

## Exact Robust Tests for Detecting Candidate-Gene Association in Case-Parents Trio Design

Zehua Chen<sup>1</sup> and Gang Zheng<sup>2</sup>

<sup>1</sup>*National University of Singapore* and

<sup>2</sup>*National Heart, Lung and Blood Institute*

*Abstract:* In the case-parents trio design for testing candidate-gene association, the distribution of the data under the null hypothesis of no association is completely known. Therefore, the exact null distribution of any test statistic can be simulated by using Monte-Carlo method. In the literature, several robust tests have been proposed for testing the association in the case-parents trio design when the genetic model is unknown, but all these tests are based on the asymptotic null distributions of the test statistics. In this article, we promote the exact robust tests using Monte-Carlo simulations. It is because: (i) the asymptotic tests are not accurate in terms of the probability of type I error when sample size is small or moderate; (ii) asymptotic theory is not available for certain good candidates of test statistics. We examined the validity of the asymptotic distributions of some of the test statistics studied in the literature and found that in certain cases the probability of type I error is greatly inflated in the asymptotic tests. In this article, we also propose new robust test statistics which are statistically more reasonable but without asymptotic theory available. The powers of these robust statistics are compared with those of the existent statistics in the literature through a simulation study. It is found that these robust statistics are preferable to the others in terms of their efficiency and robustness.

*Key words:* Asymptotic distribution, case-parents design, disease-gene association, exact robust test, Monte-Carlo simulation.

### 1. Introduction

In testing the association between a candidate gene and a disease, the case-parents trio design proposed by Schaid and Sommer (1993) has been extensively studied in recent years. The case-parents trio design is the scheme that the disease-affected children (cases) and their parents are ascertained, and then the genotypes of the cases and the parents are obtained. Statistical models were developed based on the genotype relative risks defined by ratios of penetrances of

the candidate-gene genotypes and the conditional probabilities of case genotypes given parental mating types. Various test procedures were proposed and studied for testing the candidate-gene association under different assumptions on the underlying genetic mechanism. Schaid and Sommer (1993) proposed score tests when the underlying genetic mechanism is determined by one of the four genetic models, i.e., additive, recessive, multiplicative and dominant models. Note that the score test is asymptotically optimal when the underlying genetic model is correctly specified. In practice, however, the underlying mode of inheritance is unknown for complex diseases. For this situation, using a score test may lose substantial power when the genetic model is mis-specified. Hence, it is necessary to consider some robust tests which are independent of model assumptions. Schaid (1999) considered an un-constrained likelihood ratio test (LRT) procedure, Zheng, Freidlin and Gastwirth (2002) studied a test procedure which they coined as MAX3, and Zheng, Chen and Li (2003) proposed a restricted version of likelihood ratio test (RLRT).

The robust test procedures (MAX3 and RLRT) mentioned above are based on the asymptotic distributions of the test statistics. However, under the null hypothesis of no association, the distribution of the data in the case-parents design is completely known. By this fact, the exact null distribution of any test statistic can be simulated by Monte-Carlo methods. Since we are now armed with computing facilities of great capacity, Monte-Carlo methods are feasible and we can substitute computer powers for asymptotic theories. In this article, we consider the exact robust tests using Monte-Carlo methods with case-parents design data. There are three persuasive reasons for this. (i) An asymptotic theory only provides an approximation to the null distribution of a test statistic, the accuracy of the approximation depends on sample size, and how large the sample size must be for the approximation to be satisfactorily accurate is unknown in most of cases. But the Monte-Carlo method simulates the exact null distribution, it can be made arbitrarily accurate by simply increasing simulation size. (ii) Asymptotic theory is not always available for every test statistic, especially, in cases where the regularity conditions for classical theory are not satisfied. The Monte-Carlo method, however, is not subject to such regularity conditions. (iii) Though, in principle, the exact null distribution can be determined without resorting to Monte-Carlo simulation, as in the case of Fisher's exact test, it is only feasible when the sample size is small. In the trio design, data consist of three informative parental mating types. The expected sample sizes of the informative mating types are unbalanced. Even the total sample size of the three informative mating types is relatively large, e.g. 1,000, the expected sample sizes for some informative mating types could still be small, which affects adversely the validity of asymptotic distributions of some test statistics. On the other hand, with a

total sample size 1,000, the determination of the exact null distribution without resorting to Monte-Carlo simulation is practically impossible.

It should be remarked here that, by an exact test, we mean the critical value or the p-value of the test is determined by the true null distribution rather than the asymptotic distribution of the test statistic, which differs from what are called exact tests by other authors. For example, Cleves, Olson and Jacobs (1997) considered exact transmission-disequilibrium tests, and Schaid (1999) considered exact tests for the case-parents design. However, what they referred to as an exact test is a special test whose critical region consists of points of small probabilities, that is, any point in the critical region has a probability smaller than any point outside the critical region. Simulation methods are also considered by other authors. For instance, Lazzeroni and Lange (1998) considered a simulation method for TDT. But their simulation method is based on permutation instead of the exact null distribution. In this article, we propose exact robust tests of which the asymptotic distributions are not available for the test statistics. We conduct extensive Monte-Carlo simulations to compare the power and robustness of various tests.

