We consider three tests for genetic association in data from nuclear families (the Family-Based Association Test (FBAT) test proposed by Rabinowitz and Laird ([2000] Hum. Hered. 50:211–223), a second test proposed by Rabinowitz ([2002] J. Am. Stat. Assoc. 97:742–758), and the Family Genotype Analysis Program (FGAP) nonfounder or partial score test proposed by Clayton ([1999] Am. J. Hum. Genet. 65:1170–1177) and Whittemore and Tu ([2000] Am. J. Hum. Genet. 66:1329–1340)). We show that each test statistic arises from the efficient score of the family data as the solution to a set of constraints on its null expectation. Moreover, the FBAT and Rabinowitz tests (but not the FGAP test) are locally the most powerful among all tests satisfying their constraints. We used simulations to examine how the three tests perform in situations when their assumptions are violated and the number of families is not huge. We found that the FBAT test tended to have less power than the other two tests, particularly when applied to families in whom all offspring were affected. The Rabinowitz and FGAP tests performed similarly, although the latter tended to extract more information from families containing one typed parent. While none of the tests showed good power to detect rare, recessively acting genes, the Rabinowitz test with a sample variance estimate performed particularly poorly in this case. However, the Rabinowitz test with a model-based variance had power comparable to that of the FGAP test, and more accurate type I error rates. We conclude that for the situations we considered, the Rabinowitz test with model-based variance has good power without forfeiting robustness against misspecification of parental genotype probabilities. However, its utility is limited by the lack of a simple algorithm to apply it to families with varying structures and phenotypes. Genet Epidemiol 25:80–91, 2003. © 2003 Wiley-Liss, Inc.

Key words: constrained optimization; family-based association tests; Hardy-Weinberg frequencies; statistical power

INTRODUCTION

Genetic association studies can be more powerful than allele-sharing methods for identifying genes with small to moderate effects on disease risk [Risch and Merikangas, 1996]. Case-control studies are commonly used to evaluate associations between traits and candidate genes. However, concern about the possibility of association due to population stratification rather than linkage has fostered the development of family-based association tests. In particular, the transmission disequilibrium test (TDT) and its generalizations [Falk and Rubinstein, 1987; Terwilliger and Ott, 1992; Spielman et al., 1993; Ewens and Spielman, 1995; Ott, 1989; Spielman and Ewens, 1996] provide valid tests for linkage, regardless of the distribution of parental genotypes, the family structures, or the family traits. These tests are based on the conditional distribution of genotypes in a set of unrelated affected individuals, given the genotypes of their parents. The tests compare the observed number of parental transmissions of a specific allele to that expected when the gene is unassociated with the trait of interest. However, these tests require known parental genotypes, which is a problem for conditions having a late onset, when parents of affected individuals are apt to be deceased. When not all parental genotypes are known, use of the TDT only on families with complete parental genotype data can yield the wrong type I error probabilities [Curtis and Sham, 1995]. Several researchers proposed test statistics that avoid such bias by including all families, regardless of missing parental genotypes. Here we consider three of these statistics: the
family-based association test (FBAT) of Rabinowitz and Laird [2000], the statistic recently proposed by Rabinowitz [2002], and the nonfounder or partial score statistic proposed by Clayton [1999] and Whittemore and Tu [2000]. We provide a common theoretical basis for the three statistics, and use simulations to examine how they perform.

To fix ideas, suppose we wish to use the genotype data from \( n \) unrelated nuclear families to infer a measure \( \theta \) of gene-trait association, with \( \theta = 0 \) corresponding to the null hypothesis of no association. (We restrict our attention to nuclear families to simplify the presentation; however, the theory underlying all three statistics extends to multigenerational families.) We start by specifying a model for the distribution of observed family genotypes \( G_i \) and for the probability of family trait \( T_i \) given \( G_i \), \( i = 1, \ldots, n \). To accommodate the ascertainment of families on the basis of their traits, we base the statistics on the resulting conditional distribution \( P_{G_i|T_i}^0(g|t_i) \) of \( G_i \) given the trait value \( T_i = t_i \). Each of the three statistics has the form

\[
\frac{\sum_{i=1}^{n} e_i(G_i, t_i)}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}}
\]

where \( e_i(G_i, t_i) \) is derived from \( P_{G_i|T_i}^0(g|t_i) \), and where \( \sigma_i^2 \) is the null variance of \( e_i(G_i, T_i) \) (or an estimate of this variance).

