Tests in a Case-Control Design Including Relatives

STEFANIE BIEDERMANN and EVA NAGEL

Department of Mathematics

Ruhr-University Bochum

AXEL MUNK and HAJO HOLZMANN

Department of Mathematics

Georg-August-University Göttingen

ANSGAR STELAND

Department of Mathematics

Ruhr-University Bochum

Running Heading: Case-control studies with relatives

Abstract

We present a new approach to handle dependencies within the general framework of case-control designs, illustrating our approach by a particular application from the field of genetic epidemiology. The method is derived for parent-offspring trios, which will later be relaxed to more general family structures. For applications in genetic epidemiology we consider tests on equality of allele frequencies among cases and controls utilizing well-known risk measures to test for independence of phenotype and genotype at the observed locus. These test statistics are derived as functions of the entries in the associated contingency table containing the numbers of the alleles under consideration in the case and the control group. We find the joint asymptotic distribution of these entries, which enables us to derive critical values for any test constructed on this basis. A simulation study reveals the finite sample behavior of our test statistics.

Key words: association tests, contingency tables, dependent data, risk measures

1 Introduction

The aim of this article is to give a new approach to handle data of relatives within a case-control study. We first introduce our method for the situation of parent-child trios with no missing observations, and show later in the discussion how to modify the tests for more general data situations such as more complex pedigrees and missing data. Our approach is based on commonly used tests to analyze contingency tables, namely the odds ratio, the attributable risk, and the relative risk. We derive the asymptotic distributions of these test statistics by establishing a general result on the asymptotic normality of the (appropriately standardized) entries of the contingency table. To illustrate our approach we refer to a particular application in genetic epidemiology where the null hypothesis of identical cell probabilities for cases and controls is equivalent to the independence of a specific genotype from the phenotype.

Most methods to detect association with disease in the literature are either pure population based (case-control samples with unrelated individuals only) or pure family based, where parental (founder) genotypes are used to construct tests of association that are entirely contained within the family and thus robust to population stratification. The TDT, for example, which was introduced by Spielman, McGinnis & Ewens (1993), uses nontransmitted parental alleles of a case as a control sample, analyzing the data by a McNemar statistic. Case-control studies, however, tend to have a better power than pure family based procedures; see, e.g., McGinnis, Shifman & Darvasi (2002). Risch & Teng (1998), Teng & Risch (1999), and Risch (2000) point out that using pedigrees including many cases will lead to an increase in power, due to higher expected frequencies of disease-susceptibility alleles in pedigrees with multiple cases compared to the frequencies of these alleles in population based cases. Moreover, utilizing all cases in a pedigree rather than just one per pedigree improves power by increasing the effective sample size. Several authors have therefore derived methods to combine the benefits of population based and family based approaches.

Let us briefly discuss some recent contributions of this kind. Purcell, Sham & Daly (2005) identify the ignorance of parental phenotypes as one possible source for the lower efficiency of pure family based methods. Fitting a variance component model for nuclear families (both parents and an arbitrary number of children), they break phenotypic association with genotype into two components; a within component, robust to stratification, in which association is examined within each family, and a between component, where association is examined across families. In their terminology, the TDT consists of a within-family component only, whereas a case-control study with unrelated individuals is entirely a between-family test. Risch & Teng (1998) propose a method, which is applicable to several different designs including sibships with parents, sibships without parents and unrelated controls, using DNA pooling. In Slager & Schaid (2001) the test for trend in proportions introduced by Armitage (1955) is extended to general family data, whereas Böhringer & Steland (2006) provide a very accurate version of the likelihood for parent-offspring duos. Whittemore & Tu (2000) present a class of score statistics accommodating genotypes of both unrelated individuals and families with arbitrary structures. Epstein *et al.* (2005) discuss the issue of sampling both parental and unrelated controls, modifying the approximate analysis approach of Nagelkerke *et al.* (2004), who had found under quite restrictive model assumptions that analyzing data from triads and unrelated controls together yields a higher power than separate analyses. The approach of Epstein *et al.* (2005) allows for more flexibility of modelling allele effects, and less restrictive assumptions are needed, without losing power compared with Nagelkerke *et al.* (2004). Browning *et al.* (2005) account for correlations between individuals in a case-control design by calculating an optimal weight for each individual based on IBD sharing probabilities as in McPeek, Wu & Ober (2004), who introduced these optimal weights in the context of finding the best linear unbiased estimator for allele frequencies of data where the relationships among the sampled individuals are specified by a large, complex pedigree such as in isolated founder populations, which makes the use of maximum likelihood estimation impractical.

All the methods discussed above are in fact likelihood ratio tests. Since likelihood approaches require a full specification of the genetic model, we propose nonparametric level α tests to test for independence of genotype and phenotype. This article is organized as follows. Section 2 deals with a more detailed discussion of the genetic model under consideration. The main results of our research are then stated in Section 3, where the asymptotic distributions of the test statistics are derived. Section 4 provides various important extensions of our method, e.g., polygenic disorders, different inheritance models or strategies to deal with population stratification. To assess the finite sample behavior of the tests in terms of sample size and power, we conduct a simulation study under various scenarios of practical interest in Section 5. The simulations analyze whether including relatives increases statistical power, and also provide a comparison with the TDT, which has become a kind of "benchmark" test among the family based procedures. The simulations demonstrate on the one hand that including relatives in the study always leads to a significant increase in power, and on the other hand that the nonparametric tests have increased power compared to the TDT for virtually all scenarios under consideration under the assumption of no population stratification. The discussion in Section 6 provides a more detailed insight into the problem of dealing with different family structures and the situation of missing data. The proofs of our results, finally, are deferred to an appendix.

2 Genetic model and assumptions

We assume that there are two biallelic loci $L1$ and $L2$, with $L1$ being the candidate locus where data are observed, whereas $L2$ denotes the true but unknown causal locus. The goal of the study is to ascertain if there is significant evidence for linkage disequilibrium between $L1$ and $L2$ by comparing the allele frequencies of a marker allele at $L1$ in the case and the control group, respectively. We denote the possible alleles at both $L1$ and $L2$ by A and $\bar{A}$, where the causative allele for the disease at $L2$ will be termed A in what follows, whereas A at $L1$ stands for that allele at the marker locus which is suspected to be associated with the disease. Using the same notation for the alleles at $L1$ and $L2$ does not imply that the alleles at the different loci would consist of the same sequence of nucleotides, but is merely for notational convenience. For simplicity of motivation and derivation of results, we further suppose that the underlying inheritance model is dominant, i.e. for any individual the probability of being affected given there is at least one allele A present at locus $L2$ is equal to the penetrance f with $0 < f \leq 1$, whereas individuals without any allele A at $L2$ are affected with probability zero. At first glance it seems that our assumptions concerning the number of candidate and disease loci and the mode of inheritance imposed at this stage are quite restrictive, but we show in Section 4 how to apply our approach to an arbitrary number of markers and predisposing loci as well as different modes of inheritance.

