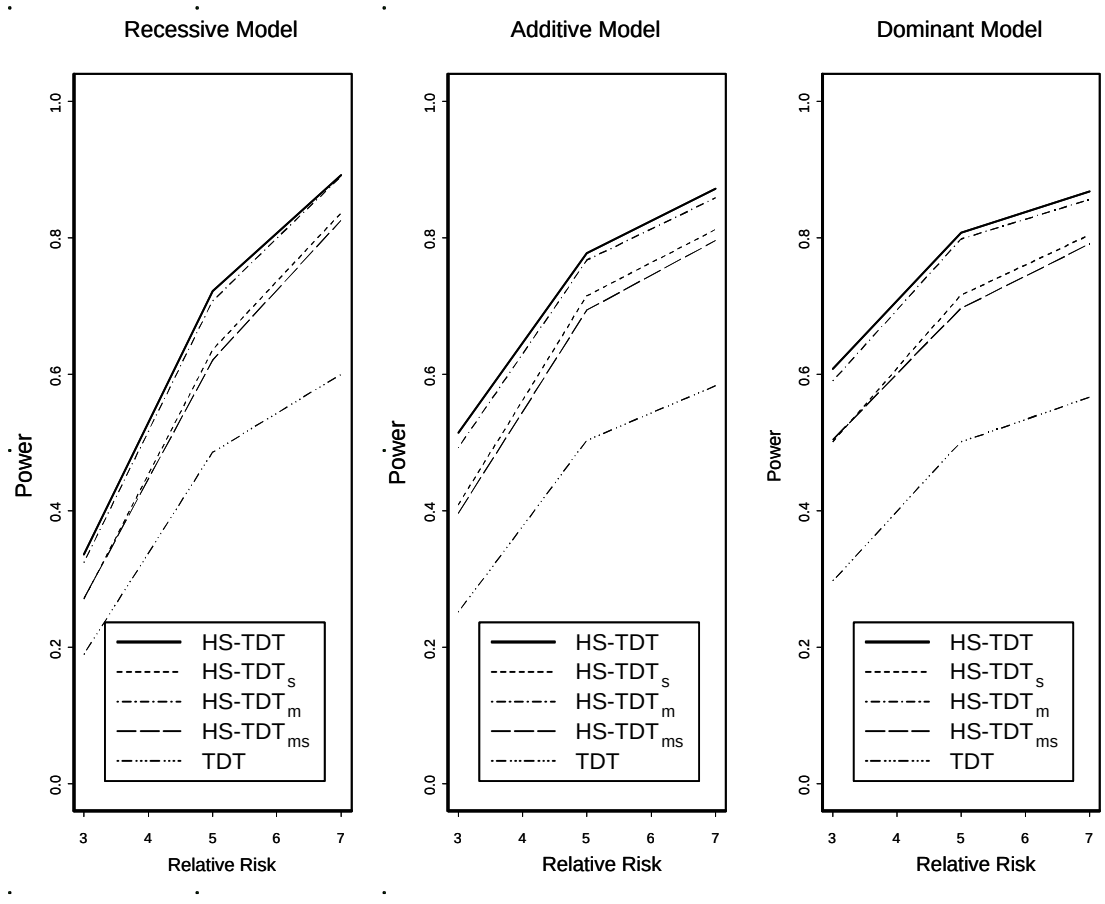


Figure 2. Power comparisons of the five tests using error-free data. The sample size N=100. Rare haplotype cut-off value (for the HS-TDT$_m$ and the HS-TDT$_{ms}$) $\alpha_0 = 2\%$.

