

Testing Gene-Gene Interactions in the Case-Parents Design

Zhaoxia Yu

Department of Statistics, University of California, Irvine, Calif., USA

Key Words

Gene-gene interaction · Robustness · Population stratification · Linkage disequilibrium · Case-parents trios

Abstract

The case-parents design has been widely used to detect genetic associations as it can prevent spurious association that could occur in population-based designs. When examining the effect of an individual genetic locus on a disease, logistic regressions developed by conditioning on parental genotypes provide complete protection from spurious association caused by population stratification. However, when testing gene-gene interactions, it is unknown whether conditional logistic regressions are still robust. Here we evaluate the robustness and efficiency of several gene-gene interaction tests that are derived from conditional logistic regressions. We found that in the presence of SNP genotype correlation due to population stratification or linkage disequilibrium, tests with incorrectly specified main-genetic-effect models can lead to inflated type I error rates. We also found that a test with fully flexible main genetic effects always maintains correct test size and its robustness can be achieved with negligible sacrifice of its power. When testing gene-gene interactions is the focus, the test allowing fully flexible main effects is recommended to be used.

Copyright © 2011 S. Karger AG, Basel

Introduction

The high-throughput technologies have greatly improved our understanding of the genetic basis of human traits; however, analyzing data generated from these new technologies has been challenging researchers in many scientific disciplines. In particular, developing methods that can efficiently capture multi-locus associations, including gene-gene interactions, has been quite difficult. Since first introduced by Bateson [1], interaction has been defined from a wide variety of perspectives. In the field of statistical analysis, interactions between genes often refer to the deviations from multiplicity or additivity, depending on the parameterization. For example, for dichotomous traits, one definition of interaction is the deviations from multiplicative penetrance. In testing two-locus gene-gene interactions in the case-parents design, we define gene-gene interactions as the deviations from multiplicative genotype relative risks. Although these interactions do not necessarily coincide with biological interactions, testing such defined interactions can nevertheless capture associations that might be missed by single-locus analyses.

Among many designs to detect genetic associations, the case-parents design is generally regarded to be robust against population stratification. A common way to analyze data generated by this design is to match each af-

affected offspring with his or her pseudo-siblings formed by combinations of alleles that are not, but can be, transmitted to the affected offspring [2–4]. Following that, tests based on conditional logistic regressions [5] are used to assess associations. One example of such tests is the transmission disequilibrium test (TDT) [6], which can be derived from a conditional logistic regression by assuming multiplicative genotype relative risks [7]. The TDT is valid no matter whether the true genetic model is dominant, multiplicative, or recessive. However, it has been observed that tests of gene-environment interactions are not robust against population stratification when the main-genetic-effect models are incorrectly specified, and conditioning on parental genotypes does not guarantee correct test size [8–14]. While the potential bias that results from the incorrect specification of main-effect models and population stratification has been studied for tests of gene-environment interactions, to the best of our knowledge there has been no systematic evaluation of robustness against population stratification for tests of gene-gene interactions. In addition, currently existing tests for gene-environment interactions of family data usually require the independence between the tested single nucleotide polymorphism (SNP) and the environment factor conditional on parental mating type or its sufficient statistic [10]. In this article, linkage disequilibrium (LD) between SNPs is allowed.

To examine the impact of misspecification of the main-genetic-effect models on robustness and efficiency of tests for gene-gene interactions, a conditional logistic regression with three components is often considered: $Z_1(G_1)$, $Z_2(G_2)$, and $Z_3(G_1, G_2)$, where G_1 and G_2 represent the genotype variables at two loci, and Z_1 , Z_2 , Z_3 denote the main- and interaction-effect models [15–17]. For example, for a SNP with the non-risk allele ‘a’ and the risk allele ‘A’, the function $Z_1(\cdot)$ defined by $Z_1(aa) = 0$ and $Z_1(aA$ or $AA) = 1$ represents a dominant main-genetic-effect model at the SNP. Under this framework, the gene-gene interactions between the two loci are assessed by examining the coefficient of $Z_3(G_1, G_2)$. However, it is unknown how misspecification of the main genetic effects, i.e., when $Z_1(G_1)$ and $Z_2(G_2)$ do not agree with the true main-effect models, affects the statistical inference of the interactions between the two loci. Here we examine the robustness and efficiency of three conditional logistic regression methods in testing gene-gene interactions. To assess the potential bias caused by misspecifying the main-genetic-effect models in the presence of SNP genotype correlation due to population stratification or LD, we calculate the type I error rates of the three methods under the null hypothesis

of no gene-gene interactions. We then compare the efficiency of the three methods under different interaction models in the absence of SNP genotype correlation due to population stratification or LD.

Methods

Conditional Logistic Regression

One common way to analyze case-parents data is to first construct matched pseudo-siblings for each affected offspring conditioning on the parental genotypes of the offspring, and then analyze the resulted matched case-control data using conditional logistic regressions [3, 4, 7]. For two unlinked SNPs, there are 16 ways that a pair of parents can transmit alleles to their offspring, including the one that gives the genotype of the affected offspring. Because the 16 possible genotypes are equally likely under the null hypothesis of no genetic effects, conditional logistic analyses can be built by matching each affected offspring to his or her 15 pseudo-siblings [16, 17].

When the two tested SNPs are linked, the transmission of alleles from parents to offspring does not follow the Mendelian transmission, no matter whether there are genetic effects or not. Thus the computation of the conditional probabilities of offspring genotypes given parental genotypes requires the recombination fractions between loci. However, these recombination fractions are usually unknown and difficult to estimate, which complicates multi-locus analyses of linked loci in the case-parents design [16, 17]. As a result, the 1:15 matching strategy used for two unlinked loci cannot be used.