We shall show that each of the statistics can be constructed from a solution to a constrained optimization problem. In the absence of constraints, the optimal solution for a specified model is the full efficient score for \( \theta \) from the likelihood of all the observed family genotypes, given all observed family traits (including data from parents). However, this solution requires specifying a null distribution for the parental genotypes, and misspecification of this null distribution can produce incorrect type I error probabilities. To reduce or eliminate sensitivity to such misspecification, the solutions are constrained to have zero expectation with respect to a given class of null probability distributions. Different classes of null distributions yield different statistics, with the class for the FBAT statistic giving the most stringent constraint, and the class for the FGAP nonfounder statistic giving the least stringent constraint. When both parental genotypes are known, all three classes of distributions reduce to Mendelian probabilities for the offspring genotypes conditional on those of the parents. Therefore, all three statistics reduce to the same statistic (the TDT statistic or another of the score statistics proposed by Schaid [1996]) when all parental genotypes are known.

Despite their common properties, relatively little is known about how the statistics perform in practical applications when their underlying assumptions are violated, the number of families is small or moderate, and the family structures and traits vary. Cervino and Hill [2000] used simulations to evaluate the bias and power of the partial score statistic of Clayton [1999] (i.e., the FGAP nonfounder statistic), and showed that the test performed well in the situations examined. To further address this issue, we used simulations to study the size and power of all three tests in various settings. We conclude with a summary of the results and a brief discussion of the relative ease of using the statistics to evaluate association, based on a sample of heterogeneous nuclear families with varying numbers of offspring, traits, and patterns of missing parental genotype data.

**THE OPTIMIZATION PROBLEM**

To describe how each of the three statistics can be constructed from solutions to a constrained optimization problem, we distinguish between the observed family genotype data \( G_i \) and the complete family data \( C_i = (G_i, M_i) \) that would have been observed had both parents been typed. Here \( M_i \) denotes the missing parental genotype data; if none is missing, \( C_i = G_i \). The constraints associated with each statistic are implemented via a set of family-specific matrices \( X_1, \ldots, X_n \). Here \( X_i \) is an \( r_i \times s_i \) matrix whose rows index the \( r_i \) possible values of the observed genotype data \( G_i \) for family \( i \), given its number of offspring and missing parental genotype pattern. For a family with one offspring and one untyped parent, for instance, the first two columns of Table I show the \( r_i = 7 \) possible values \( g \) for the observed family genotype \( G_i \), when genotypes are evaluated with respect to a single diallelic polymorphism. The columns of \( X_i \) are indexed by \( \ell \) which takes the \( s_i \) possible values of a function \( L_i = f(C_i) \) of the complete data \( C_i \). The \((g, \ell)_{th}\) entry of \( X_i \) is

\[
x_{ig\ell} = P_{G_i|L_i}^0(g|\ell), \quad g = 1, \ldots, r_i, \quad \ell = 1, \ldots, s_i,
\]

where \( P_{G_i|L_i}^0(g|\ell) \) denotes the null probability that \( G_i = g \) given the value \( \ell \) of \( L_i \). For the FBAT
To motivate the objective function (3), let $C$ be the minimal sufficient statistic for the null hypothesis (see Rabinowitz and Laird [2000] for further discussion). For the Rabinowitz statistic, $L_i$ is the complete parental genotype data. For the FGAP statistic, $L_i$ is the observed value of the parental genotypes; if neither parental genotype is observed, $x_{ig}$ is the observed value of $x_{ig}$. Thus in the presence of missing parental genotypes, the $L_i$ for the FBAT and FGAP statistics depend only on the observed data $G_{ij}$ while the $L_i$ for the Rabinowitz statistic depends also on the missing parental data $M_i$.

Let $E_i$ denote an $r_i$-dimensional column vector whose $g^{th}$ component is $e_i(g, t_i)$. We say that a solution $E_1, \ldots, E_n$ to the equations

$$X_i^TE_i = 0, \quad i = 1, \ldots, n$$

(2)
is optimal if it maximizes

$$\lim_{\theta \to 0} \left[ \theta^{-1} \sum_{i=1}^{n} \sum_{g=1}^{r_i} e_i(g, t_i)P_{G_i|T_i}(g|t_i) \right]$$

(3)
among all solutions of (2) having the property that

$$\sum_{i=1}^{n} \sum_{g=1}^{r_i} e_i^2(g, t_i)P_{G_i|T_i}(g|t_i) = C$$

(4)

where $C$ is the same constant for all solutions. To motivate the objective function (3), let $\phi(\theta) = \sum_g \sum_{t_i} e_i(g, t_i)P_{G_i|T_i}(g|t_i)$ denote the expected value of the numerator of (1). We show in the Appendix that the constraints (2) insure that $\phi(0) = 0$. Thus (3) gives the slope of the power function $\phi(\theta)$ at $\theta = 0$, and optimality corresponds to local efficiency [Cox and Hinkley, 1974].