We consider the following study setup. A sample of n_1 affected children (cases) is randomly collected. Similarly, a random sample of n_2 unaffected children forms the basis of the control group. Denote by $n = n_1 + n_2$ the total number of children at stage 1. There are no degrees of relationship allowed among these children to avoid dependencies within the data at this stage of the experiment. To each child from this basic sample we assign a random variable C_i, $i = 1, \ldots, n$, which counts the number of occurrences of allele A at the candidate locus $L1$, i.e. $C_i = 2$, 1 or 0, accordingly. In what follows, we will consider the case that data from family trios, i.e. from the children and their parents are available, bringing the total number of individuals taking part in the study to $3n$. The random variables corresponding to parental observations are defined equivalently to the offspring data and denoted by A_i and B_i for the first and second parent of the i^{th} trio, respectively. From the above study setup, it follows that there are dependencies among the data and, as a further complication, the number of parents belonging to either group is also random.

In what follows, we suppose that the following standard assumptions on the various processes involved in inheritance are satisfied.

(A1) The random processes yielding the phenotype given the genotype are independent for different individuals.

(A2) Gamete formation is independent of phenotype.

(A3) Hardy-Weinberg equilibrium holds for each locus.

(A4) Offspring data are randomly sampled from the two groups in the population, and the children in the basic sample are unrelated.

Our interest is on testing whether there is an influence of the observed genotype at locus $L1$ on the phenotype. We therefore test the null hypothesis H_0 that there is no linkage disequilibrium (LD) between the observed locus $L1$ and the disease locus $L2$ against the alternative H_1 of the existence of LD between these two loci. Note that the LD coefficient δ describing the LD between alleles at the two loci $L1$ and $L2$ within the population is given by $\delta = \delta_{AA} = h_{AA} - p_{1A}p_{2A}$, where h_{AA} denotes the haplotype frequency of alleles A, A at the loci $L1$ and $L2$, and p_{1A}, p_{2A} stand for the allele frequencies of A at $L1$ and $L2$, respectively. In terms of the LD coefficient δ, the testing problem is thus given by testing the null hypothesis $H_0: \ \delta = 0$ against $H_1: \ \delta > 0$ or $H_1: \ \delta \neq 0$, depending on the experimenter's preference for either a one- or a two-sided alternative. In order to construct tests for these hypotheses, we reformulate the hypotheses in terms of the allele frequencies of A at $L1$ in the two respective groups. Let p_v denote the frequency of A at $L1$ among the affected individuals in the population, and p_w the corresponding term among the unaffected individuals. Then p_v and p_w can be expressed in terms of the model parameters δ, f, p_{1A} and p_{2A} by

$$p_v \ = \ p_{1A} + \delta \, \frac{1 - p_{2A}}{p_{2A}(2 - p_{2A})}, \quad p_w = p_{1A} - \delta \, \frac{f(1 - p_{2A})}{1 - f p_{2A}(2 - p_{2A})}, \tag{1}$$

where a derivation of formula (1) can be found in the appendix. The hypotheses can thus be reformulated equivalently by $H_0 : p_v = p_w = p_{1A}$ against $H_1 : p_v > p_w$ or $H_1 : p_v \neq p_w$.

3 The test statistics and their asymptotic distributions

The null hypothesis H_0 implies that the allele frequency of A at locus $L1$ is the same among affected and unaffected individuals in the population. It is therefore reasonable to consider test statistics based on differences or ratios of estimators for the allele frequencies p_v and p_w to detect deviations from H_0. We choose as estimators $\hat{p}_v$ and $\hat{p}_w$ the empirical counterparts of p_v and p_w in the two respective groups, which can be obtained from the corresponding contingency table with entries $N_1,\ N_2,\ N_3$ and N_4, where N_1 denotes the number of alleles A at $L1$ among the affected individuals, N_2 is the number of A at $L1$ among the unaffected individuals, and $N_3,\ N_4$ are the corresponding numbers of alleles $\bar{A}$ at $L1$. Substituting $\hat{p}_v$ and $\hat{p}_w$ in the formulae for the risk measures yields the following test statistics

$$\begin{aligned}
\text{attributable risk} \qquad T_{n1} &= \hat{p}_v - \hat{p}_w = \frac{N_1}{N_1 + N_3} - \frac{N_2}{N_2 + N_4}, \\
\text{odds ratio} \qquad T_{n2} &= \frac{\hat{p}_v(1 - \hat{p}_w)}{\hat{p}_w(1 - \hat{p}_v)} = \frac{N_1 N_4}{N_2 N_3}, \\
\text{relative risk} \qquad T_{n3} &= \frac{\hat{p}_v}{\hat{p}_w} = \frac{N_1(N_2 + N_4)}{N_2(N_1 + N_3)}.
\end{aligned}$$

We are now ready to present the main result of this article, which gives the joint asymptotic distribution of the empirical allele frequencies $\hat{p}_v$ and $\hat{p}_w$ under the null hypothesis H_0. Since under H_0 we have $\mathrm{Var}(C_1) = \mathrm{Var}(A_1) = \mathrm{Var}(B_1)$ and $\mathrm{Cov}(C_1, A_1) = \mathrm{Cov}(C_1, B_1)$ the results are given in terms of $\mathrm{Var}(C_1)$ and $\mathrm{Cov}(C_1, A_1)$ instead of these five different expressions for brevity of notations. The proof of Theorem 1 is deferred to the appendix.

Theorem 1 *Suppose the null hypothesis H_0 is valid, assumptions (A1)-(A4) are satisfied and the ratio n_1/n converges to a positive constant $c \in (0,1)$ for $n \to \infty$. Then the joint asymptotic distribution of $\hat{p}_v$ and $\hat{p}_w$ is given by*

$$\sqrt{n}\left(\hat{p}_v - p_{1A},\ \hat{p}_w - p_{1A}\right)^T \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma), \tag{2}$$

where the entries of the covariance matrix Σ are given by

$$\Sigma_{1,1} = \frac{Var(C_1)}{12\,t_1} + \frac{c\,p_1\,Cov(C_1,A_1)}{9\,t_1^2}, \quad \Sigma_{2,2} = \frac{Var(C_1)}{12\,(1-t_1)} + \frac{(1-c)(1-p_2)\,Cov(C_1,A_1)}{9\,(1-t_1)^2},$$
$$\Sigma_{1,2} = \Sigma_{2,1} = \frac{Cov(C_1,A_1)(c(1-p_1)+(1-c)p_2)}{18\,t_1(1-t_1)}$$

where p_1 and p_2 denote the probabilities that a parent is affected given the corresponding child is affected or unaffected, respectively. $Var(C_1) = 2p_{1A}(1-p_{1A})$, $Cov(C_1,A_1) = p_{1A}(1-p_{1A})$, and t_1 is the asymptotic expected percentage of affected individuals in the study, i.e. $t_1 = (c + 2c\,p_1 + 2(1-c)p_2)/3$.