One less efficient but still valid way of constructing matched siblings in the presence of linkage is to match each affected offspring with only one pseudo-sibling using the alleles that are not transmitted to the affected offspring. For example, Kottli et al. [18] proposed a four-degree-of-freedom likelihood ratio test (LRT) of gene-gene interactions using this 1:1 matching. Their method models four parameters for the main genetic effects and another four parameters for the gene-gene interactions, and the method is valid in the presence of SNP genotype correlation due to population stratification or LD. Using this 1:1 matching to create pseudo-siblings, we can also derive other LRTs, with fully flexible or constrained main or interaction effects.

Suppose that we are testing whether there are gene-gene interactions between two SNPs underlying a disease. Let ‘a’ and ‘A’ denote the two alleles at the first SNP, and ‘b’ and ‘B’ denote the two alleles at the second SNP. Let g represent the numerically coded values for genotypes at the first SNP based upon the number of copies of allele ‘A’, and h be the numerically coded values of genotypes at the second SNP. We define subscripts f , m , o and i to be the indices for father, mother, offspring, and case-parents trio, respectively. Let $g_i = (g_{f,i}, g_{m,i}, g_{o,i})$ be the vector of the genotypes of the i -th trio at the first SNP, $h_i = (h_{f,i}, h_{m,i}, h_{o,i})$ be the vector at the second SNP. Let $H(g_i, h_i)$ be a set of genotypes for the offspring in the i -th trio. For two *unlinked* SNPs, the set H includes the genotype of the affected offspring and the other 15 combinations of parental alleles [16, 17]. For two *linked* SNPs, we match each affected offspring with the pseudo-sibling formed by the non-transmitted alleles; hence the set H consists of two genotypes, one for the affected offspring and the other for the pseudo-sibling.

Table 1. Genotype relative risks under the saturated model

	SNP2		
	bb	Bb	BB
SNP1			
aa	1	e^{β_1}	e^{β_2}
Aa	e^{α_1}	$e^{\alpha_1 + \beta_1 + \gamma_{11}}$	$e^{\alpha_1 + \beta_2 + \gamma_{12}}$
AA	e^{α_2}	$e^{\alpha_2 + \beta_1 + \gamma_{21}}$	$e^{\alpha_2 + \beta_2 + \gamma_{22}}$

Table 2. Multiplicative genotype relative risks

	SNP2		
	bb	Bb	BB
SNP1			
aa	1	e^{β_1}	e^{β_2}
Aa	e^{α_1}	$e^{\alpha_1 + \beta_1}$	$e^{\alpha_1 + \beta_2}$
AA	e^{α_2}	$e^{\alpha_2 + \beta_1}$	$e^{\alpha_2 + \beta_2}$

Table 3. Genotype relative risks with constrained interactions

	SNP2		
	bb	Bb	BB
SNP1			
aa	1	e^{β_1}	e^{β_2}
Aa	e^{α_1}	$e^{\alpha_1 + \beta_1 + \gamma}$	$e^{\alpha_1 + \beta_2 + 2\gamma}$
AA	e^{α_2}	$e^{\alpha_2 + \beta_1 + 2\gamma}$	$e^{\alpha_2 + \beta_2 + 4\gamma}$

Table 1 shows the genotype relative risks based on the two SNPs for a general two-locus model. This model corresponds to a saturated model. Using the conditional logistic regression for matched data, the likelihood function based on n case-parents trios is:

$$L_8 = \prod_{i=1}^n \frac{\exp\left(\alpha_1 I_{(g_{o,i}=1)} + \alpha_2 I_{(g_{o,i}=2)} + \beta_1 I_{(h_{o,i}=1)} + \beta_2 I_{(h_{o,i}=2)} + \sum_{k=1}^2 \sum_{l=1}^2 \gamma_{kl} I_{(g_{o,i}=k)} I_{(h_{o,i}=l)}\right)}{\sum_{(g'_{o,i}, h'_{o,i}) \in H(g_i, h_i)} \exp\left(\alpha_1 I_{(g'_{o,i}=1)} + \alpha_2 I_{(g'_{o,i}=2)} + \beta_1 I_{(h'_{o,i}=1)} + \beta_2 I_{(h'_{o,i}=2)} + \sum_{k=1}^2 \sum_{l=1}^2 \gamma_{kl} I_{(g'_{o,i}=k)} I_{(h'_{o,i}=l)}\right)}$$

where the subscript ‘8’ in L_8 reflects the number of parameters in the saturated model. α_1 and α_2 are the coefficients of the main genetic effects at the first SNP, β_1 and β_2 are the coefficients of the main genetic effects at the second SNP, and γ_{kl} for $k = 1, 2$ and $l = 1, 2$ are the coefficients of the interactions between the two SNPs. $I_{(\cdot)}$ is an indicator function, for example, $I_{(g_{o,i}=1)}$ is 1 if $g_{o,i} = 1$ and is 0 if otherwise. The numerator in the function is equal to the

genotype relative risk for a subject with genotypes $g_{o,i}$ at the first SNP and $h_{o,i}$ at the second SNP.

Likelihood Ratio Tests

Different models can be constructed by placing restrictions on the main and interaction effects. By comparing different models, we can develop LRTs to assess the significance of gene-gene interactions.