**RELATIONSHIPS AMONG STATISTICS**

The following theorem describes the locally most efficient statistic satisfying a given set of constraints on its null expectation. The theorem also describes how the constraints for the three statistics are related.

**THEOREM**

a) The optimal solution $E_1, \ldots, E_n$ to the maximization problem (3), subject to (2) and (4), is given by

$$E_i = V_i - W_i^{-1}X_i^TW_i^{-1}X_i^TV_i, \quad i = 1, \ldots, n.$$  

(5)

Here $V_i$ is the $r_i$-dimensional vector of efficient scores whose $g^{th}$ component is $v_{ig} = \frac{d}{d\theta} \log p_{G_i|T_i}(g|t_i)_{\theta=0}$, and $W_i$ is the $r_i \times r_i$ diagonal matrix with entries $p_{G_i|T_i}(g|t_i)$. The test statistic (1), with $e_i(G_i, t_i)$ given by the appropriate component of (5) and with

$$\sigma^2_i = V_i^T[W_i - X_i^TW_i^{-1}X_i^TV_i]^TV_i$$

is locally most efficient subject to the constraint (2).

b) Each of the FBAT, Rabinowitz, and FGAP nonfounder statistics is constructed from a solution to (2) for a specific set of matrices $X_i$. The solutions for the FBAT and Rabinowitz statistics are optimal, but the solution for the FGAP nonfounder statistic is not.

c) When both parental genotypes are known, the set of matrices $X_i$ is the same for all three statistics. The columns of $X_i$ correspond to the $s_i$ possible parental genotypes $L_i$, as illustrated in Table II for a diallelic polymorphism. In this case, (5) gives the efficient score vector with components

$$e_i(g, t_i) = \frac{d}{d\theta} \log p_{G_i|T_i}(g|L_i = \ell, t_i)_{\theta=0}.$$  

(6)
d) The column space of an FBAT matrix $X_i$ contains that of the corresponding Rabinowitz
matrix $X_{ij}$, which in turn contains the column space of the FGAP matrix $X_i$.

The Appendix contains a proof of this theorem. To illustrate the constraint matrices $X_i$ corresponding to the three statistics, we show in Table I the Rabinowitz and FGAP matrices for a diallelic polymorphism, based on a family with one offspring and one untyped parent. The $s_i = 9$ columns of the Rabinowitz matrix correspond to the nine possible pairs of parental genotypes. The $(g, \ell)^{th}$ entry gives the null probability of observing family genotype $g$ conditional on parental genotype $\ell$, $g = 1, \ldots, 7$, $\ell = 1, \ldots, 9$. The $s_i = 3$ columns in the FGAP matrix correspond to the three possible genotypes of the typed parent, and the $(g, \ell)^{th}$ entry gives the null probability of family genotype $g$ conditional on the observed parental genotype. Note that the FGAP matrix requires specification of a null distribution for the parental genotypes. The $s_i = 7$ columns of the FBAT matrix $X_i$ (not shown) correspond to the seven possible values of the minimal sufficient statistic $L_i$ for the null hypothesis. For this family structure and missingness pattern and for diallelic polymorphisms, the minimal sufficient statistic is the entire observed family genotype $G_i$. Thus for FBAT, $X_i$ is the identity matrix of dimension 7, and so the constraint (2) implies $E_i = 0$. Hence for diallelic polymorphisms, families with this structure and missingness pattern do not contribute to the FBAT statistic.