The article is arranged as follows. In section 2, the background of the case-parents design is given and various test statistics are discussed including the new ones we propose. In section 3, the method of exact test using Monte-Carlo simulation is described, and the issue on the validity of the asymptotic approximations is addressed. In section 4, the results on the power comparison of various test statistics are presented. Conclusions are given in section 5.

## 2. Background

Suppose the candidate-gene of concern has two different alleles denoted by  $A$  and  $a$ , which can form three possible genotypes  $aa$ ,  $Aa$  and  $AA$ . One of them has higher risk than the other. Denote the penetrances by

$$f_0 = \Pr(\text{disease}|aa), \quad f_1 = \Pr(\text{disease}|Aa), \quad f_2 = \Pr(\text{disease}|AA),$$

which are probabilities of developing disease conditional on the genotype. By taking  $aa$  as the base genotype, the relative risks of genotypes  $Aa$  and  $AA$  are defined as  $r_1 = f_1/f_0$  and  $r_2 = f_2/f_0$  respectively. Genetic models can be specified using the genotype relative risks. A genetic model is called recessive (rec), additive (add), multiplicative (mul), or dominant (dom) if  $r_1 = 1$ ,  $r_1 = (1+r_2)/2$ ,  $r_2 = r_1^2$ , or  $r_1 = r_2$ , respectively. The null hypothesis of no association refers to that the disease status is independent of the genotypes. In terms of the relative risks, the null hypothesis can be stated as  $H_0 : r_1 = r_2 = 1$ , i.e.,  $f_0 = f_1 = f_2$ , that is, individuals with different genotypes have the same risk to develop the

disease. Given parental mating types, the conditional probabilities of case genotypes can be derived in terms of  $r_1$  and  $r_2$ , which provides the foundation for testing association between candidate-gene and disease in the case-parents design. For a gene with two alleles, there are six possible combinations of parental genotypes referred to as mating types: 1)  $AA \times AA$ , 2)  $AA \times Aa$ , 3)  $AA \times aa$ , 4)  $Aa \times Aa$ , 5)  $Aa \times aa$ , and 6)  $aa \times aa$ . The possible case (offspring) genotypes of each mating type are given in Table 1 (second column). For example, the case genotype can only be  $AA$  for mating type 1), since both parents can only transmit an  $A$  allele to the offspring. For mating type 2), there are two possible case genotypes:  $Aa$  (one parent transmits  $A$  and the other transmits  $a$ ) and  $AA$  (both parents transmit  $A$ ). Note that for mating types 1), 3) and 6), each of them can produce only one possible genotype for the offsprings. Since whether or not there exists association between the disease and the candidate gene is reflected by whether or not the probabilities of different case genotypes are different given a parental mating type, mating types 1), 3) and 6) are non-informative. The third column of Table 1 consists of counts for different combinations of genotypes of trios. The conditional probabilities of case genotypes given parental mating types are given in columns 4 to 8 of Table 1 in terms of  $r_1$  and  $r_2$ : column 4 for general genetic model (no assumption of relationship between genotype relative risks) and columns 5 to 8 for the four particular genetic models mentioned earlier. The conditional case genotype probabilities can be obtained as follows. Let  $g_0$ ,  $g_1$  and  $g_2$  stand for case genotypes  $aa$ ,  $Aa$  and  $AA$ , respectively. Let  $D$  and  $MT$  stand for disease and mating type, respectively. Denote by  $\Pr(g_i|MT, D)$  the conditional probability of case genotype  $g_i$  given mating type  $MT$ . Then by the Bayes formula we have

$$\begin{aligned} \Pr(g_i|MT, D) &= \frac{\Pr(g_i, D|MT)}{\Pr(D|MT)} = \frac{\Pr(g_i|MT)\Pr(D|MT, g_i)}{\sum_j \Pr(g_j|MT)\Pr(D|MT, g_j)} \\ &= \frac{\Pr(g_i|MT)\Pr(D|g_i)}{\sum_j \Pr(g_j|MT)\Pr(D|g_j)} = \frac{f_i\Pr(g_i|MT)}{\sum_j f_j\Pr(g_j|MT)} \end{aligned}$$

where  $\Pr(D|MT, g_i) = \Pr(D|g_i)$  because the disease status only depends on the genotype  $g_i$  and  $\Pr(g_j|MT)$  can be obtained by the Mendelian rule: a parent transmits his or her two alleles to a offspring equally likely. For example,  $\Pr(g_1 = Aa|MT = AA \times Aa) = 1/2$ ,  $\Pr(g_1 = AA|MT = Aa \times Aa) = 1/4$  and  $\Pr(g_1 = Aa|MT = Aa \times Aa) = 1/2$ .

In the case-parents trio design, the cases (affected children) and their parents are ascertained and their genotypes are obtained. The data from the design consists of the numbers of case genotypes for each parental mating type. Let  $n_k, k = 2, 4, 5$  be the total number of cases with the  $k$  th parental mating type.