Test 1: LRT_4

The first test we consider compares the saturated model to a main-effect-only model. As mentioned earlier, in this article we define gene-gene interactions as the deviations from multiplicative genotype relative risks. Under the null hypothesis of no gene-gene interaction, $\gamma_{kl} = 0$ for $k = 1, 2$ and $l = 1, 2$, and the genotype relative risks can be formulated by four parameters, as shown in table 2. The likelihood function for this reduced model therefore is

$$L_4 = \prod_{i=1}^n \frac{\exp\left(\alpha_1 I_{(g_{o,i}=1)} + \alpha_2 I_{(g_{o,i}=2)} + \beta_1 I_{(h_{o,i}=1)} + \beta_2 I_{(h_{o,i}=2)}\right)}{\sum_{(g'_{o,i}, h'_{o,i}) \in H(g_i, h_i)} \exp\left(\alpha_1 I_{(g'_{o,i}=1)} + \alpha_2 I_{(g'_{o,i}=2)} + \beta_1 I_{(h'_{o,i}=1)} + \beta_2 I_{(h'_{o,i}=2)}\right)}$$

where the subscript ‘4’ in L_4 reflects the number of parameters in the reduced model. To test the significance of the interaction parameters, we compare the saturated model L_8 to the multiplicative model L_4 . The resulted test has four degrees of freedom, and we denote it as LRT_4 . A similar four-degree-of-freedom test was used in Marchini et al. [19] for case-control data.

Test 2: LRT_1

The second test we consider compares a restricted-interaction-effect model to the main-effect-only model. To reduce the number of degrees of freedom and therefore to improve power, tests with parsimonious parameterizations have been studied in the past. For example, one test is to use a model with fully flexible main effects but constrained interactions [20], as shown in table 3. The resulted five-parameter likelihood function is

$$L_5 = \prod_{i=1}^n \frac{\exp\left(\alpha_1 I_{(g_{o,i}=1)} + \alpha_2 I_{(g_{o,i}=2)} + \beta_1 I_{(h_{o,i}=1)} + \beta_2 I_{(h_{o,i}=2)} + \gamma g_{o,i} h_{o,i}\right)}{\sum_{(g'_{o,i}, h'_{o,i}) \in H(g_i, h_i)} \exp\left(\alpha_1 I_{(g'_{o,i}=1)} + \alpha_2 I_{(g'_{o,i}=2)} + \beta_1 I_{(h'_{o,i}=1)} + \beta_2 I_{(h'_{o,i}=2)} + \gamma g'_{o,i} h'_{o,i}\right)}$$

Again, to test the significance of the interaction parameter, we compare L_5 to the one with no interaction, i.e. L_4 . The resulted test has one degree of freedom, and we denote it as LRT_1 .

Test 3: $LRT_{1,mul}$

The third test we consider compares a restricted-main-and-interaction-effect model to a restricted-main-effect-only model. When testing the association between a single SNP and a disease using the case-parents design, one often uses the TDT, which tests for allelic effects. This is equivalent to assuming that the effects of alleles on genotype relative risks are multiplicative. As the multiplicative model is located between the two extreme models (the

Table 4. Genotype relative risks with constrained main effects and constrained interactions

	SNP2		
	bb	Bb	BB
SNP1			
aa	1	e^β	$e^{2\beta}$
Aa	e^α	$e^{\alpha + \beta + \gamma}$	$e^{\alpha + 2\beta + 2\gamma}$
AA	$e^{2\alpha}$	$e^{2\alpha + \beta + 2\gamma}$	$e^{2\alpha + 2\beta + 4\gamma}$

dominant model and the recessive model), the TDT is convenient and robust, and it is common to use this test when the underlying genetic model is not known. When the genotype relative risks have an approximately multiplicative pattern, the TDT is statistically more powerful than tests with two degrees of freedom, as it uses fewer degrees of freedom. This strategy has been applied to reduce the number of degrees of freedom in testing gene-gene interactions [21] for case-control data. As LRT_1 uses four nuisance parameters for the main effects, one can restrict the main effects to be multiplicative and therefore reduce the number of nuisance parameters from four to two. The resulted test, $LRT_{1,mul}$, compares the following two likelihood functions,

$$L_3 = \prod_{i=1}^n \frac{\exp(\alpha g_{o,i} + \beta h_{o,i} + \gamma g_{o,i} h_{o,i})}{\sum_{(g'_{o,i}, h'_{o,i}) \in H(g_i, h_i)} \exp(\alpha g'_{o,i} + \beta h'_{o,i} + \gamma g'_{o,i} h'_{o,i})}$$

$$L_2 = \prod_{i=1}^n \frac{\exp(\alpha g_{o,i} + \beta h_{o,i})}{\sum_{(g'_{o,i}, h'_{o,i}) \in H(g_i, h_i)} \exp(\alpha g'_{o,i} + \beta h'_{o,i})}$$

The corresponding genotype relative risks are shown in tables 4 and 5.

Population Stratification and LD

The population stratification we consider assumes that case-parents trios are sampled from two random mating populations. In population 1, the allele frequencies are fixed, $\Pr(A) = 0.2$, and $\Pr(B) = 0.2$, while in population 2, the allele frequencies of A and B, denoted by p_2 and q_2 , respectively, vary from 0.1 to 0.9. We assume that the two SNPs are *unlinked*. We sample 500 trios from population 1, and another 500 trios from population 2.