Please note that the condition $n_1/n \rightarrow c \in (0,1)$ ensures that the number of data in one of the groups is not outbalanced by the corresponding quantity in the other group. The condition is required in this particular form due to the asymptotic nature of the result of Theorem 1. In a real data application where of course both n and n_1 are finite, this means that the experimenter should make sure that the ratio n_1/n is not too close to either zero or one to avoid situations where there are almost no cases or no controls in the study.

3.1 Asymptotic distributions of the test statistics

To find asymptotic critical values for the tests under consideration, we utilize the result of Theorem 1 to derive the asymptotic distributions of the test statistics T_{n1}, $\log(T_{n2})$ and $\log(T_{n3})$. Instead of T_{n2} and T_{n3}, we will use the logarithms of these test statistics in the ensuing derivations since $\log(T_{n2})$ and $\log(T_{n3})$ were found to preserve the nominal significance level α more precisely in simulations than the original versions of the tests.

Corollary 1 *Under the assumptions of Theorem 1 the asymptotic distributions of the test statistics T_{n1}, $\log(T_{n2})$ and $\log(T_{n3})$ under H_0 are given by*

$$\sqrt{n}\,T_{n1} \xrightarrow{\mathcal{D}} \mathcal{N}(0,\sigma_1^2), \quad \sqrt{n}\,\log(T_{n2}) \xrightarrow{\mathcal{D}} \mathcal{N}(0,\sigma_2^2), \quad \sqrt{n}\,\log(T_{n3}) \xrightarrow{\mathcal{D}} \mathcal{N}(0,\sigma_3^2),$$

where the asymptotic variances σ_1^2, σ_2^2 *and* σ_3^2 *are obtained as*

$$\begin{aligned}
\sigma_1^2 &= \frac{Var(C_1)}{12\,t_1(1-t_1)} + \frac{Cov(C_1,A_1)(cp_1-(c+cp_1+(1-c)p_2)t_1+t_1^2)}{9\,t_1^2(1-t_1)^2} \\
&= \frac{p_{1A}(1-p_{1A})}{18\,t_1^2(1-t_1)^2}\Big(2cp_1+(3-c)t_1-4t_1^2\Big),
\end{aligned}$$

$$\begin{aligned}
\sigma_2^2 &= \frac{Var(C_1)}{12\,t_1(1-t_1)p_{1A}^2(1-p_{1A})^2} + \frac{Cov(C_1,A_1)(cp_1-(c+cp_1+(1-c)p_2)t_1+t_1^2)}{9\,t_1^2(1-t_1)^2p_{1A}^2(1-p_{1A})^2} \\
&= \frac{1}{18\,t_1^2(1-t_1)^2p_{1A}(1-p_{1A})}\Big(2cp_1+(3-c)t_1-4t_1^2\Big),
\end{aligned}$$

$$\begin{aligned}
\sigma_3^2 &= \frac{Var(C_1)}{12\,t_1(1-t_1)p_{1A}^2} + \frac{Cov(C_1,A_1)(cp_1-(c+cp_1+(1-c)p_2)t_1+t_1^2)}{9\,t_1^2(1-t_1)^2p_{1A}^2} \\
&= \frac{(1-p_{1A})}{18\,t_1^2(1-t_1)^2p_{1A}}\Big(2cp_1+(3-c)t_1-4t_1^2\Big)
\end{aligned}$$

The assertions of Corollary 1 follow from Theorem 1 and a straightforward application of the Δ-method; see, e.g., Serfling (1980).

3.2 Estimation of unknown parameters

Since the asymptotic variances of the test statistics depend on the unknown parameters p_{1A}, p_1 and p_2, these parameters have to be estimated in practice. By Slutsky's theorem, the results given in Corollary 1 still hold if σ_1^2, σ_2^2 and σ_3^2 are replaced by consistent estimators. Under the null hypothesis H_0, the allele frequency p_{1A} of A at $L1$ in the two respective groups and the probabilities p_1 and p_2 of a parent being affected given the corresponding child is affected or unaffected, respectively, can be estimated $\sqrt{n}$-consistently by the sample means of the corresponding random variables.

For estimating $\mathrm{Var}(C_1)$ and $\mathrm{Cov}(C_1, A_1)$, there are several approaches. Firstly, one could simply replace p_{1A} by its estimator $\hat{p}_{1A}$ in the relations $\mathrm{Var}(C_1) = 2p_{1A}(1-p_{1A})$ and $\mathrm{Cov}(C_1, A_1) = p_{1A}(1-p_{1A})$. Alternatively, one could use the estimators $\hat{v}$ for $\mathrm{Var}(C_1)$ and $\hat{cov}$ for $\mathrm{Cov}(C_1, A_1)$ given by

$$\hat{v} = \frac{1}{3n}\sum_{i=1}^{n}(C_i^2+A_i^2+B_i^2) - \frac{1}{9n(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}^{n}(C_i+A_i+B_i)(C_j+A_j+B_j)$$

$$\hat{cov} = \frac{1}{2n}\sum_{i=1}^{n}(C_iA_i + C_iB_i) - \frac{1}{9n(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}^{n}(C_i + A_i + B_i)(C_j + A_j + B_j).$$

Straightforward calculations show that these estimators are unbiased with variance converging to 0 at a rate of $1/n$.

4 Extensions

We demonstrate our method for the case of a monogenic disease and a single marker locus for the sake of clearness and brevity of this article. The proposed tests, however, can readily be extended to genetically more complex models. If a set of $k \geq 1$ candidate loci has been chosen we can test as follows whether a specific allele combination (genotype), say $G = (g_1, \ldots, g_k)$, at these k loci contributes to the phenotype expression. We define a virtual biallelic locus L with alleles G and $\bar{G}$ where $\bar{G}$ consists of all allele combinations specified by the k markers except the candidate genotype G. The proposed tests can then be applied to L where the candidate allele A from the original test is replaced by G. Some slight changes have to be accounted for in the calculation of the variances of the test statistics. In this way, all different allele combinations can be tested with an appropriately corrected significance level α, which can, e.g., be found by Bonferroni correction or more sophisticated (more powerful) techniques such as Holm's (1979) or Hochberg's (1988) procedures or modifications of these. This approach will provide us with p-values for forward and backward selection procedures with respect to single allele combinations and/or marker loci. If the k markers are closely located on the same chromosome one could also think of conducting test procedures on haplotypes since classical genetics has demonstrated that the phenotypic effect of several mutations at different loci can sensitively depend on whether the mutations occur in cis or in trans position; see, e.g., Schaid *et al.* (2002). Again, we can test if a specific haplotype, say H, has an influence on the phenotype by defining a virtual biallelic locus with alleles H and $\bar{H}$ and applying our method. To identify the unknown phase the method of molecular haplotyping can be used if the DNA sequence containing the markers is not too long; see Michalatos-Beloin *et al.* (1996). Otherwise haplotypes can either be determined by pedigree analysis or haplotype frequencies can be estimated for example by an EM-algorithm; see, e.g., Fallin *et al.* (2001). A more detailed exploration of this issue will be the subject of further research.