Table 1: Conditional probabilities of case genotypes given parental mating types for various genetic models

Parental Mating Type	Case Genotype	Count	Pr(case genotype mating type, disease)				
			General ( $r_1, r_2$ )	REC $r_1 = 1$ $r_2 = r$	ADD $r_1 = r$ $r_2 = 2r - 1$	MUL $r_1 = r$ $r_2 = r^2$	DOM $r_1 = r$ $r_2 = r$
1) $AA \times AA$	$AA$	$n_{12}$	1	1	1	1	1
2) $AA \times Aa$	$AA$	$n_{22}$	$\frac{r_2}{r_1+r_2}$	$\frac{r}{r+1}$	$\frac{2r-1}{3r-1}$	$\frac{r}{r+1}$	$\frac{1}{2}$
	$Aa$	$n_{21}$	$\frac{r_1}{r_1+r_2}$	$\frac{1}{r+1}$	$\frac{r}{3r-1}$	$\frac{r}{r+1}$	$\frac{1}{2}$
3) $AA \times aa$	$Aa$	$n_{31}$	1	1	1	1	1
4) $Aa \times Aa$	$AA$	$n_{42}$	$\frac{r_2}{r_2+2r_1+1}$	$\frac{r}{r+3}$	$\frac{2r-1}{4r}$	$\frac{r^2}{(r+1)^2}$	$\frac{r}{3r+1}$
	$Aa$	$n_{41}$	$\frac{2r_1}{r_2+2r_1+1}$	$\frac{2}{r+3}$	$\frac{1}{2}$	$\frac{2r}{(r+1)^2}$	$\frac{2r}{3r+1}$
	$aa$	$n_{40}$	$\frac{1}{r_2+2r_1+1}$	$\frac{1}{r+3}$	$\frac{1}{4r}$	$\frac{1}{(r+1)^2}$	$\frac{1}{3r+1}$
5) $Aa \times aa$	$Aa$	$n_{51}$	$\frac{r_1}{r_1+1}$	$\frac{1}{2}$	$\frac{r}{r+1}$	$\frac{r}{r+1}$	$\frac{r}{r+1}$
	$aa$	$n_{50}$	$\frac{1}{r_1+1}$	$\frac{1}{2}$	$\frac{1}{r+1}$	$\frac{1}{r+1}$	$\frac{1}{r+1}$
6) $aa \times aa$	$aa$	$n_{60}$	1	1	1	1	1

Then  $n_2 = n_{21} + n_{22}$ ,  $n_4 = n_{40} + n_{41} + n_{42}$  and  $n_5 = n_{50} + n_{51}$ . Let  $n = n_2 + n_4 + n_5$ . Given  $(n_2, n_4, n_5)$ , the numbers of cases with the three informative parental mating types,  $(n_{21}, n_{22})$ ,  $(n_{40}, n_{41}, n_{42})$  and  $(n_{51}, n_{50})$  are conditionally independent and follow binomial or trinomial distributions with cell probabilities given in Table 1. The likelihood function of  $(r_1, r_2)$  is therefore given by

$$L(r_1, r_2) = \frac{r_1^{n_{21}+n_{41}+n_{51}} r_2^{n_{22}+n_{42}}}{(1+r_1)^{n_5} (1+2r_1+r_2)^{n_4} (r_1+r_2)^{n_2}}. \quad (2.1)$$

In the remainder of this section, we discuss various test statistics for testing the null hypothesis of no association based on the above likelihood function.

When the alternative hypothesis is specified as one of the four genetic models mentioned in the last paragraph, Schaid (1999) studied the likelihood ratio test (LRT) and Schaid and Sommer (1993) considered the score test for each of the four models. They also showed that the score statistics for the additive model and the multiplicative model are the same. These score statistics are given below:

$$Z_{\text{rec}} = (4n_{22} + 4n_{42} - 2n_2 - n_4) / \sqrt{4n_2 + 3n_4}, \quad (2.2)$$

$$Z_{\text{dom}} = (n_4 + 2n_5 - 4n_{40} - 4n_{50}) / \sqrt{3n_4 + 4n_5}, \quad (2.3)$$

$$\begin{aligned} Z_{\text{add}} &= Z_{\text{mul}} \\ &= (n_{22} + 2n_{42} + n_{51} - n_{21} - 2n_{40} - n_{50})/\sqrt{n_2 + 2n_4 + n_5}. \end{aligned} \quad (2.4)$$

It is noted that the score statistic  $Z_{\text{add}}$  for the additive model is also the statistic of the transmission disequilibrium test (Spielman, McGinnis and Ewens, 1993). Under the null hypothesis, the score statistics follow asymptotically standard normal distributions. The critical values of the score tests can therefore be determined by the standard normal distribution. When the high risk allele (either  $A$  or  $a$ ) is specified, a one-sided test is to be carried out. If the allele status is unspecified, a two-sided test is to be carried out.