Since population stratification may create an artificial correlation between two unlinked SNPs in a mixed population, it is reasonable to suspect that, for *linked* SNPs, test sizes might also be affected by the misspecification of the main-genetic-effect models, even if the data are sampled from a random mating population. To mimic linkage between two SNPs, we fix the allele frequencies to be 0.2, i.e., $\Pr(A) = \Pr(B) = 0.2$, and vary the LD coefficient D from 0 to its maximum value of 0.16, where D is defined as

$$D = \Pr(\text{haplotype } AB) - 0.2 \times 0.2.$$

In our computation, we ignore the recombination from parental haplotypes to offspring haplotypes.

Table 5. Genotype relative risks with constrained main effects and no interactions

	SNP2		
	bb	Bb	BB
SNP1			
aa	1	e^β	$e^{2\beta}$
Aa	e^α	$e^{\alpha + \beta}$	$e^{\alpha + 2\beta}$
AA	$e^{2\alpha}$	$e^{2\alpha + \beta}$	$e^{2\alpha + 2\beta}$

To examine the impact of the misspecification of the main-genetic-effect models on testing gene-gene interactions in the presence of SNP genotype correlation due to population stratification or LD, we consider the following main-effect-only models,

- (1) $\alpha_1 = 0, \alpha_2 = \ln 2; \beta_1 = 0, \beta_2 = \ln 2$
(recessive-recessive model)
- (2) $\alpha_1 = 0, \alpha_2 = \ln 2; \beta_1 = \ln 2, \beta_2 = \ln 4$
(recessive-multiplicative model)
- (3) $\alpha_1 = \alpha_2 = \ln 2; \beta_1 = \beta_2 = \ln 2$
(dominant-dominant model)
- (4) $\alpha_1 = \alpha_2 = \ln 2; \beta_1 = \ln 2, \beta_2 = \ln 4$
(dominant-multiplicative model)
- (5) $\alpha_1 = \ln 2, \alpha_2 = \ln 4; \beta_1 = \ln 2, \beta_2 = \ln 4$
(multiplicative-multiplicative model)

In all of those models, $\gamma_{11} = \gamma_{12} = \gamma_{21} = \gamma_{22} = 0$.

Genetic Interaction Models to Compute Power

Since robustness is usually achieved at the sacrifice of some efficiency, it is important to investigate whether tests that are robust against population stratification or LD have substantially decreased efficiency, under the condition that there is no population stratification or LD. To do so, we compute the power of the three LRTs under four gene-gene interaction models.

In the first interaction model we assume that the main genetic effects are multiplicative at both SNPs and the four interaction parameters γ_{ki} are not all the same. Specifically, we let $\gamma_{11} = 0$ and $\gamma_{12} = \gamma_{21} = \gamma_{22} = \gamma$. Then the corresponding genotype relative risk for a subject with genotype (g, h) is

$$\frac{\exp(g \ln(1.2) + h \ln(1.2) + \gamma [I_{(g=2)} I_{(h=1)} + I_{(h=1)} I_{(g=2)} + I_{(g=2)} I_{(h=2)}])}{I_{(g=2)} I_{(h=2)}} \quad (\text{Model 1})$$

In all the other three models, we assume that the gene-gene interactions can be modeled by a single parameter. What distinguishes those three models from each other are the restrictions on the main genetic effects. Specifically, the dominant, recessive, and multiplicative main-effect models at both SNPs are assumed in Model 2, Model 3, and Model 4, respectively. The genotype relative risks of Models 2–4 are as follows,

$$\frac{\exp(I_{(g=1)} \ln(2) + I_{(g=2)} \ln(2) + I_{(h=1)} \ln(2) + I_{(h=2)} \ln(2) + \gamma gh)}{I_{(g=1)} I_{(h=1)}} \quad (\text{Model 2})$$

$$\frac{\exp(I_{(g=2)} \ln(2) + I_{(h=2)} \ln(2) + \gamma gh)}{I_{(g=2)} I_{(h=2)}} \quad (\text{Model 3})$$

$$\frac{\exp(g \ln(1.2) + h \ln(1.2) + \gamma gh)}{I_{(g=1)} I_{(h=1)}} \quad (\text{Model 4})$$

In particular, Model 3 is the same as the two-locus interaction multiplicative effects model in Marchini et al. [19], and Models 2 and 4 differ from Model 3 in the assumptions of the main genetic effects. For Models 1, 2, and 3, we compute statistical power using 1,000 trios drawn from a random mating population with $\Pr(A) = \Pr(B) = 0.2$. To show the subtle difference between the two one-degree-of-freedom tests for Model 4, we reduce the number of trios to 200. We vary the parameter γ to depict the patterns of power increase with the gene-gene interactions. We also assume that the two SNPs are unlinked.

Computation of Asymptotic Power and Type I Error Rates

To compute the power or type I error rates, for a given main-effect-only model and a set of parameters at two SNPs, we first compute the expected number of different trio types, where a trio type is a specific combination of genotypes of the father, mother, and the affected offspring in a family. Using these expected numbers, we then maximize the five likelihoods $L_2, L_3, L_4, L_5,$ and L_8 , and we denote their maximum values by $\max L_2, \max L_3, \max L_4, \max L_5,$ and $\max L_8$, respectively. Following that, we calculate the non-centrality parameters for the three LRTs as

$$\begin{aligned} ncp_4 &= -2[\ln(\max L_4) - \ln(\max L_8)], \\ ncp_1 &= -2[\ln(\max L_4) - \ln(\max L_5)], \\ ncp_{1,mul} &= -2[\ln(\max L_2) - \ln(\max L_3)], \end{aligned}$$

where $ncp_4, ncp_1,$ and $ncp_{1,mul}$ are for $LRT_4, LRT_1,$ and $LRT_{1,mul}$, respectively. These non-centrality parameters are then used to compute the power or type I error rates using asymptotic χ^2 distributions [17, 22]. For example, the power or type I error rate of LRT_4 is defined as $\Pr(X_4(ncp_4) > \chi_{4,0.95}^2)$, where $X_4(ncp_4)$ is a χ^2 -distributed random variable with four degrees of freedom and a non-centrality parameter of ncp_4 , and $\chi_{4,0.95}^2$ is the 95th percentile of the χ^2 distribution with four degrees of freedom. The power or type I error rates of the other two tests are computed similarly. Note that using the 95th percentile is equivalent to using a nominal p value cutoff of 0.05.