In the situation of a highly inhomogeneous population some care is needed when collecting the data to avoid possible biases of the tests created by population stratification, which is a potential worry in case-control studies; see, e.g., Thomas & Witte (2002). The simplest ways to cope with this situation would be to include only individuals of homogeneous ethnic and geographic origin, or to categorize the data into the different subpopulations, which are then analyzed separately. In case of almost equal allele frequencies at the predisposing locus (percentages of affected individuals in each subpopulation are about equal), it is also feasible to "match" the controls from the basic sample to the cases such that the percentages of different subpopulations are equal in both groups. This will then on average also be true for the parents in the two groups, so that the percentages of different subpopulations will be about equal among the affected and the unaffected individuals. A different strategy to robustify the tests against stratification bias would be to modify the test statistics by estimating p_v and p_w for all subpopulations separately and defining $\hat{p}_v$, $\hat{p}_w$ as the sums (or weighted sums) of the respective estimates. The tests can then be carried out on these modified estimators as before, taking into account the changes in the asymptotic variances. All the methods proposed above, however, rely on knowing the strata, i.e. categorizing individuals into the wrong ethnic group will lead to biased results. Pritchard, Stephens & Donnelly (2000) developed a Bayesian method for the estimation of ethnic origins using genomic information from polymorphic markers that are not linked with the candidate genes under study. The use of this method (or another genomic adjustment approach) can be a helpful supplement to our tests. The

issue of population stratification may, however, not be dismissed.
At first glance it seems that the genetic inheritance model plays a crucial role, and the question arises, whether the proposed tests also work if the assumption of a dominant mode of inheritance is violated. An inspection of the proof shows that Theorem 1 remains valid even in this general setting. The underlying mode of inheritance influences the asymptotic null distributions only through the values of the parameters p_1 and p_2. These parameters, however, can be consistently estimated from phenotype data, so that the tests can be applied to data from any type of inheritance model.

5 Simulation Study

To assess the finite sample behavior of our tests, we carried out a simulation study for various parameter settings. We were, firstly, interested in investigating whether the inclusion of parental data yields a substantial gain in power, and, secondly, in comparing our tests with the TDT as an already established method. Thirdly, we examined the sensitivity of the real significance level with respect to the dependency structure of the data. For brevity, we restricted ourselves to simulate the test $H_0 : \delta = 0$ against the one-sided alternative $H_1 : \delta > 0$, which corresponds to the scenario that the experimenter's belief is in positive LD between A at $L1$ and A at $L2$. To simulate random variables with the same distributions as C_i, A_i and B_i, $i = 1, \ldots, n$, four parameters have to be specified in advance. In the study at hand we fixed the values of the two allele frequencies p_{1A}, p_{2A}, the penetrance f and the LD coefficient δ. Due to the dependency of the LD coefficient δ on the haplotype frequency and the allele frequencies, it is difficult to compare the amount of LD between different pairs of loci using the respective LD coefficients. We therefore use an appropriately standardized version, i.e. Lewontin's $D' = |\delta|/\delta_{max}$ where δ_{max} is defined by $\min\{p_{1A}(1-p_{2A}), (1-p_{1A})p_{2A}\}$ if $\delta > 0$, and $\delta_{max} = \min\{p_{1A}p_{2A}, (1 - p_{1A})(1 - p_{2A})\}$ if $\delta < 0$; see, e.g., Devlin & Risch (1995). In what follows, the parameters will be given in terms of D'

instead of δ for comparability.

Using a nominal type I error of $\alpha = 0.05$ we report the empirical rejection rate based on 10,000 runs. Unkown parameters are estimated as explained in Subsection 3.2. We used the sample sizes $n = 60$, $n = 100$ and $n = 200$ for the basic sample, respectively, and the value of n_1 was chosen according to the "expected equal allocation rule", i.e. the expected number of affected individuals in the entire sample is equal to the expected number of unaffected individuals. For some parameter combinations, this choice was not admissible due to the restriction that $n_1 < n$. In these situations, we simulated the tests for several choices of n_1 with $0.75n \leq n_1 \leq 0.95n$, and found that the tests are not very sensitive with respect to the particular choice of n_1 within this range. Table 1 provides results under $H_0 : D' = 0$ to assess the accuracy of the type I error. Since in what follows we will compare the performance of our tests with the performance of the corresponding tests including either data from unrelated subjects only or data from children with one parent each, we label by $(**)$ the tests including data from both parents for each child.

Table 1 here

It can be seen that our tests preserve the α-level quite well, even for the small sample size of $n = 60$. Only in the situation of a very small (< 0.05) allele frequency p_{1A}, the tests do not preserve the α-level when the sample size is small. In this scenario, the test based on the attributable risk appears to be most robust among the three tests under consideration. Our theoretical results provide an explanation for this fact. Indeed, noting that the variances have the form $\sigma_1^2 = p_{1A}(1-p_{1A}) \times f(p_1, p_2)$ (attr. risk) and $\sigma_2^2 = 1/\{p_{1A}(1-p_{1A})\} \times f(p_1, p_2)$ (odds ratio), where $f(p_1, p_2)$ is a term depending on p_1 and p_2 only, we see that σ_2^2 is more sensitive for such p_{1A}. Increasing the sample size, however, leads to reliable results for all three tests.

Tables 2, 3 and 4 show the simulated powers of the tests on trios under the alternative hypothesis H_1. For comparison, we also display the corresponding values when only data of unrelated subjects are included in the test statistics. In this case, the number of affected individuals n_1 was chosen according to the equal allocation rule $n_1 = n/2$. The tests including offspring data only, thus having a total sample size of n individuals, are not given any label. The tests labeled by $(*)$ denote the tests including the data from the basic sample plus the data from one randomly chosen parent for each child, bringing the total sample size in this situation to $2n$. For these tests, the expected equal allocation rule was used to determine n_1. In Table 2, we chose relatively small allele frequencies of A at the two loci $L1$ and $L2$, a small amount of LD measured by $D' = 0.2$ and a high penetrance of $f = 0.6$.