As mentioned in Section 1, when the underlying genetic model is unknown, the use of any of the above score tests  $Z_{\text{rec}}$ ,  $Z_{\text{add}}$  and  $Z_{\text{dom}}$  may lose substantial power when the model is mis-specified. This is because these three tests correspond to three extreme cases. Robust tests are therefore desirable. That a test is robust is in the sense that its power is not affected much by the underlying genetic model. The use of robust tests was introduced in Gastwirth (1985) and Freidlin, Podgor and Gastwirth (1999) in a general context. In the context of case-parents trio design, when the genetic model is unspecified under the alternative hypothesis but the allele risk status is specified, say,  $A$  is specified as the high risk allele, several test statistics which share certain robust properties have been proposed in the literature. Zheng et al. (2002) considered the MAX3 statistic given by

$$\text{MAX3}_1 = \max\{Z_{\text{rec}}, Z_{\text{add}}, Z_{\text{dom}}\}. \quad (2.5)$$

The asymptotic distribution of  $\text{MAX3}_1$  is determined by the joint asymptotic distribution of  $Z_{\text{rec}}$ ,  $Z_{\text{add}}$  and  $Z_{\text{dom}}$ . Note that  $Z_{\text{add}} = (c_1 Z_{\text{dom}} + c_2 Z_{\text{rec}})/(2c_3)$ , where  $c_1 = (3n_4 + 4n_5)^{1/2}$ ,  $c_2 = (4n_2 + 3n_4)^{1/2}$  and  $c_3 = (n_2 + 2n_4 + n_5)^{1/2}$ . Therefore the asymptotic distribution of  $\text{MAX3}_1$  is indeed determined by the joint asymptotic distribution of  $Z_{\text{dom}}$  and  $Z_{\text{rec}}$ . They derived the exact correlation between  $Z_{\text{rec}}$  and  $Z_{\text{dom}}$  under the null hypothesis, which is given by  $\rho = n_4/\{(3n_4+4n_5)(3n_4+4n_2)\}^{1/2}$ . Though there is no closed form for the asymptotic distribution of  $\text{MAX3}_1$ , the distribution can be simulated easily. Zheng et al. (2003) studied a restricted likelihood ratio test (RLRT). Note that when  $A$  is specified as the high risk allele, we have  $r_2 \geq r_1 \geq 1$ . It is then more reasonable to consider the likelihood ratio test under this restriction. The test statistic for the RLRT is given by

$$\text{RLRT}_1 = 2 \log \left\{ \max_{(r_1, r_2) \in T_1} L(r_1, r_2)/L(1, 1) \right\}, \quad (2.6)$$

where  $L(r_1, r_2)$  is the likelihood function given in (2.1), and  $T_1$  is the cone on  $r_1$ - $r_2$  plane determined by the restriction  $r_2 \geq r_1 \geq 1$ . By a result from Self and

Liang (1987), the asymptotic distribution of  $\text{RLRT}_1$  is a mixture of a degenerated distribution at zero and two chi-square distributions with degrees of freedom 1 and 2.

In practice, it is more realistic that both the genetic model under the alternative hypothesis and the risk status of the alleles cannot be specified. In this case, the  $\text{MAX3}_1$  statistic can be extended straightforwardly to the case of unspecified allele risk status as follows:

$$\text{MAX3}_2 = \max \{ |Z_{\text{rec}}|, |Z_{\text{add}}|, |Z_{\text{dom}}| \}. \quad (2.7)$$

Moreover, the range of  $(r_1, r_2)$  is the union of the cone  $T_1$  and the triangle  $T_2$  determined by the restriction  $0 \leq r_2 \leq r_1 \leq 1$ . Let  $T = T_1 \cup T_2$ . Note that  $T$  is not a cone, hence, differs from the parameter space considered by Self and Liang (1987). It is then natural to consider the restricted likelihood ratio test with the alternative that  $(r_1, r_2) \in T \setminus \{(1, 1)\}$ . The test statistic for this RLRT is given below:

$$\text{RLRT}_2 = 2 \log \left\{ \max_{(r_1, r_2) \in T} L(r_1, r_2) / L(1, 1) \right\}. \quad (2.8)$$

We can also consider, as a test statistic, the maximum of the four LRT statistics obtained by restricting to each of the four genetic models. We refer to this statistic as  $\text{MLRT}_1$  when the range of  $(r_1, r_2)$  is confined to  $T_1$ , as  $\text{MLRT}_2$  otherwise. All these robust statistics ( $\text{MAX3}_2$ ,  $\text{RLRT}_2$ ,  $\text{MLRT}_1$ , and  $\text{MLRT}_2$ ), though arising naturally, have not been considered in the literature yet. A difficulty with these statistics is that we cannot find appropriate approximations to the null distributions of these statistics. But, exact null distributions of these robust tests have not been examined.

### 3. Exact Tests and Validity of Asymptotic Distributions

As discussed in the last section, the data from a case-parents design,  $(n_{21}, n_{22})$ ,  $(n_{40}, n_{41}, n_{42})$  and  $(n_{51}, n_{52})$ , follow binomial or trinomial distributions while conditioning on the numbers of cases with the three informative parental mating types. The cell probabilities of these distributions given in Table 1 are determined by  $r_1$  and  $r_2$ . Under the null hypothesis of no association,  $r_1 = r_2 = 1$ . Therefore, the cell probabilities of the binomial and trinomial distributions are completely determined under the null hypothesis. Specifically,  $(n_{21}, n_{22})$  follows the binomial distribution  $B(n_2, 1/2)$ ,  $(n_{40}, n_{41}, n_{42})$  follows the trinomial distribution  $\text{Mul}(n_4, 1/4, 1/2, 1/4)$  and  $(n_{50}, n_{51})$  follows the binomial distribution  $B(n_5, 1/2)$ . As a consequence of this fact, the null distribution of any test statistic can be simulated and exact test can be carried out. Let  $\mathbf{N} = (n_{21}, n_{22}, n_{40}, n_{41}, n_{42}, n_{51}, n_{52})$ . For any test statistic  $S(\mathbf{N})$ , its null distribution