Results

Robustness

We found that both LRT_4 and LRT_1 always have correct type I error rates for all the five main-effect-only models, even in the presence of a SNP genotype correlation due to population stratification or LD. The type I error rates of $LRT_{1,mul}$ for two unlinked loci are illustrated in figure 1 using 3-D plots. Because $LRT_{1,mul}$ correctly specifies the main effects under the multiplicative-multiplicative model, it has correct test size (results not shown). However, its type I error rates are inflated in the other four main-effect-only models, with the error rates increase with the difference in allele frequencies between the two populations. For example, when the allele frequencies between the two populations are different by 0.7 at both loci, the type I error rates are greater than 0.95 under the re-

cessive-recessive and dominant-dominant models. The inflation in the type I error rates of $LRT_{1,mul}$ is also noticeable under a moderate difference in allele frequencies between the two populations. For example, when $p_2 = q_2 = 0.4$, the type I error rate is about 0.13 under the dominant-dominant model and 0.11 under the recessive-recessive model. Although the inflation is less severe when one of the SNPs has a multiplicative main effect, the maximum type I error rates are greater than 0.5 for both the recessive-multiplicative and dominant-multiplicative models, as shown in figure 1b and d.

Another interesting observation from figure 1 is that $LRT_{1,mul}$ has correct size when $p_2 = 0.2$ or $q_2 = 0.2$. To see this better, we plotted the type I error rates of $LRT_{1,mul}$ when both SNPs have a recessive genetic effect in figure 2. This figure clearly shows that the type I error rates are equal to the nominal p value cutoff of 0.05 when the allele frequencies between the two populations agree at at least one SNP. Suppose that $c \times 100\%$ trios are sampled from population 1 and $(1 - c) \times 100\%$ are sampled from population 2. Based on our results, the condition that $LRT_{1,mul}$ has correct size is $c(1 - c)(p_1 - p_2)(q_1 - q_2) = 0$, where p_i is the allele frequency of the allele 'A' in population i and q_i is the allele frequency of allele 'B' in population i . We can show that this is equivalent to no correlation between the two SNPs in the mixed population. When the main effects are not multiplicative and need to be presented using two parameters, the model with multiplicative restriction on the main effects underfits the true model. When fitting a linear model, the independence of regressors ensures unbiasedness of the least square estimates of the regression coefficients. However, this type of collapsibility is not guaranteed in nonlinear regressions [23]. Nevertheless, our results indicate that there are no inflated type I error rates when two unlinked SNPs show no LD in the mixed population.

When the main-effect models are incorrectly specified, the type I error rates are also inflated if the SNPs under study are in LD, even if population stratification is not present. As shown in figure 3, when the main effects are not multiplicative at at least one SNP, the type I error rates of $LRT_{1,mul}$ using the 1:1 matching increases with the LD coefficient. When the main effects at both loci are dominant or recessive, a moderate LD between the two SNPs can lead to large type I error rates for $LRT_{1,mul}$. Note that the computation was done under the assumption of no population stratification, therefore, unless two loci are known to be unlinked, tests with fully flexible main effects should be used to test gene-gene interactions, no matter whether population stratification is a concern or not.