Table 2 here

We observe that the inclusion of parental data increases the powers of the tests considerably. We further added the powers of the two-sided tests ($H_1 : \ \delta \neq 0$) to Table 2 (scenarios: children only (no label) and both parents included $(**)$) to show that also in this testing problem the inclusion of parental data significantly increases the capability to detect deviation from H_0. The same holds true for medium scale values of the allele frequencies p_{1A} and p_{2A} as can be seen from Table 3.

Table 3 here

Table 4, finally, displays the values of the simulated powers when the value for the penetrance f is chosen relatively small.

Table 4 here

To further assess the practical relevance of a newly developed method it is of great importance to compare its performance with an already established procedure. We chose the TDT as the competing method in this simulation study since it has become the benchmark method for surveys on trios. Table 5 shows the simulated powers of the TDT for the parameter combinations used above so that they can readily be compared with the powers of the tests (with trios) proposed in this article given in Tables 2, 3 and 4. For convenience, the corresponding results for the test based on the attributable risk are given again in Table 5. The number n in this table refers to the number of trios, on which we carried out the tests. Furthermore, we simulated a scenario where the penetrance f is small ($f = 0.1$) to assess the power of our tests if the number of affected parents is likely to be small.

Table 5 here

We observe from Table 5 that for all four scenarios our tests perform between "slightly" better to "significantly" better than the TDT. In particular the test based on the attributable risk has a considerably higher capability to detect deviations from the null hypothesis H_0. In the last scenario where p_{1A} is small, we find that the tests based on the relative risk and the odds ratio are less stable than the test based on the attributable risk, which is again due to the form of the asymptotic variances of these tests. In this situation, these two tests are comparable with the TDT.
To provide an example for another mode of inheritance than the dominant, we also simulated a recessive inheritance model, and again compared our tests with the TDT. The results, which are very similar to those for the dominant model, are given in Table 6.

Table 6 here

The three tests proposed in this article are asymptotically equivalent, but considering all the tables in this study, we observe that for some scenarios the test based on the attributable risk has a higher power than the other two tests in a finite sample. For many scenarios, however, there is virtually no difference in performance between the three tests. This simulation study therefore indicates that in practice one would best use the test based on the attributable risk, which is also hinted at by further simulation results that are not presented here for brevity.

6 Discussion

The simulation study reveals that our tests preserve the nominal significance level very well and are quite robust with respect to the genetic parameters. The statistical power to detect linkage disequilibrium, i.e. to detect predisposing genes, is substantially increased by including parental information. Moreover, a comparative power simulation study with respect to the TDT, which has become a benchmark procedure in practical applications, reveals a superiority of our tests (in terms of power) for many scenarios, demonstrating the practical relevance of our approach. These findings are in line with the theoretical results of McGinnis, Shifman & Darvasi (2002) (for unrelated cases and controls), who found that in general fewer case-control samples are required to achieve the same power as the TDT, suggesting greater genotyping efficiency with the case-control design. The TDT, however, was invented as a test robust to population stratification, which is a potential problem in case-control studies. We have provided some robustification strategies for our tests, which are described in Section 4, but there is no complete solution for this problem. It therefore mainly depends on the structure of the population, which test (our tests or the TDT) might be more appropriate for a particular problem. The reason why our method is more powerful than the TDT seems to be as follows. In contrast to the TDT or other methods based on transmission of marker alleles, we use all available data, i.e. parental phenotypes as well as all parental genotypes, whereas the TDT discards both parental phenotypes and the genotype data of those parents that are homozygote at the observed locus.

The results presented in Section 3 cover the important case that data from parent-offspring trios are available. However, the methodology established in the appendix can be extended to cope with more general family data and the situation of missing data. For example, from a practical viewpoint the generalization to more general types of relatives as the additional inclusion of sibs' data may be a concern. As an example to get the ideas, we first consider the case where for each child from the basic sample the required phenotype and genotype data of both parents and one sib are available. In this scenario, the number of participants in the study is still fixed. From the proof of Theorem 1 in the appendix, we conclude that it is sufficient to extend the random vectors $X_i,\ i = 1, \ldots, n_1$, and $Y_j,\ j = 1, \ldots, n_2$, each corresponding to a child from the two respective groups, by three additional entries giving the phenotype and genotype status of the sib as well as a combination of both exactly analogous to the entries corresponding to the parents. An asymptotic result similar to Lemma 1 can then be proven immediately, from which then the asymptotic null distributions of the test statistics can be derived. Other types of relatives can be added to the study analogously, further increasing the sample size and hence the power to detect deviations from the null hypothesis.

Another strategy to increase power would be to sample on the one hand families with many cases and on the other hand unrelated controls, as, under the alternative hypothesis, this will increase the allele frequency difference between the two groups; see, e.g., Risch & Teng (1998), Teng & Risch (1999), and Risch (2000). It is even possible to allow for missing data under a missing at random assumption. The problem of missing data can be addressed adequately by introducing a random variable indicating if a particular relative takes part in the study or not. Again, the random vectors $X_i,\ i = 1, \ldots, n_1$, and $Y_j,\ j = 1, \ldots, n_2$, are extended by these indicators as well as combinations of the indicator variables with random variables giving the genotype and phenotype status of the respective relative. If the study is designed for data from related cases and unrelated controls it is only necessary to modify the vectors $X_i,\ i = 1, \ldots, n_1$, corresponding to affected children by the indicators. The vectors $Y_j,\ j = 1, \ldots, n_2$, will then be one-dimensional consisting of the offspring genotype data $C_i,\ i = n_1 + 1, \ldots, n$. Note that in these scenarios the number of individuals taking part in the study is random. We therefore have to include the entry N_4 when we calculate the joint asymptotic distribution of the entries of the contingency table since the distribution of N_4 is no longer given by the joint distribution of the other entries N_1, N_2 and N_3.

Note that in models extended in such a way additional nuisance parameters describing the dependence structure and the missing data mechanism, respectively, shall appear, and for a practical implementation of the tests the number of parameters to be estimated (complexity of the model), and the increase of data have to be balanced carefully, since estimating too many parameters compared to the increase in available data can lead to poor results. Therefore more detailed investigations of these issues should be a subject of future research.

Acknowledgements

The authors are indebted to two unknown referees, an Associate Editor and the Editor for some really helpful comments, which improved this article and its presentation substantially. This research was supported by the Deutsche Forschungsgemeinschaft (project: Statistical methods for the analysis of association studies in human genetics). H. Holzmann and A. Munk also acknowledge financial support of the Deutsche Forschungsgemeinschaft. The authors are also grateful to S. Böhringer for some interesting discussions and helpful comments on an earlier version of this article.

References

Armitage, P. (1955). Tests for linear trends in proportions and frequencies. *Biometrics* **11**, 375–386.