**IMPLICATIONS OF THEOREM**

Assertion (a) states that the test statistic that is locally most powerful for a given set of constraints (2) is based on the projections $E_1, \ldots, E_n$ of the efficient score vectors $V_i$ onto the spaces $S_i$ orthogonal to the column spaces of the constraint matrices $X_i$. In the absence of constraints, the optimal statistic is based on those vectors $E_i$ in $S_i$ that are "closest" to $V_i$ in the metric induced by the norm $\|E\|^2 = E^T W_i E$. A constrained statistic based on such $E_1, \ldots, E_n$ is less powerful than the one based on $V_1, \ldots, V_n$, and the more stringent the constraints (i.e., the larger the column spaces of $X_1, \ldots, X_n$), the less powerful the statistic. Assertion (d) states that the constraints for the FBAT statistic are most stringent, and the ones for the FGAP nonfounder statistic least stringent. However, any increased efficiency gained by the FGAP statistic is purchased at the price of a smaller set of null parental genotype distributions for which the statistic yields correct $P$-values. Thus, of the three statistics, the FGAP nonfounder test maintains its nominal size for the smallest set of null parental genotype distributions. (This set can be expanded somewhat by specifying its distributions in terms of unknown parameters that are estimated by maximum likelihood.) In contrast, the constraints for the Rabinowitz statistic are the least stringent needed to insure an unbiased test in the family of null distributions restricted only by Mendelian parent-to-offspring allele transmission.

Note from (5) that although the statistics’ optimality depends on correct specification of the null probabilities $P_{G_i}(g)$ in the matrices $W_i$, the constraints (2) hold regardless of the values of the $W_i$. This means that the tests using these statistics retain a specified size, regardless of how the

**TABLE II. Matrix $X_i$ for a Family With One Offspring and Both Parental Genotypes Observed, Based on Genotypes at a Diallelic Polymorphism With Alleles A,B**

<table>
<thead>
<tr>
<th>Observed genotype</th>
<th>Parental mating type&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parents&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Offspring</td>
</tr>
<tr>
<td>1</td>
<td>AA,AA</td>
</tr>
<tr>
<td>2</td>
<td>AA,AB</td>
</tr>
<tr>
<td>3</td>
<td>AA,AB</td>
</tr>
<tr>
<td>4</td>
<td>AA,BB</td>
</tr>
<tr>
<td>5</td>
<td>AB,AB</td>
</tr>
<tr>
<td>6</td>
<td>AB,AB</td>
</tr>
<tr>
<td>7</td>
<td>AB,AB</td>
</tr>
<tr>
<td>8</td>
<td>AB,AB</td>
</tr>
<tr>
<td>9</td>
<td>AB,AB</td>
</tr>
<tr>
<td>10</td>
<td>AB,AB</td>
</tr>
</tbody>
</table>

<sup>a</sup>With pooling rows and columns corresponding to parental genotypes AA, AB and AB, AA, etc.
probabilities \( P^0_{G_i}(g) \) in \( W_i \) are specified. Moreover, the \( P^0_{G_i\mid t_i}(g|\ell) \) for the FBAT and Rabinowitz statistics involve only Mendelian parent-to-offspring transmission probabilities. However, the null variances depend on the \( P^0_{G_i}(g) \) that appear in \( W_i \) and therefore their misspecification could induce incorrect \( P \)-values by inducing misspecified variances. Rabinowitz [2002] circumvents this problem by replacing the variance of \( \sum_{i=1}^n e_i(G_i, t_i) \) by its sample variance \( \sum_{i=1}^m \hat{\sigma}_i^2(G_i, t_i) \). However, this strategy may be problematic when sample sizes are small. In the simulations described below, we evaluate the performance of the Rabinowitz test with sample variance estimate and also with the model-based variance \( \sigma^2 = \sum_{i=1}^m \hat{\sigma}_i^2 \), where the \( \hat{\sigma}_i^2 \) are given by expression (6).

Assertion (c) affirms the established fact that when all parental genotypes are known, the locally most powerful statistic for testing the null hypothesis \( \theta = 0 \), among all statistics having zero null expectation given the parental genotypes, is the efficient score statistic based on the conditional distribution of \( G_i \) given the parental genotypes [Cox and Hinkley, 1974]. In the absence of complete parental genotype data, the FBAT statistic is also an efficient score statistic, based on the conditional distribution of \( G_i \) given the offspring traits and the minimal sufficient statistic for the null hypothesis (see Appendix). In contrast, neither the Rabinowitz statistic nor the FGAP nonfounder statistic is an efficient score statistic obtained by conditioning on some function of the observed data.

Finally, it should be noted that the property of maximum local efficiency is convincing only in so far as it leads to near-optimal power in situations when \( \theta \) is large in magnitude. Below, we use simulations to evaluate the size and power of the three statistics in a range of conditions of practical interest in applications.