Table 2: The expected sample sizes  $n_2$ ,  $n_4$  and  $n_5$ , conditional on the total sample size  $n = 100$  for various genetic models and allele frequencies

Genetic model ( $r_1, r_2$ )	Allele freq $p$	Expected sample sizes		
		$n_2$	$n_4$	$n_5$
NULL (1, 1)	.01	0.01	1.00	98.99
	.10	1.10	9.89	89.01
	.30	11.39	26.58	62.03
	.50	33.33	33.33	33.33
REC (1, 2)	.01	0.02	1.25	98.74
	.10	1.60	12.00	86.40
	.30	15.21	29.58	55.21
	.50	40.00	33.33	26.67
ADD (2, 3)	.01	0.02	1.33	98.65
	.10	1.76	12.68	85.56
	.30	16.30	30.43	53.26
	.50	41.67	33.33	25.00
MUL (2, 4)	.01	0.02	1.49	98.49
	.10	2.07	13.99	83.94
	.30	18.27	31.98	49.75
	.50	44.44	33.33	22.22
DOM (2, 2)	.01	0.01	1.16	98.82
	.10	1.44	11.31	87.25
	.30	14.04	28.65	57.31
	.50	38.10	33.33	28.57

can be simulated as follows. Given  $n_2, n_4$ , and  $n_5$ , we generate  $\mathbf{N}$  from the above binomial and trinomial distributions for  $m$  times. Here  $m$  is the simulation size and is determined by the accuracy we desire for the simulated null distribution to approximate the theoretical null distribution. Denote these simulated values by  $\mathbf{N}_i, i = 1, \dots, m$ . Then the theoretical null distribution of  $S(\mathbf{N})$  is approximated by

$$F_S(s) = \frac{1}{m} \sum_{i=1}^m \{S(\mathbf{N}_i) \leq s\}.$$

The accuracy of this approximation does not depend on the sample sizes  $n_2, n_4$  and  $n_5$ . In principle, the approximation can be made arbitrarily accurate by simply increasing the simulation size  $m$ .

Although the exact test described above can be carried out no matter whether or not an asymptotic null distribution of the test statistic is available and whether or not the sample sizes are large, one might still prefer an asymptotic test when



the asymptotic null distribution of the test statistic is available because of its simplicity. However, caution must be taken on the validity of the asymptotic distributions, especially when the sample sizes are small or moderate. In the case-parents design, the validity of an asymptotic distribution is affected by the effective sample sizes  $n_2$ ,  $n_4$  and  $n_5$ . These effective sample sizes are indeed random variables and their expected values are greatly affected by the allele frequencies ( $\Pr(A) = p$  and  $\Pr(a) = q = 1 - p$ ) of the population. Under the assumption of Hardy-Weinberg equilibrium, the genotype frequencies are determined by the allele frequencies as:  $\Pr(aa) = q^2$ ,  $\Pr(Aa) = 2pq$  and  $\Pr(AA) = p^2$ . Given the total number  $n$  of cases with the three informative parental mating types ( $n = n_2 + n_4 + n_5$ ), the expected values are given by  $E(n_k) = np_k / (p_2 + p_4 + p_5)$ ,  $k = 2, 4, 5$ , where  $p_2 = 2p^3q(r_1 + r_2)/R$ ,  $p_4 = p^2q^2(r_2 + 2r_1 + 1)/R$ ,  $p_5 = 2pq^3(r_1 + 1)/R$  and  $R = p^2r_2 + 2pqr_1 + q^2$ . The conditional expected sample sizes of each informative parental mating type given  $n = 100$  for various genetic models and allele frequencies are given in Table 2. It can be seen from Table 2 that when  $p$  is small the expected effective sample sizes are very un-balanced, say, when  $p = 0.3$ , the expected sample sizes are 11, 27 and 62. When  $n$  increases, the proportion  $n_k/n$ ,  $k = 2, 4, 5$ , will stay the same as in Table 2. This unbalancedness can affect adversely the validity of the asymptotic distribution, as will be seen in the results to be reported later.

For the test statistics of which asymptotic distributions are available, the validity of the asymptotic distributions for given sample sizes can be verified by comparing with the simulated exact null distributions. In the remainder of this section, we report some results on the comparison between the exact and the asymptotic null distributions for the test statistics  $Z_{\text{rec}}$ ,  $Z_{\text{add}}$ ,  $Z_{\text{dom}}$  and  $\text{MAX3}_1$ . For total sample size  $n = 100, 500$  and  $1,000$ , the effective sample sizes  $n_2, n_4$  and  $n_5$  are calculated using the  $A$  allele frequencies  $p = 0.01, 0.1, 0.3$  and  $0.5$  under the null hypothesis  $r_1 = r_2 = 1$ . For each set of effective sample sizes so obtained, the exact critical values for a one-sided size  $\alpha = 0.05$  test are simulated for the four statistics. The simulation size  $m = 10,000$ . The asymptotic critical values for  $\text{MAX3}_1$  are also obtained by simulation. These critical values are given in Table 3. For the three score statistics, the critical values are to be compared with 1.645 which is the critical value for the asymptotic test. In Table 3, those exact critical values that differ from the corresponding asymptotic critical values by at least 0.05 are marked with an asterisk. While, generally speaking, the asymptotic critical values provide reasonably accurate approximations to the exact critical values, there are a few cases where the asymptotic critical values differ substantially from the exact critical values. The problem is most prominent with the statistic  $Z_{\text{rec}}$ , especially when the allele frequency  $p$  is small.