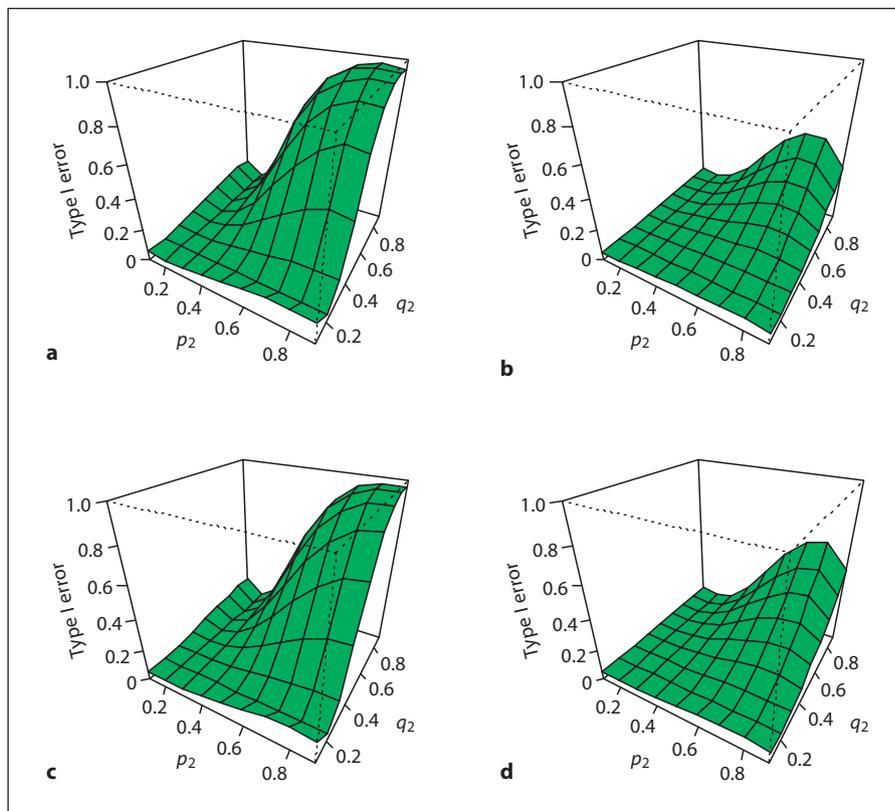


Fig. 1. Type I error rates of $LRT_{1,mul}$ as a function of allele frequencies in population 2. **a** The main effects are recessive at both SNPs; **b** the main effect is recessive at SNP 1 and multiplicative at SNP 2; **c** the main effects are dominant at both SNPs; **d** the main effect is dominant at SNP 1 and multiplicative at SNP 2.

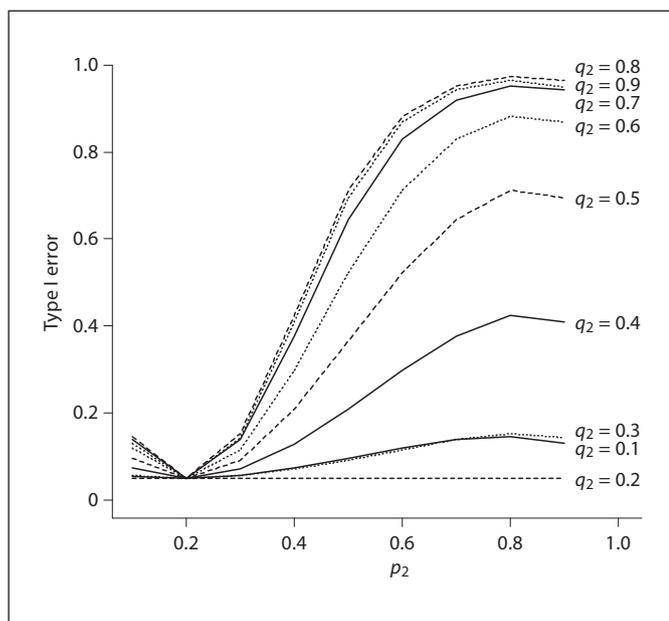


Fig. 2. Type I error rates of $LRT_{1,mul}$ as a function of allele frequencies in population 2 when the main effects at both SNPs are recessive.

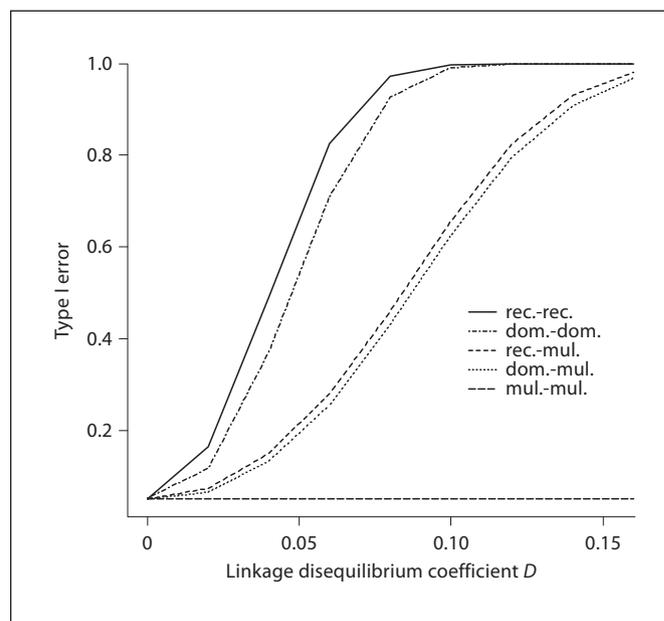


Fig. 3. Type I error rates of $LRT_{1,mul}$ as a function of the LD coefficient. rec. = recessive; dom. = dominant; mul. = multiplicative.