**SIMULATIONS**

We evaluated the three statistics’ type I error rates and their power to detect association between a trait and a variant allele of a diallelic polymorphism, based on simulated genotype data for \( n=100 \) nuclear families, each containing three offspring. We generated values for the complete family genotypes, conditional on one of two common family traits \( T_j \equiv t \). Here \( t \) is a five-dimensional vector of trait indicators for offspring and parents. In both family traits, one parent was affected and one was unaffected. However, in one family trait, all three offspring were affected, while in the other, two offspring were affected and one was unaffected. (We chose these heavily loaded trait families to obtain sharp power comparisons for the three statistics, and to compare the power of trait-concordant and trait-discordant sibships when parental genotypes are missing.) For each simulation, we: a) used one of 24 models (described below) to generate the family genotypes; b) deleted the genotypes of one or both parents; c) calculated each test statistic and flagged it if it exceeded the critical value 1.645 for a one-tailed test of nominal size 0.05; and d) repeated steps (a–d) 1,000 times, and tabulated size or power for a test as the proportion of the 1,000 replications in which it exceeded this critical value.

**CALCULATING THE STATISTICS**

We describe the forms of the three statistics as applied to families with two affected and one unaffected offspring, and with one untyped parent. Specifically, we describe the numerator in (1) for each of the three statistics, focusing on a single contribution \( e(G, t) \), where we have dropped the subscript \( i \). Let \( G = (G_P, G_1, G_2, G_3) \), where \( G_P \) denotes the genotype of the typed parent, \( G_t \) denotes the genotype of offspring, and \( G_\ell \) denotes the genotype of offspring \( \ell, \ell = 1, 2, 3 \), and an individual’s genotype is coded as 0, 1, or 2, depending on the number of variant alleles he or she carries. Each of the statistics requires the user to specify a weight \( w_i(t) \) for offspring \( \ell \) that depends on his trait status, as well as a count \( c(G_\ell) \) for his genotype that depends on its effects on his trait risk. Under a dominant model, for example, \( c(G)=1 \) if \( G=AA \) or \( AB \) and \( c(G)=0 \) otherwise, while for a recessive model, \( c(G)=1 \) only if \( G=AA \), and for an additive model, \( c(G) \) is the number of \( A \) alleles in \( G \). The offspring weights are typically taken as \( 1-v \) for affected offspring and \( -v \) for unaffected offspring, where \( v \) is an arbitrarily chosen “offset” [Lange and Laird, 2002]. In calculating all statistics, we chose \( v=0.06 \) (the null trait prevalence used to generate the data). We also chose the genotypic count to match the genetic model (additive or recessive) used to generate the data. These choices of weights and genotypic counts are locally optimal [Whittemore and Tu, 2000; Lange and Laird, 2002; Rabinowitz, 2002].
For the FBAT statistic
\[
e(G, t) = \sum_{t=1}^{3} w_t(t) \{ c(G_t) - E_0[c(G_t); G_P] \}
\]
where \( E_0 \) is the null expectation of the genotype count for an offspring, conditional on the minimal sufficient statistic for the null hypothesis. Rabinowitz and Laird [2000] gave an algorithm for computing \( E_0 \). For the FGAP nonfounder statistic, \( e(G, t) \) is given by (7), but \( E_0 \) is the null expectation of an offspring’s genotype count, conditional on all the available parental genotype information obtained from the observed genotypes in the entire family. This information is specified as a posterior distribution of parental genotypes, given the observed family genotypes, and requires specification of a prior distribution for the genotype of the untyped parent. For the Rabinowitz statistic, \( e(G, t) \) is given by (7) with the same expectation \( E_0 \) as that used for the FGAP statistic, but with an additional correction term subtracted:
\[
e(G, t) = \sum_{t=1}^{3} w_t(t) \{ c(G_t) - E_0[c(G_t); G_P] \} - e_{corr}(G).
\]
The correction term \( e_{corr}(G) \) insures that \( e(G, t) \) has zero null expectation, whatever the genotype of the untyped parent. When both parental genotypes are known, \( e_{corr}(G) = 0 \).

We computed the FGAP nonfounder statistic using two sets of null models: i) those specifying random mating and Hardy-Weinberg (HW) frequencies (called FGAP-HW), and ii) those specifying only random mating (called FGAP-NHW). The Rabinowitz statistic was computed using the model with only random mating, and using both sample- and model-based variance estimates.