Böhringer, S. & Steland, A. (2006). Testing genetic causality for binary traits using parent-offspring data. *Technical Report.* Available at http://www.ruhr-uni-bochum.de/mathematik3/steland.htm#DFG-Project_on_statistical_genetics

Browning, S. R., Briley, J. D., Briley, L. P., Chandra, G., Charnecki, J. H., Ehm, M. G., Johansson, K. A., Jones, B. J., Karter, A. J., Yarnall, D. P. & Wagner, M. J. (2005). Case-control single marker and haplotypic association analysis of pedigree data. *Genetic Epidemiology* **28**, 110–122.

Devlin, B. & Risch, N. (1995). A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics* **29(2)**, 311–22.

Epstein, M. P., Veal, C. D., Trembath, R. C., Barker, J. N. W. N., Li, C. & Satten, G. A. (2005). Genetic association analysis using data from triads and unrelated subjects. *American Journal of Human Genetics* **76**, 592–608.

Fallin, D., Cohen, A., Essioux, L., Chumakov, I., Blumenfeld, M., Cohen, D. & Schork, N. J. (2001). Genetic analysis of case/control data using estimated haplotype frequencies: Application to APOE locus variation and Alzheimer's disease. *Genome Research* **11**, 143–151.

Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* **75(4)**, 800–802.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Statist.* **6**, 65–70.

McGinnis, R., Shifman, S. & Darvasi, A. (2002). Power and Efficiency of the TDT and Case-Control Design for Association Scans. *Behavior Genetics* **32(2)**, 135–144.

McPeek, M. S., Wu, X. & Ober, C. (2004). Best linear unbiased allele-frequency estimation in complex pedigrees. *Biometrics* **60**, 359–367.

Michalatos-Beloin, S., Tishkoff, S., Bentley, K., Kidd, K. & Ruano, G. (1996). Molecular haplotyping of genetic markers 10 kb apart by allele-specific long-range PCR. *Nucleic Acids Research* **24**, 4841–4843.

Nagelkerke, N. J., Hoebee, B., Teunis, P. & Kimman, T. G. (2004). Combining the transmission disequilibrium test and case-control methodology using generalized logistic regression. *European Journal of Human Genetics* **12**, 964–970.

Pritchard, J. K., Stephens, M. & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.

Purcell, S., Sham, P. & Daly, M. J. (2005). Parental phenotypes in family-based association analysis. *American Journal of Human Genetics* **76**, 249–259.

Risch, N. (2000). Searching for genetic determinants in the new millenium, *Nature* **405**, 847–856.

Risch, N. & Teng, J. (1998). The relative power of family-based and case-control designs for linkage disequilibrium studies of complex human diseases. I. DNA pooling. *Genome Research* **8**, 1273–1288.

Schaid, D. J., Rowland, C. M., Tines, D. E., Jacobson, R. M. & Poland, G. A. (2002). Score tests for association between traits and haplotypes when linkage phase is ambiguous. *American Journal of Human Genetics* **70**, 425–434.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

Slager, S. L. & Schaid, D. J. (2001). Evaluation of candidate genes in case-control studies: a statistical method to account for related subjects. *American Journal of Human Genetics* **68**, 1457–1462.

Spielman, R. S., McGinnis, R. E. & Ewens, W. J. (1993). Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *American Journal of Human Genetics* **52(3)**, 501–16.

Teng, J. & Risch, N. (1999). The relative power of family-based and case-control designs for linkage disequilibrium studies of complex human diseases. II. Individual genotyping. *Genome Research* **9**, 234–241.

Thomas, D. C. & Witte, J. S. (2002). Point: Population Stratification: A Problem for Case-Control Studies of Candidate-Gene Associations? *Cancer Epidemiology Biomarkers & Prevention* **11**, 505–512.

Whittemore, A. S. & Tu, I. P. (2000). Detection of disease genes by use of family data. I. Likelihood-based theory. *American Journal of Human Genetics* **66**, 1328–1340.

Ansgar Steland, Department of Mathematics, Ruhr-University Bochum, 44780 Bochum, Germany.
E-mail: ansgar.steland@rub.de

Appendix: Proofs

Derivation of formula (1): Without loss of generality, we consider the "first" allele l_1 at locus $L1$. Then $p_v = P(l_1 = A \mid \text{aff})$, and we obtain (with l_2 denoting the "first" allele at $L2$ so that l_1 and l_2 form a haplotype, and pre standing for the prevalence):

$$\begin{aligned} p_v &= P(l_1 = A \mid \text{aff}, l_2 = A)P(l_2 = A \mid \text{aff}) + P(l_1 = A \mid \text{aff}, l_2 = \bar{A})P(l_2 = \bar{A} \mid \text{aff}) \\ &= \Big(\frac{\delta}{p_{2A}} + p_{1A}\Big)\Big(\frac{fp_{2A}}{pre}\Big) + \Big(p_{1A} - \frac{\delta}{1-p_{2A}}\Big)\Big(1 - \frac{fp_{2A}}{pre}\Big) = p_{1A} + \delta\frac{f - pre}{pre(1-p_{2A})} \end{aligned}$$

This formula holds since given l_2, l_1 does no longer depend on the phenotype. Analogously, $p_w = p_{1A} - \delta(f - pre)/\{(1-pre)(1-p_{2A})\}$. With $pre = fp_{2A}(2-p_{2A})$, we obtain (1).

Proof of Theorem 1: The main difficulty in the proof is that N_1 and N_2 contain sums with a random number of random variables A_i and B_i. However, these can be reconstructed as sums with a deterministic number of random variables as follows. To each affected child, we assign a seven-dimensional random vector $X_i = (X_i^{(1)}, X_i^{(2)}, X_i^{(3)}, X_i^{(4)}, X_i^{(5)}, X_i^{(6)}, X_i^{(7)})^T$, $i = 1, \dots, n_1$, where $X_i^{(1)} = C_i$, $X_i^{(2)} = I\{\text{first parent of child } i \text{ is affected}\}$, $X_i^{(3)} = A_i$, $X_i^{(4)} = X_i^{(2)}X_i^{(3)}$, $X_i^{(5)} = I\{\text{second parent of child } i \text{ is affected}\}$, $X_i^{(6)} = B_i$ and $X_i^{(7)} = X_i^{(5)}X_i^{(6)}$. Analogously, we define the vectors Y_j, $j = 1, \ldots, n_2$, for each child from the control group. Since we did not allow for any degree of relationship among the children, the random vectors X_i, $i = 1, \ldots, n_1$, as well as Y_j, $j = 1, \ldots, n_2$, are iid. Lemma 1 gives the joint asymptotic distribution of the sums of the vectors under the null hypothesis H_0. For brevity of notation, we denote the expectation of C_i, A_i and B_i by p_A, i.e. $p_A = 2\,p_{1A}$.