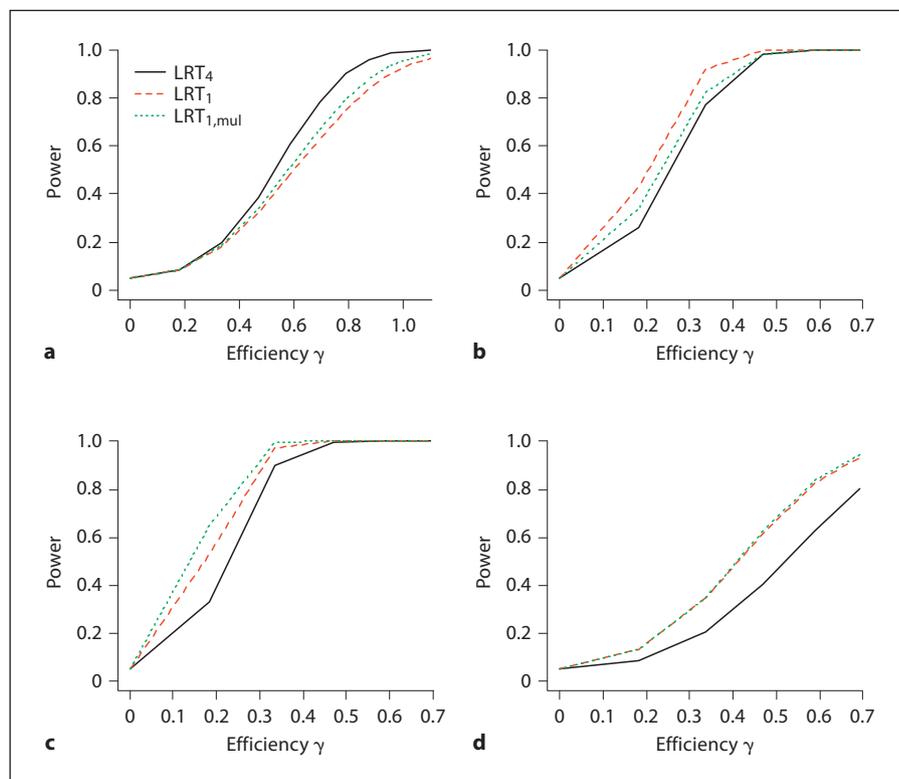


Fig. 4. Efficiency of the three likelihood ratio tests under four gene-gene interaction models.

Again, these results demonstrate that $LRT_{1,mul}$ has inflated type I error rates when two tested SNPs show correlation. Combined with the results when there is population stratification but no LD, we can conclude that $LRT_{1,mul}$ has inflated type I error rates in the presence of a SNP genotype correlation due to population stratification or LD.

Efficiency

The efficiency of the three tests under the four different models of gene-gene interactions is shown in figure 4. Because the four interaction parameters of Model 1 are not all the same, the test with fully flexible interaction effects, i.e. LRT_4 , provides the highest power. For the one-degree-of-freedom test, the power is slightly higher when main effects are assumed to be multiplicative ($LRT_{1,mul}$) than when they are assumed to be fully flexible (LRT_1).

For Models 2, 3 and 4, because of the restriction on the gene-gene interactions, the two one-degree-of-freedom tests, LRT_1 and $LRT_{1,mul}$, are parsimonious in modeling the gene-gene interactions and expected to be more powerful than the four-degree-of-freedom test LRT_4 . As shown in figure 4b–d, LRT_1 and $LRT_{1,mul}$ are indeed more powerful than the four-degree-of-freedom test LRT_4 . Although in the absence of a SNP genotype correlation due

to population stratification or LD, misspecifying the main effects does not lead to inflated type I error rates, it nevertheless affects the power, with the relative power of LRT_1 and $LRT_{1,mul}$ depending on the true main and interaction effects.

As shown above, LRT_1 is robust against population stratification or LD, while $LRT_{1,mul}$ is not. One concern is how much efficiency of LRT_1 is traded-off to gain robustness. To assess that, we compared its power to that of $LRT_{1,mul}$, as $LRT_{1,mul}$ is the most powerful test when there is no population stratification and the true main genetic effects are multiplicative (Model 4). The results are shown in figure 4d. The difference in power between the two tests is negligible, which demonstrates that the sacrifice of LRT_1 in efficiency to gain robustness against population stratification is probably small.

Discussion

In this article we examined the robustness and efficiency of three tests for two-locus gene-gene interactions in the case-parents design. Our results show that in the presence of a SNP genotype correlation due to popula-

tion stratification or LD, the parsimonious test $LRT_{1,mul}$ with multiplicative main effects is likely to have inflated type I error rates. In contrast, the one-degree-of-freedom test LRT_1 with fully flexible main effects is robust against a SNP genotype correlation. With respect to efficiency, there are situations favoring either of the two tests. However, in the situation that surely favors the most parsimonious test $LRT_{1,mul}$, we found that the power of $LRT_{1,mul}$ is similar to that of LRT_1 . This indicates that modeling two extra nuisance parameters does not lead to substantial loss in efficiency. In genome-wide studies where a large number of loci are examined, the main effect of each SNP is usually unknown. Therefore, the one-degree-of-freedom test with fully flexible main effects, LRT_1 , is recommended to detect potential gene-gene interactions.

Our examination of robustness against a SNP genotype correlation due to population stratification or LD is based upon a conditional logistic regression framework, and we conclude that correctly specifying the main-effect models is crucial to ensure correct test size. This conclusion may also apply to other tests with incorrectly specified main-genetic-effect models. For example, one may first compute the transmission rate of an allele at one SNP for the affected children in each of the three genotype groups at the second SNP, and then examine whether the three transmission rates are the same. The resulted test is not robust against a SNP genotype correlation, as it is equivalent to assuming a multiplicative main effect at the first locus. Therefore, as a test of gene-gene interactions, it is not robust against a SNP genotype correlation. The conclusion can also be extended to settings where genetic interactions of higher orders are focused on. For example, when a three-way genetic interaction is the primary interest, parsimonious parameterization of lower-order interactions may lead to incorrect inference in the presence of a SNP genotype correlation due to population stratification or LD.

In the study of robustness against population stratification, we assumed different populations share the same genotype relative risks and they are different only in allele frequencies. This is a strong assumption that might not always hold. When data are known to be from genetically distinct populations, such as African-American families and European-American families, it is advised to analyze these two groups separately and then examine whether equal genotype relative risks between the two populations can be assumed.

Our article focuses on the case-parents design. During the preparation of the article, we noticed some recent

studies on testing gene-environment interactions for the case-only and the case-control designs [24–26]. Tchetgen Tchetgen and Kraft [24] proposed to use a robust sandwich estimator of variance to protect inflated type I error rates. The articles by Mukherjee and Chatterjee [25] and by Li and Conti [26] aim to improve power by combining a method that is robust against the violation of model assumptions and another method that is more powerful when the underlying assumption is satisfied. Applying these ideas to the case-parents design might improve the efficiency of testing gene-gene interactions.