**GENERATING FAMILY GENOTYPE DATA**

A complete family genotype is a five-dimensional vector of genotypes for the individual family members. We indexed the \( 3^5 = 243 \) possible complete family genotypes by \( g = 1, \ldots, 243 \). We generated family genotype indices \( g_1, \ldots, g_{100} \), given their common family trait value \( t_1 = \cdots = t_{100} = t \), using the distribution
\[
P_{G|T}(g|t) = \frac{P_{G|T}(g|t|g)}{\sum_g P_{G|T}(g'|t|g')}.
\]
In all models, we assumed random parental mating with respect to genotypes. We modeled the null probability \( \pi \) that a parent has genotype \( j \), \( j = 0, 1, 2 \), as
\[
\pi_2 = F_p + (1 - F)p^2, \quad \pi_1 = (1 - F)2pq, \quad \pi_0 = Fq + (1 - F)q^2, \quad 0 \leq F < 1.
\]
Here \( p = 1 - q \) is the null frequency of the variant allele among parents, and \( F \) is the inbreeding coefficient [Ott, 1999]. HW genotype frequencies correspond to \( F = 0 \). Positive values of \( F \) can occur when there is inbreeding or stratification in the parental population.

We used an HW model and two types of departure from this model to generate the family genotypes. The HW model assumed HW frequencies for the null parental genotype distribution \( F = 0 \) with either \( p = 40\% \) (common variant allele) or \( p = 5\% \) (rare variant allele). The first type of departure from the HW model corresponds to population stratification. For this, we assumed that parents were drawn from a 50:50 admixture of two subpopulations, each with HW genotype frequencies, but with different variant allele frequencies. The frequencies of the variant allele in subpopulations 1 and 2 were, respectively, 20\% and 60\% (common variant) or 2.5\% and 7.5\% (rare variant). These assumptions imply that the genotype frequencies in the admixed population are given by (8) with \( p = 40\% \) or \( p = 5\% \), and with \( F = 1/6 \). In population 1, the null trait prevalence was 0.5\%, while in population 2 it was 11.5\%. With these assumptions, the null trait prevalence in the admixed population is 6\%. This departure represents strong population stratification that, when \( \theta = 0 \) and when unacknowledged in a case-control study of the admixed population, would almost certainly produce a spurious association. For example, when the variant allele is common, the pooled odds ratio (OR) in the admixed population is 3.22 (dominant model) or 2.25 (recessive model), despite the absence of association (OR=1) in both subpopulations.

In the absence of such strong population stratification, departures from HW frequencies corresponding to \( F = 1/6 \) are unlikely to occur in practice (see Discussion). Nevertheless, to allow for even more extreme HW deviations, we considered a second departure from the HW model, which assumed equations (8) and (9), with the same values for \( p \), intercept, and \( \theta \), but with \( F = 1/4 \).
In all models, we assumed independence of traits among relatives, conditional on their genotypes. We modelled the trait probability \( \varphi_j \) among individuals with genotype \( j \) as the logistic function

\[
\logit(\varphi_j) = -2.75 + \theta z_j.
\]  

(9)

Here \( z_j \) is one of two specified functions of the number \( j \) of variant alleles: \( z_j = j \) (the logit of risk in heterozygotes is half that of homozygous variant carriers, an additive genetic effect) or \( z_2 = 2, z_1 = z_0 = 0 \) (heterozygote risk equal to wild-type homozygote risk, a recessive genetic effect). For the alternative models, we took \( \theta = 0.576 \) (giving a relative risk of 2.8 for homozygous carriers relative to carriers of the wild type). The logistic intercept value \(-2.75\) corresponds to a null trait prevalence of 6% in the parental population.

In summary, the simulations used HW genotype frequencies and two types of departure, each with a common or a rare variant, each with \( \theta = 0 \) or \( \theta = 0.576 \), and each with an additive or a recessive genetic effect, for a total of \( 3 \times 2 \times 2 \times 2 = 24 \) models. Thus there were \( 24 \times 2 \times 2 = 96 \) simulations, corresponding to 24 models, two family traits \( t \), and two patterns of missing parental data.

RESULTS

TEST SIZE. When data were generated under one of the null models specifying \( \theta = 0 \), the FBAT and FGAP-NHW statistics performed well, with rejection rates close to the nominal type I error rate of 5%. However, for rare variant alleles (frequency \( p = 5\% \)), the FBAT recessive statistic tended to be conservative, rejecting in only 1% of the replications. In contrast, both the FGAP-HW statistic and the Rabinowitz statistic with sample variance showed inflated size in some circumstances. The size of the FGAP-HW additive statistic was inflated (with actual rejection rates as large as 26%) when parental genotypes were generated, with an inbreeding coefficient of \( F = 1/4 \). This inflation reflects overestimation of the number of heterozygote parents in the sample, with consequent underestimation of the variance of the number of variant allele transmissions and inflation of the test statistic. The size of the Rabinowitz recessive statistic with sample variance estimate was inflated (with rejection rates as large as 18%) when the variant allele frequency \( p \) was 5%. Since the Rabinowitz statistic is designed to have the correct asymptotic type I error rate, this inflation reflects downward bias in the sample variance estimate in the presence of rare alleles in samples of moderate size. In contrast, the Rabinowitz test with model-based variance exhibited accurate type I error rates in all situations studied.