Lemma 1 *Let n_1/n converge to some constant $c \in (0,1)$ for $n \to \infty$. Denote by m_x and m_y the means of X_1 and Y_1, respectively, and by K_x, K_y the corresponding covariance matrices. Then*

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n_1} X_i - \frac{n_1}{n} m_x,\ \frac{1}{n}\sum_{i=1}^{n_2} Y_i - \frac{n_2}{n} m_y\right)^T \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma_K),$$

where the covariance matrix Σ_K is a non-degenerate block matrix with blocks cK_x and $(1-c)K_y$. Explicitely, we have that $m_x = (p_A, p_1, p_A, p_1 p_A, p_1, p_A, p_1 p_A)^T$, $m_y = (p_A, p_2, p_A, p_2 p_A, p_2, p_A, p_2 p_A)^T$, and K_x is

$$\begin{pmatrix}
Var(C_1) & 0 & Cov(C_1,A_1) & p_1\,Cov(C_1,A_1) & 0 & Cov(C_1,B_1) & p_1\,Cov(C_1,B_1)\\
0 & p_1(1-p_1) & 0 & p_A p_1(1-p_1) & p_3 & 0 & p_3 p_A\\
Cov(C_1,A_1) & 0 & Var(C_1) & p_1\,Var(C_1) & 0 & 0 & 0\\
p_1\,Cov(C_1,A_1) & p_A p_1(1-p_1) & p_1\,Var(C_1) & p_1 p_A(1+(0.5-p_1)p_A) & p_3 p_A & 0 & p_3 p_A^2\\
0 & p_3 & 0 & p_3 p_A & p_1(1-p_1) & 0 & p_A p_1(1-p_1)\\
Cov(C_1,B_1) & 0 & 0 & 0 & 0 & Var(C_1) & p_1\,Var(C_1)\\
p_1\,Cov(C_1,B_1) & p_3 p_A & 0 & p_3 p_A^2 & p_A p_1(1-p_1) & p_1\,Var(C_1) & p_1 p_A(1+(0.5-p_1)p_A)
\end{pmatrix}.$$

K_y is of the same form as K_x with p_1 replaced by p_2, and the parameter p_3 describing the covariance structure between parental phenotypes given the offspring is affected, replaced by a parameter, say p_4, for the respective covariance given the child is unaffected.

Proof of Lemma 1: The expectations and covariance matrices of X_1 and Y_1 under H_0 are obtained by straightforward calculations. Note that under the null hypothesis the random variables $X_i^{(2)}$, $X_i^{(5)}$ and $Y_j^{(2)}$, $Y_j^{(5)}$ corresponding to phenotype data are independent of the random variables $X_i^{(1)}$, $X_i^{(3)}$, $X_i^{(6)}$ and $Y_j^{(1)}$, $Y_j^{(3)}$, $Y_j^{(6)}$ describing genotype features of the individuals. Asymptotic normality follows by applying the multivariate central limit theorem for iid vectors to both sequences separately, and exploiting that the sequences X_i, $i = 1, \ldots, n_1$, and Y_j, $j = 1, \ldots, n_2$, are independent.

To prove that the covariance matrix K_x is non-degenerate we calculate its determinant $|K_x| = 0.5p_1^2(1-p_1)^2\mathrm{Var}(C_1)^5(p_1^2(1-p_1)^2 - p_3^2)$, where we used that $\mathrm{Var}(C_1) = 2\,p_{1A}(1-p_{1A})$ and $\mathrm{Cov}(C_1, A_1) = p_{1A}(1-p_{1A})$. Since p_{1A}, p_1 and p_2 are assumed to lie in $(0,1)$ it remains to show that $p_3^2 < p_1^2(1-p_1)^2$. Recall that $p_1^2(1-p_1)^2 - p_3^2 = \mathrm{Var}(X_1^{(2)})\mathrm{Var}(X_1^{(5)}) - [\mathrm{Cov}(X_1^{(2)}, X_1^{(5)})]^2 \geq 0$ by Hölder's inequality. As $X_1^{(2)}$ and $X_1^{(5)}$ are not linearly dependent (the four possible outcome combinations for $X_1^{(2)}$ and $X_1^{(5)}$ each occur with positive probability) even the strict inequality holds and thus the determinant of K_x is positive. Analogously, $|K_y| > 0$ with p_1 and p_3 replaced by p_2 and p_4, respectively.

Lemma 2 *For $\lim_{n\to\infty} n_1/n = c \in (0,1)$, we obtain under H_0:*

$$\sqrt{n}\Big(\frac{N_1}{n} - \frac{E[N_1]}{n},\ \frac{N_2}{n} - \frac{E[N_2]}{n},\ \frac{N_3}{n} - \frac{E[N_3]}{n}\Big)^T \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma_N),$$

with expectations $E[N_1] = (n_1 + 2n_1p_1 + 2n_2p_2)p_A$, $E[N_2] = (3n - n_1 - 2n_1p_1 - 2n_2p_2)p_A$, $E[N_3] = (n_1 + 2n_1p_1 + 2n_2p_2)(2 - p_A)$, and the entries $\Sigma_{N,i,j}$, $i, j = 1, 2, 3,$ of the covariance matrix Σ_N are given by

$$\begin{aligned}
\Sigma_{N,1,1} &= \mathit{Var}(C_1)\{c + 2cp_1 + 2(1-c)p_2\} + 4\mathit{Cov}(C_1, A_1)cp_1 \\
&\quad +2p_A^2\{cp_1(1-p_1) + (1-c)p_2(1-p_2) + cp_3 + (1-c)p_4\} \\
\Sigma_{N,1,2} &= 2\mathit{Cov}(C_1, A_1)\{c(1-p_1) + (1-c)p_2\} \\
&\quad -2p_A^2\{cp_1(1-p_1) + (1-c)p_2(1-p_2) + cp_3 + (1-c)p_4\} \\
\Sigma_{N,1,3} &= -\mathit{Var}(C_1)\{c + 2cp_1 + 2(1-c)p_2\} - 4\mathit{Cov}(C_1, A_1)cp_1 \\
&\quad +2p_A(2-p_A)\{cp_1(1-p_1) + (1-c)p_2(1-p_2) + cp_3 + (1-c)p_4\} \\
\Sigma_{N,2,2} &= \mathit{Var}(C_1)\{3 - c - 2cp_1 - 2(1-c)p_2\} + 4\mathit{Cov}(C_1, A_1)(1-c)(1-p_2) \\
&\quad +2p_A^2\{cp_1(1-p_1) + (1-c)p_2(1-p_2) + cp_3 + (1-c)p_4\} \\
\Sigma_{N,2,3} &= -2\mathit{Cov}(C_1, A_1)\{c(1-p_1) + (1-c)p_2\} \\
&\quad -2p_A(2-p_A)\{cp_1(1-p_1) + (1-c)p_2(1-p_2) + cp_3 + (1-c)p_4\}
\end{aligned}$$

$$\begin{aligned}\Sigma_{N,3,3} &= Var(C_1)\{c + 2cp_1 + 2(1-c)p_2\} + 4Cov(C_1, A_1)cp_1 \\ &\quad +2(2-p_A)^2\{cp_1(1-p_1) + (1-c)p_2(1-p_2) + cp_3 + (1-c)p_4\}.\end{aligned}$$