To further investigate the properties of the exact null distribution of  $Z_{\text{rec}}$

Table 3: The simulated exact and asymptotic critical values of score statistics and  $\text{MAX3}_1$  for tests of size  $\alpha = 0.05$  (simulation size  $m = 10,000$ )

$n$	$p$	$Z_{\text{rec}}$	$Z_{\text{add}}$	$Z_{\text{dom}}$	$\text{MAX3}_1$	
					Exact	Asymp.
100	.01	1.8898*	1.6751	1.6439	1.8898*	1.9749
	.10	1.9612*	1.6772	1.6998*	1.9612	1.9932
	.30	1.5689*	1.6570	1.6454	1.9651	2.0179
	.50	1.6181	1.6432	1.6181	1.9882	2.0394
500	.01	1.9415*	1.6429	1.6529	1.9415	1.9674
	.10	1.6590	1.6387	1.6443	2.0494*	1.9886
	.30	1.6343	1.6526	1.6430	2.0072	2.0206
	.50	1.6698	1.6416	1.6541	2.0362	2.0412
1,000	.01	1.7056*	1.6461	1.6471	1.9082*	1.9684
	.10	1.5986	1.6459	1.6427	2.0025	1.9887
	.30	1.6559	1.6471	1.6414	2.0030	2.0187
	.50	1.6421	1.6440	1.6272	2.0173	2.0432

Note: The asymptotic critical value of size  $\alpha = 0.05$  test for score statistics is 1.645.

when the allele frequency  $p$  is small, we simulated the exact null distribution of  $Z_{\text{rec}}$  with  $p = 0.01$ ,  $n = 100, 500, 1,000$  and  $10,000$  by a simulation of size  $m = 10,000$ . The simulated results reveal that the probability mass of the exact null distribution concentrates on only a few points. For  $n = 100$ , the probability mass concentrates on three points. For  $n = 500$ , the probability mass concentrates on five points. Even for  $n = 1,000$ , the probability mass concentrates on only eight points. This suggests that the null distribution is quite discrete so that a continuous approximation such as the standard normal distribution might not be appropriate. The simulated null distributions are presented in Table 4. For each  $n$ , the left column gives the points at which there is a positive mass, the right column gives the corresponding probabilities. For  $n = 10,000$ , only the upper tail of the distribution is presented. Also shown in Table 4 are the type I errors when the asymptotic critical value 1.645 is used. It can be seen that the true type I error is greatly inflated when the asymptotic test is carried out.

#### 4. Comparison of Test Statistics

With the feasibility of exact tests, we have a much richer repertoire of test statistics for testing gene-disease association in a case-parents design. In this section, we compare, under some specified alternative hypotheses, the powers of the following test statistics:  $Z_{\text{rec}}$ ,  $Z_{\text{add}}$ ,  $Z_{\text{dom}}$ ,  $\text{MAX3}_1$ ,  $\text{MLRT}_1$ ,  $\text{RLRT}_1$ ,  $\text{MAX3}_2$ ,

Table 4: The simulated null distribution of  $Z_{\text{rec}}$  when  $p = .01$  based on 10,000 replications

		Total sample size $n$							
		100		500		1,000		10,000	
		$Z_{\text{rec}}$	%	$Z_{\text{rec}}$	%	$Z_{\text{rec}}$	%	$Z_{\text{rec}}^*$	%
		-1.134	37.14	-1.387	20.45	-1.706	8.65	1.655	3.18
		0.378	50.52	-0.277	41.85	-0.853	26.64	1.970	1.92
		1.890	12.34	0.832	28.85	0.000	33.31	2.286	1.21
				1.941	8.04	0.853	20.72	2.601	0.57
				3.051	0.81	1.706	8.50	2.916	0.28
						2.558	1.83	3.230	0.28
						3.411	0.33	3.546	0.08
						4.264	0.02	3.862	0.01
Type I									
Error**		0.1234		0.0885		0.1068		0.0753	

\* Only upper tail of the null distribution is reported.

\* Type I error is calculated when 1.645 is used as the critical value.

MLRT<sub>2</sub> RLRT<sub>2</sub>. We consider two situations: (a) the allele risk status is specified and (b) the allele risk status is unspecified. In the first situation,  $Z_{\text{rec}}$ ,  $Z_{\text{add}}$ ,  $Z_{\text{dom}}$ , MAX3<sub>1</sub>, MLRT<sub>1</sub> and RLRT<sub>1</sub> are compared. To fix point, allele  $A$  is assumed to be the high risk allele. In the second situation,  $Z_{\text{rec}}$ ,  $Z_{\text{add}}$ ,  $Z_{\text{dom}}$ , MAX3<sub>2</sub>, MLRT<sub>2</sub> and RLRT<sub>2</sub> are compared. In the comparison, the total sample size is taken as  $n = 200$ , the allele frequency  $p$  is taken as  $p = 0.01, 0.1, 0.3$  and  $0.5$ , and the following five alternative hypotheses are considered: (i)  $(r_1, r_2) = (1, 2)$ , a recessive model, (ii)  $(r_1, r_2) = (1.5, 1.5)$ , a dominant model, (iii)  $(r_1, r_2) = (1.5, 2.25)$ , a multiplicative model, (iv)  $(r_1, r_2) = (1.5, 2.0)$ , a additive model and (v)  $(r_1, r_2) = (1.3, 2.3)$ , a model which does not fall into the four categories of genetic models. The effective sample sizes  $n_1, n_2, n_3$  are taken as their conditional expectations under each of the five models. In all cases, the size of the test is fixed at  $\alpha = 0.05$  and the simulated exact critical value is used so that the type I error for all the tests are controlled at the same level. The power of each test statistic at a given alternative hypothesis is simulated with a simulation of size  $m = 10,000$ .