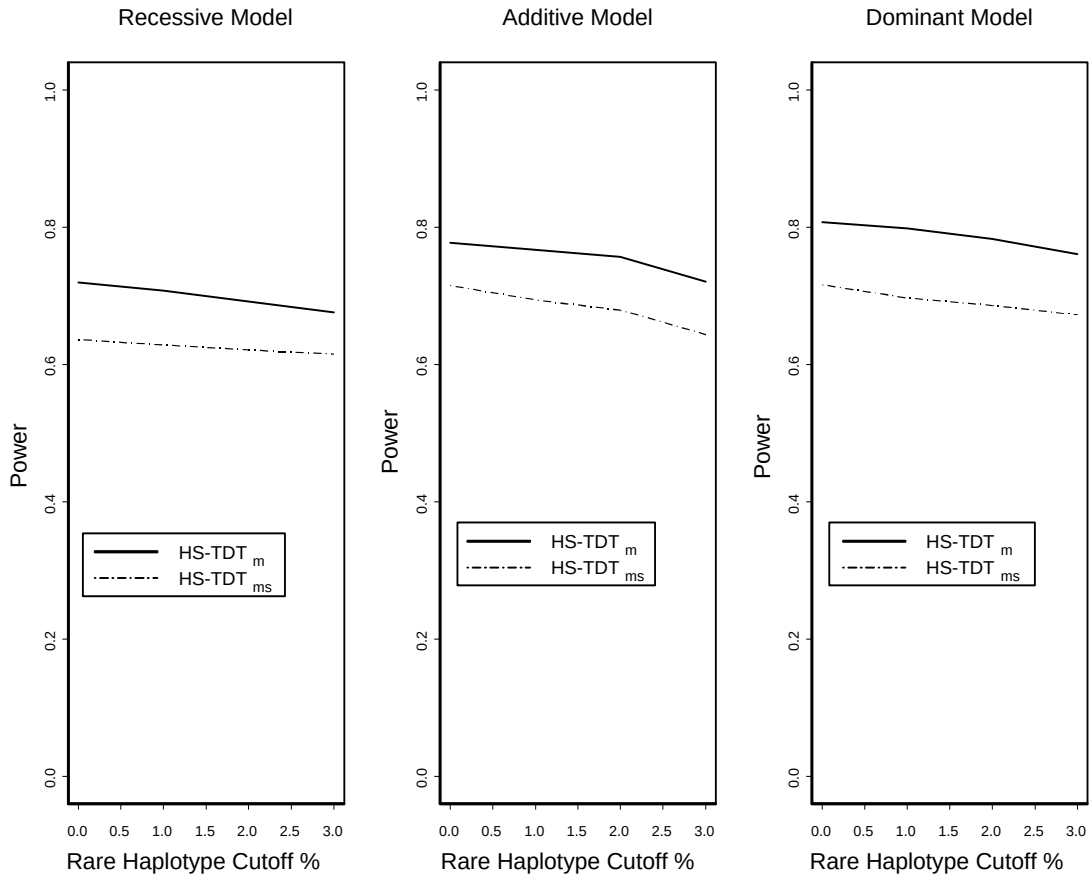


Figure 3. Power comparisons of the HS-TDT$_m$ and the HS-TDT$_{ms}$ for different rare haplotype cutoff values. The sample size N=100. Genotypic relative risk $RR = 5$. Genotyping error rate $\epsilon = 1\%$.

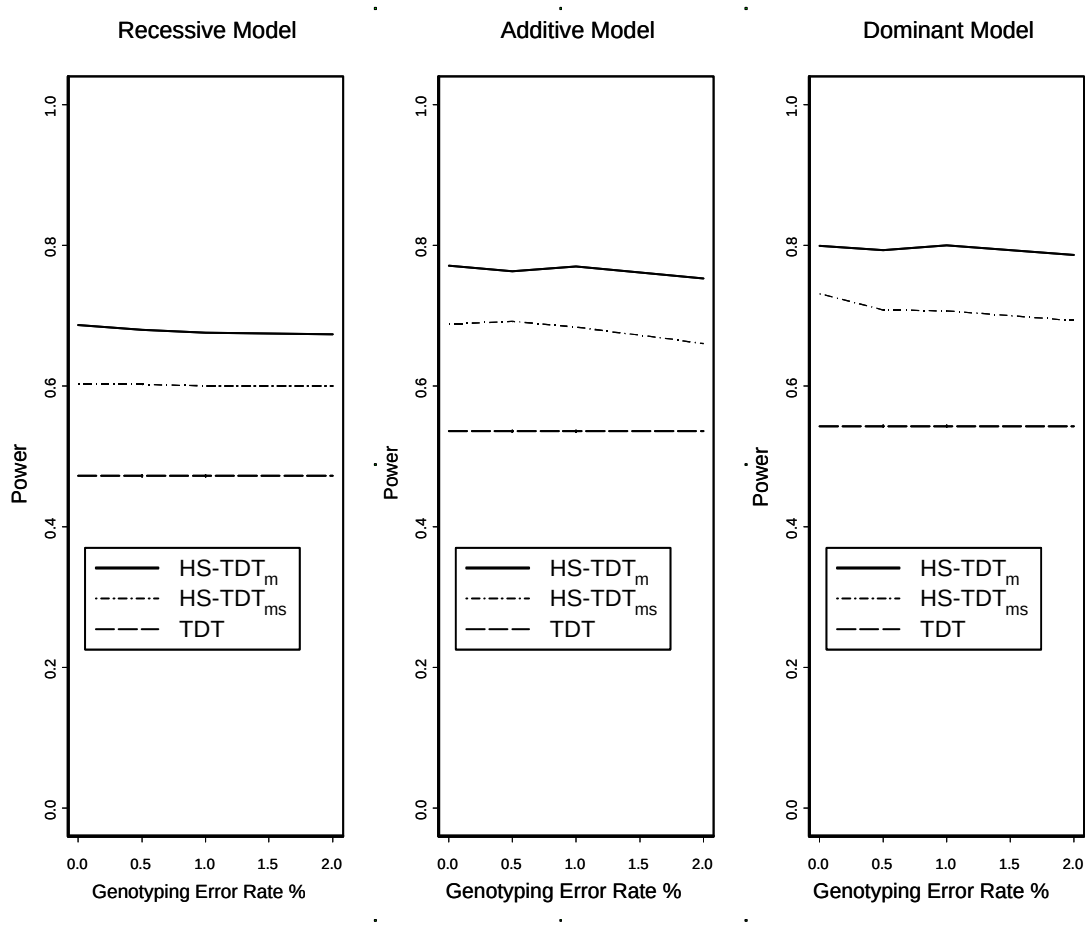


Figure 4. Power comparisons of the three tests for different genotyping error rates. The sample size N=100. Genotypic relative risk $RR = 5$. Rare haplotype cut-off value $\alpha_0 = 2\%$. Single-marker TDT test (TDT) uses error-free data.