TEST POWER

As expected, the FGAP-HW test had more power than the other tests when the parental genotypes were generated under HW and somewhat lower power when they were generated in other ways. Since the FGAP-HW and FGAP-NHW tests tended to have similar power, we restrict attention to the latter. Tables III and IV show the power of test statistics for data generated by models with \( F = 0, 1/6, \) or \( 1/4 \), with variant allele frequency 40% or 5%, and for an additive genetic effect (relative risk in AB heterozygotes compared to (wild-type) BB homozygotes given by \( R_{het} = 1.7 \)) and for a recessive genetic effect (\( R_{het} = 1 \)). In all models, the trait risk in wild-type homozygotes was 6%, and that in homozygous variant carriers was 17%. Table III gives results when both parental genotypes are missing, while Table IV refers to results for one missing parental genotype.

As can be seen from Tables III and IV, the FBAT test had less power than did the other two tests, with a greater power disadvantage when all three offspring were affected than when only 2 of 3 offspring were affected. Tables III and IV also show that the Rabinowitz test tended to have power comparable to that of the FGAP statistic, with some exceptions. First, the power of the Rabinowitz test was relatively poor when genotypes were missing for only one parent, regardless of how the variance was handled. (In Tables III and IV, column R1 gives results for the Rabinowitz test with sample variance, while column R2 gives results for the Rabinowitz test with model-based variance.) Moreover, although all tests had limited power to detect association with a recessively-acting variant allele of low (5%) prevalence, in this situation, the Rabinowitz statistic with a sample variance estimate performed particularly poorly, often rejecting more frequently under the null hypothesis than under the alternatives. This poor performance was not observed, however, when the test included the model-based variance.

When neither parent was typed, all three statistics had greater power when 2 of 3 offspring were affected than when all 3 were affected. This
phenomenon may reflect the fact that when all offspring are affected their genotypes tend to be similar, making it difficult to infer the parental genotypes. In contrast, the phenomenon was not evident when one parent was typed, probably because the additional typed parent provided

### TABLE III. Power of Association Tests (of Nominal Size $\alpha=0.05$) in Nuclear Families With Two Untyped Parents

<table>
<thead>
<tr>
<th>Parental genotype frequencies (%)</th>
<th>Two affected, one unaffected offspring</th>
<th>Three affected offspring</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterozygote relative risk $b$</td>
<td>FBAT</td>
<td>FGAP</td>
</tr>
<tr>
<td>HW (F=0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.7</td>
<td>77</td>
<td>80</td>
</tr>
<tr>
<td>0.2</td>
<td>42</td>
<td>48</td>
</tr>
<tr>
<td>Population stratification (F=1/6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.7</td>
<td>64</td>
<td>69</td>
</tr>
<tr>
<td>0.3</td>
<td>36</td>
<td>42</td>
</tr>
<tr>
<td>NHW (F=1/4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.7</td>
<td>66</td>
<td>73</td>
</tr>
<tr>
<td>0.2</td>
<td>67</td>
<td>94</td>
</tr>
<tr>
<td>Population stratification (F=1/6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.7</td>
<td>46</td>
<td>75</td>
</tr>
<tr>
<td>0.3</td>
<td>40</td>
<td>68</td>
</tr>
<tr>
<td>NHW (F=1/4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.7</td>
<td>70</td>
<td>96</td>
</tr>
<tr>
<td>0.2</td>
<td>70</td>
<td>96</td>
</tr>
</tbody>
</table>

*aFGAP and Rabinowitz tests are constructed without assuming Hardy-Weinberg (HW) parental genotype frequencies.

$b$Relative to BB genotype. Disease risk is 6% for BB genotype, and 17% for AA genotype.

$c$Rabinowitz statistic with sample variance estimate.

$d$Rabinowitz statistic with model-based variance.