The simulated powers of the test statistics for situations (a) and (b) are given in Tables 5 and 6 respectively. By comparing the powers in these two tables, we observe the following features. (i) If the genetic model is correctly specified under the alternative hypothesis, the score statistic derived by assuming the specified genetic model is generally more powerful than the other test statistics. (ii) The performance of the MLRT and RLRT statistics in both situations are comparable. They are robust in the sense that their powers are only smaller than the score

Table 5: The powers at five specified alternatives of one-sided tests using six test statistics (total sample size  $n = 200$  and simulation size  $m = 10,000$ )

$p$	$(r_1, r_2)$	$Z_{\text{rec}}$	$Z_{\text{add}}$	$Z_{\text{dom}}$	MAX3 <sub>1</sub>	MLRT <sub>1</sub>	RLRT <sub>1</sub>
.01	(1.0, 2.0) <sup>1</sup>	0.2610	0.0552	0.0547	0.0676	0.2626	0.2543
	(1.5, 1.5) <sup>2</sup>	0.0995	0.8810	0.9032	0.8106	0.9119	0.8932
	(1.5, 2.25) <sup>3</sup>	0.1979	0.8881	0.8998	0.8136	0.9165	0.9032
	(1.5, 2.0) <sup>4</sup>	0.1692	0.8849	0.8965	0.8110	0.9127	0.8974
	(1.3, 2.3) <sup>5</sup>	0.2401	0.6005	0.6104	0.4711	0.6883	0.6620
.10	(1.0, 2.0)	0.6111	0.1962	0.0784	0.4810	0.4872	0.4941
	(1.5, 1.5)	0.0980	0.8563	0.8883	0.7936	0.8035	0.8094
	(1.5, 2.25)	0.4377	0.9232	0.8939	0.8635	0.8770	0.8845
	(1.5, 2.0)	0.3083	0.8992	0.8869	0.8439	0.8584	0.8689
	(1.3, 2.3)	0.4658	0.7608	0.6207	0.6602	0.6950	0.7252
.30	(1.0, 2.0)	0.9145	0.6804	0.1245	0.8834	0.8881	0.8915
	(1.5, 1.5)	0.0878	0.6717	0.8244	0.7047	0.7569	0.7200
	(1.5, 2.25)	0.6432	0.9465	0.8700	0.9163	0.9334	0.9299
	(1.5, 2.0)	0.4407	0.9030	0.8608	0.8627	0.8766	0.8700
	(1.3, 2.3)	0.8782	0.9255	0.6325	0.9228	0.9228	0.9253
.50	(1.0, 2.0)	0.9951	0.9436	0.1579	0.9824	0.9814	0.9819
	(1.5, 1.5)	0.1111	0.7533	0.9485	0.9028	0.9300	0.9321
	(1.5, 2.25)	0.8677	0.9481	0.7555	0.9214	0.9118	0.9179
	(1.5, 2.0)	0.6672	0.8645	0.6589	0.8187	0.8063	0.8182
	(1.3, 2.3)	0.9663	0.9752	0.5579	0.9686	0.9665	0.9667

<sup>1</sup> Recessive; <sup>2</sup> Dominant; <sup>3</sup> Multiplicative; <sup>4</sup> Additive; <sup>5</sup> Arbitrary

statistic optimal for the specified genetic model but are larger than or comparable with all the other test statistics. (3) The MLRT and RLRT statistics are generally more powerful than the MAX3 statistics. We deliberately used the word “generally” because there are a few discrepancies from the above statements in the two tables. These discrepancies might be caused by the simulation errors. It should be noticed that the performance of MAX3 is only slightly worse than MLRT and RLRT but it is easier to compute. If one wishes to compromise the power a little bit for the ease of computation, MAX3 is also a good choice for testing the gene-disease association in case-parents design.