Proof of Lemma 2: We can express N_1, N_2 and N_3 as functions of the sums of the entries of the vectors $X_i,\ i = 1, \ldots, n_1,\ Y_j,\ j = 1, \ldots, n_2$, i.e.

$$N_1 = \sum_{i=1}^{n_1} X_i^{(1)} + \sum_{i=1}^{n_1} X_i^{(4)} + \sum_{j=1}^{n_2} Y_j^{(4)} + \sum_{i=1}^{n_1} X_i^{(7)} + \sum_{j=1}^{n_2} Y_j^{(7)}$$

and analogously for N_2 and N_3. Interpreting the vector $(N_1, N_2, N_3)^T$ as a (measurable and differentiable) function from $I\!R^{14}$ to $I\!R^3$, we obtain the statement of Lemma 2 by applying the Δ-method and exploiting that $\mathrm{Cov}(C_1, A_1) = \mathrm{Cov}(C_1, B_1)$.

Applying the Δ-method to the function $(\hat{p}_v, \hat{p}_w)^T$, where $\hat{p}_v = N_1/(N_1 + N_3)$ and $\hat{p}_w = N_2/(6n - N_1 - N_3)$ then yields (2). The formulae for $\mathrm{Var}(C_1)$, $\mathrm{Cov}(C_1, A_1)$ and t_1 are obtained by straightforward calculations.

Table 1: Left columns: $p_{1A} = 0.2$, $p_{2A} = 0.015$, $D' = 0$, $f = 0.5$; Right columns: $p_{1A} = 0.2$, $p_{2A} = 0.1$, $D' = 0$, $f = 0.3$

test	$n = 60$	$n = 100$	$n = 200$	$n = 60$	$n = 100$	$n = 200$
attributable risk ($**$)	5.59%	5.56%	5.53%	5.52%	5.18%	5.04%
log odds ratio ($**$)	4.73%	4.74%	4.88%	5.12%	4.86%	4.78%
log relative risk ($**$)	4.91%	5.01%	5.04%	5.13%	4.95%	4.86%

Table 2: $p_{1A} = 0.05$, $p_{2A} = 0.001$, $D^{'} = 0.2$, $f = 0.6$

test	$n = 60$	$n = 100$	$n = 200$
attributable risk	58.44%	76.65%	95.95%
attributable risk (∗)	69.46%	86.61%	98.71%
attributable risk (∗∗)	80.53%	94.19%	99.79%
log odds ratio	58.68%	77.72%	95.85%
log odds ratio (∗)	69.76%	86.21%	98.65%
log odds ratio (∗∗)	76.75%	92.29%	99.73%
log relative risk	60.82%	78.11%	95.97%
log relative risk (∗)	70.70%	86.72%	98.70%
log relative risk (∗∗)	79.05%	93.38%	99.76%
two-sided tests			
attributable risk	44.20%	64.58%	79.51%
attributable risk (∗∗)	71.07%	89.41%	99.52%
log odds ratio	48.82%	63.81%	85.54%
log odds ratio (∗∗)	65.95%	86.42%	99.19%
log relative risk	51.37%	66.00%	87.07%
log relative risk (∗∗)	67.53%	87.16%	99.35%

Table 3: $p_{1A} = 0.2$, $p_{2A} = 0.1$, $D' = 0.2$, $f = 0.5$

test	$n = 60$	$n = 100$	$n = 200$
attributable risk	30.08%	40.67%	64.36%
attributable risk (∗)	36.42%	51.80%	76.94%
attributable risk (∗∗)	46.45%	63.84%	86.52%
log odds ratio	30.30%	41.10%	63.29%
log odds ratio (∗)	36.50%	51.29%	76.47%
log odds ratio (∗∗)	42.43%	60.02%	84.51%
log relative risk	30.82%	41.41%	64.21%
log relative risk (∗)	36.99%	51.82%	76.68%
log relative risk (∗∗)	44.56%	62.33%	85.54%

Table 4: $p_{1A} = 0.4$, $p_{2A} = 0.2$, $D' = 0.4$, $f = 0.2$

test	$n = 60$	$n = 100$	$n = 200$
attributable risk	36.51%	50.62%	75.41%
attributable risk (∗)	40.07%	57.96%	81.76%
attributable risk (∗∗)	55.72%	74.72%	94.18%
log odds ratio	35.28%	49.06%	73.88%
log odds ratio (∗)	38.90%	56.10%	80.45%
log odds ratio (∗∗)	51.05%	71.16%	93.01%
log relative risk	38.99%	49.97%	74.04%
log relative risk (∗)	40.47%	57.51%	81.54%
log relative risk (∗∗)	51.47%	71.87%	93.25%

Table 5: Simulated power of the tests based on the risk measures compared with the TDT

test	$n = 60$	$n = 100$	$n = 200$
$p_{1A} = 0.05,\ p_{2A} = 0.001,\ D' = 0.2,\ f = 0.6$			
attributable risk ($**$)	80.53%	94.19%	99.79%
TDT	76.08%	92.41%	99.78%
$p_{1A} = 0.2,\ p_{2A} = 0.1,\ D' = 0.2,\ f = 0.5$			
attributable risk ($**$)	46.45%	63.84%	86.52%
TDT	31.17%	45.04%	74.75%
$p_{1A} = 0.4,\ p_{2A} = 0.2,\ D' = 0.4,\ f = 0.2$			
attributable risk ($**$)	55.72%	74.72%	94.18%
TDT	41.07%	60.62%	88.31%
$p_{1A} = 0.05,\ p_{2A} = 0.1,\ D' = 0.1,\ f = 0.1$			
attributable risk ($**$)	18.55%	22.79%	34.39%
log odds ratio ($**$)	11.24%	16.04%	27.39%
log relative risk ($**$)	12.22%	17.27%	28.53%
TDT	12.68%	17.69%	26.97%

Table 6: Simulated power of the tests based on the risk measures compared with the TDT for a recessive inheritance model and parameters $p_{1A} = 0.2$, $p_{2A} = 0.2$, $D' = 0.1$, $f = 0.5$.

test	$n = 60$	$n = 100$	$n = 200$
attributable risk $(**)$	41.14%	56.67%	81.87%
log odds ratio $(**)$	36.03%	52.53%	79.47%
log relative risk $(**)$	36.79%	53.04%	80.05%
TDT	33.84%	50.13%	77.58%