While the misspecification of the main-genetic-effect models may lead to bias in testing gene-gene interactions, the case-parents design is still an efficient and robust design. When there is no association between the disease and any combinations of the alleles at the two loci, all the three LRTs have correct type I error rates. And using the parsimonious test with multiplicative main effects, i.e. $LRT_{1,mul}$, can provide information about the joint effects of multiple SNPs. It is under the situation that when there are some types of associations, not necessarily deviations from multiplicative genotype relative risks, misspecification of main-effect models may lead to inflated type I error rates in testing deviations from multiplicity. Thus, when testing gene-gene interactions is the focus of a study, it is advised to avoid the parsimonious test that has multiplicative constraints on both the main and the interaction effects.

Acknowledgements

We thank the reviewers for their helpful comments. The author is grateful to Daniel Schaid, and Li Deng for helpful discussions. The research was supported in part by grant NIH/R01 HG004960.

References

- 1 Bateson W: Mendel's Principles of Heredity. Cambridge, Cambridge University Press, 1909.
- 2 Falk CT, Rubinstein P: Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. *Ann Hum Genet* 1987;51:227–233.
- 3 Self SG, Longton G, Kopecky KJ, Liang KY: On estimating HLA/disease association with application to a study of aplastic anemia. *Biometrics* 1991;47:53–61.
- 4 Schaid DJ: General score tests for associations of genetic markers with disease using cases and their parents. *Genet Epidemiol* 1996;13:423–449.

- 5 Breslow NE, Day NE: *Statistical Methods in Cancer Research, vol 1: The Analysis of Case-Control Studies*. Lyon, World Health Organization, 1980.
- 6 Spielman RS, McGinnis RE, Ewens WJ: Transmission test for linkage disequilibrium: The insulin gene region and insulin-dependent diabetes mellitus (iddm). *Am J Hum Genet* 1993;52:506–516.
- 7 Schaid DJ: Likelihoods and TDT for the case-parents design. *Genet Epidemiol* 1999;16:250–260.
- 8 Umbach DH, Weinberg CR: The use of case-parent triads to study joint effects of genotype and exposure. *Am J Hum Genet* 2000;66:251–261.
- 9 Shi M, Umbach DM, Weinberg CR: Testing haplotype-environment interactions using case-parent triads. *Hum Hered* 2010;70:23–33.
- 10 Hoffmann TJ, Lange C, Vansteelandt S, Laird NM: Gene-environment interaction tests for dichotomous traits in trios and sibships. *Genet Epidemiol* 2009;33:691–699.
- 11 Vansteelandt S, DeMeo DL, Lasky-Su J, Smoller JW, Murphy AJ, McQueen M, Schneiter K, Celedon JC, Weiss ST, Silverman EK, Lange C: Testing and estimating gene-environment interactions in family-based association studies. *Biometrics* 2008;64:458–467.
- 12 Chatterjee N, Kalaylioglu Z, Carroll RJ: Exploiting gene-environment independence in family-based case-control studies: increased power for detecting associations, interactions and joint effects. *Genet Epidemiol* 2005;28:138–156.
- 13 Cordell HJ: Estimation and testing of gene-environment interactions in family-based association studies. *Genomics* 2009;93:5–9.
- 14 Moerkerke B, Vansteelandt S, Lange C: A doubly robust test for gene-environment interaction in family-based studies of affected offspring. *Biostatistics* 2010;11:213–225.
- 15 Thomas DC: *Statistical Methods in Genetic Epidemiology*. Oxford, Oxford University Press, 2004.
- 16 Cordell HJ, Clayton DG: A unified stepwise regression procedure for evaluating the relative effects of polymorphisms within a gene using case/control or family data: application to HLA in type 1 diabetes. *Am J Hum Genet* 2002;70:124–141.
- 17 Gauderman WJ: Sample size requirements for association studies of gene-gene interaction. *Am J Epidemiol* 2002;155:478–484.
- 18 Kotti S, Bickeboller H, Clerget-Darpoux F: Strategy for detecting susceptibility genes with weak or no marginal effect. *Hum Hered* 2007;63:85–92.
- 19 Marchini J, Donnelly P, Cardon LR: Genome-wide strategies for detecting multiple loci that influence complex diseases. *Nat Genet* 2005;37:413–417.
- 20 Vanderweele TJ, Laird NM: Tests for compositional epistasis under single interaction-parameter models. *Ann Hum Genet* 2010.
- 21 Millstein J, Conti DV, Gilliland FD, Gauderman WJ: A testing framework for identifying susceptibility genes in the presence of epistasis. *Am J Hum Genet* 2006;78:15–27.
- 22 Wang S, Zhao H: Sample size needed to detect gene-gene interactions using association designs. *Am J Epidemiol* 2003;158:899–914.
- 23 Gail MH, Wieand S, Piantadosi S: Biased estimates of treatment effect in randomized experiments with nonlinear regressions and omitted covariates. *Biometrika* 1984;71:431–444.
- 24 Tchetgen Tchetgen EJ, Kraft P: On the robustness of tests of genetic associations incorporating gene-environment interaction when the environmental exposure is misspecified. *Epidemiology* 2011;22:257–261.
- 25 Mukherjee B, Chatterjee N: Exploiting gene-environment independence for analysis of case-control studies: an empirical Bayes-type shrinkage estimator to trade-off between bias and efficiency. *Biometrics* 2008;64:685–694.
- 26 Li DL, Conti DV: Detecting gene-environment interactions using a combined case-only and case-control approach. *Am J Epidemiol* 2009;169:497–504.