## 5. Conclusions

We draw briefly our conclusions in this section. For testing the gene-disease association in case-parents trio designs, exact tests using any test statistics are feasible because of our great computational capacity and the fact that the distribution of the data under the null hypothesis of no association is completely

Table 6: The powers at five specified alternatives of two-sided tests using six test statistics (total sample size  $n = 200$  and simulation size  $m = 10,000$ )

$p$	$(r_1, r_2)$	$Z_{\text{rec}}$	$Z_{\text{add}}$	$Z_{\text{dom}}$	MAX3 <sub>2</sub>	MLRT <sub>2</sub>	RLRT <sub>2</sub>
.01	(1.0, 2.0) <sup>1</sup>	0.3290	0.0517	0.0565	0.0649	0.0447	0.1110
	(1.5, 1.5) <sup>2</sup>	0.2992	0.8105	0.8353	0.7129	0.8006	0.8601
	(1.5, 2.25) <sup>3</sup>	0.3005	0.8221	0.8380	0.7214	0.8104	0.8731
	(1.5, 2.0) <sup>4</sup>	0.2929	0.8241	0.8427	0.7281	0.8119	0.8736
	(1.3, 2.3) <sup>5</sup>	0.3264	0.4905	0.4995	0.3667	0.4572	0.5954
.10	(1.0, 2.0)	0.4621	0.1092	0.0591	0.3462	0.3374	0.3483
	(1.5, 1.5)	0.0565	0.7317	0.8131	0.7435	0.6924	0.7053
	(1.5, 2.25)	0.2800	0.8437	0.8283	0.8097	0.7714	0.7872
	(1.5, 2.0)	0.1911	0.7949	0.8088	0.7895	0.7476	0.7579
	(1.3, 2.3)	0.3188	0.6213	0.5068	0.6187	0.5322	0.5922
.30	(1.0, 2.0)	0.8692	0.5282	0.0533	0.8753	0.8317	0.8350
	(1.5, 1.5)	0.0714	0.5208	0.6723	0.5780	0.5888	0.5996
	(1.5, 2.25)	0.5569	0.8926	0.7314	0.8586	0.8737	0.8762
	(1.5, 2.0)	0.3548	0.7902	0.7213	0.7520	0.7735	0.7860
	(1.3, 2.3)	0.8317	0.8329	0.4310	0.8544	0.8426	0.8599
.50	(1.0, 2.0)	0.9779	0.8907	0.0563	0.9781	0.9704	0.9620
	(1.5, 1.5)	0.0697	0.6141	0.8894	0.6770	0.8991	0.8727
	(1.5, 2.25)	0.7343	0.8997	0.5194	0.8085	0.8821	0.8789
	(1.5, 2.0)	0.5185	0.7689	0.5071	0.6291	0.7554	0.7539
	(1.3, 2.3)	0.9385	0.9481	0.2985	0.9467	0.9548	0.9476

<sup>1</sup> Recessive; <sup>2</sup> Dominant; <sup>3</sup> Multiplicative; <sup>4</sup> Additive; <sup>5</sup> Arbitrary

known. In general, exact tests should be preferred to asymptotic tests, especially when the effective sample sizes are quite unbalanced, which could have been resulted from a small allele frequency. When the allele frequency is small, some of the asymptotic tests differ substantially from their corresponding exact tests, as demonstrated in our simulation study. The robust test statistics MLRT or RLRT should be preferred to all the other test statistics because of their efficiency and robustness unless one is particularly interested in testing a specific genetic model or is quite certain what specific genetic model might be under the alternative hypothesis.

### Acknowledgments

We thank one referee for careful reading and helpful suggestions that improved the presentation.

**References**

- Cleves, M. A. and Olson, J. M. and Jacobs, K. B. (1997). Exact transmission disequilibrium tests with multiallelic markers. *Genet. Epidemiol.* **14**, 337-347.
- Concove, W. J. (1980). *Practical Nonparametric Statistics*. 2nd ed. Wiley.
- Freidlin, B., Podgor, M. J. and Gastwirth, J. L. (1999). Efficiency robust tests for survival or ordered categorical data. *Biometrics* **55**, 883-886.
- Gastwirth, J. L. (1985). The use of maximin efficiency robust tests in combining contingency tables and survival analysis. *J. Am. Statist. Assoc.* **80**, 380-384.
- Lazzeroni, L. C. and Lange, K. (1998). A conditional inference framework for extending the transmission/disequilibrium test. *Hum. Hered.* **48**, 67-81.
- Schaid, D. J. (1999). Likelihoods and TDT for the case-parents design. *Genet. Epidemiol.* **16**, 250-260.
- Schaid, D. J. and Sommer, S. S. (1993). Genotype relative risks: Methods for design and analysis of candidate-gene association studies. *Am. J. Hum. Genet.* **53**, 1114-1126.
- Self, S. G. and Liang, K. Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *J. Am. Statist. Assoc.* **82**, 605-610.
- Spielman, R. S., McGinnis, R. E. and Ewens, W. J. (1993). Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am. J. Hum. Genet.* **52**, 506-516.
- Zheng, G., Freidlin, B. and Gastwirth, J. L. (2002). Robust TDT-type candidate-gene association tests. *Ann. Hum. Genet.* **66**, 145-155.
- Zheng, G., Chen, Z. and Li, Z. (2003). Tests for candidate-gene association using case-parents design. *Ann. Hum. Genet.* **67**, 589-597.

Received August 18, 2003; accepted December 10, 2003.

Zehua Chen  
Department of Statistics and Applied Probability  
National University of Singapore  
10 Kent Ridge Crescent  
Singapore 119260, Republic of Singapore  
stachenz@nus.edu.sg

Gang Zheng  
Office of Biostatistics Research  
National Heart, Lung and Blood Institute  
6701 Rockledge Drive, MSC 7938  
Bethesda, MD 20892, U.S.A.  
zhengg@nhlbi.nih.gov