### TABLE IV. Power of Association Tests (of Nominal Size $\alpha=0.05$) in Nuclear Families With One Untyped Parent

<table>
<thead>
<tr>
<th>Parental genotype frequencies (%)</th>
<th>Two affected, one unaffected offspring</th>
<th>Three affected offspring</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterozygote relative risk $b$</td>
<td>FBAT</td>
<td>FGAP</td>
</tr>
<tr>
<td>HW (F=0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.7</td>
<td>79</td>
<td>99</td>
</tr>
<tr>
<td>0.2</td>
<td>46</td>
<td>75</td>
</tr>
<tr>
<td>Population stratification (F=1/6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.7</td>
<td>67</td>
<td>94</td>
</tr>
<tr>
<td>0.3</td>
<td>40</td>
<td>68</td>
</tr>
<tr>
<td>NHW (F=1/4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.7</td>
<td>70</td>
<td>96</td>
</tr>
<tr>
<td>0.2</td>
<td>70</td>
<td>96</td>
</tr>
</tbody>
</table>

*aFGAP and Rabinowitz tests are constructed without assuming Hardy-Weinberg (HW) parental genotype frequencies.

$b$Relative to BB genotype. Disease risk is 6% for BB genotype, and 17% for AA genotype.

$c$Rabinowitz statistic with sample variance estimate.

$d$Rabinowitz statistic with model-based variance.
enough family genotype data to infer the unknown parental genotype with little uncertainty.

**DISCUSSION**

We considered three statistics for detecting trait-gene association in families, and showed that each can be constructed from a solution to a set of constraints on its null expectation. We also showed that while the FBAT and Rabinowitz statistics are optimal solutions for their constraints, the FGAP nonfounder statistic is not. In addition, we used simulations to examine the size and power of the three tests to detect association between a trait and a variant allele of a candidate gene, using nuclear families with three offspring. We considered several models for the variant allele frequencies in the parental population and for the effects of the variant allele on trait risk. The families varied with respect to the offspring traits and the number of missing parental genotypes. Several general conclusions seem warranted.

The simulations suggest that for the limited family structures and phenotypes considered, the Rabinowitz test and the FGAP test have greater power than does the FBAT test. The relatively poor power of FBAT reflects its heavy conditioning on the minimal sufficient statistic. For some pedigree/parameter sets, this conditioning excluded all but a few (<5) of the 100 pedigrees in many of the 1,000 iterations. For those sets in which FBAT included at least 8 of the 100 pedigrees in most iterations, it performed well in comparison with the FGAP and Rabinowitz statistics. FBAT's exclusion of a high proportion of pedigrees in some instances may also explain its conservative tendency in those cases, rejecting $H_0$ less often than its nominal type I error rate. A practical advantage of the FBAT test is the flexibility of the FBAT software, which can handle pedigrees more complicated than nuclear families. The good power of the Rabinowitz test is noteworthy, given that the statistic also has the correct asymptotic type I error rate, whatever the unknown parental genotypes, and whatever is specified for their distributions. This combination of good power and robustness to model misspecification recommends the Rabinowitz test as the method of choice, at least for data similar to those generated in these simulations. Moreover, the test with model-based variance performed better in detecting effects of rare alleles than did the one with sample variance estimate. At present, however, there is no simple algorithm for applying the test to a diverse set of nuclear families with different sizes, traits, and patterns of missing parental genotypes. It is now necessary to write programs specifically tailored to the families at hand.

The test based on the FGAP nonfounder statistic has two limitations. First, its type I error rate may not equal its nominal size when the family of distributions for the missing parental genotypes is misspecified. Second, under the restricted family for which its nominal and actual sizes agree, it is suboptimal. Nevertheless, the test without the assumption of HW parental genotype frequencies had the correct type I error rate and, overall, it performed best in the limited types of simulations considered here. The FGAP test assuming HW parental genotype frequencies showed inflated size in the presence of extreme departures from HW parental genotypes ($F=1/4$), but not in the presence of population stratification severe enough to produce an inbreeding coefficient of $F=1/6$. HW departures of such magnitude are unlikely to arise in practice. Estimates from human populations indicate that values for $F$ due to inbreeding are considerably smaller [Vogel and Motulsky, 1986]. (When genotypes of a sample of unrelated unaffected individuals are available, they can be used to evaluate the presence and extent of HW departures in the population from which the parents have been drawn.) Nevertheless, since little power is lost when the FGAP test is used without the HW assumption, there seems little reason to impose the assumption. The software to implement the FGAP nonfounder test also allows computation of the test statistic, based on the full efficient score for both parental and offspring genotypes. An advantage of working with the full efficient scores is that the observed or inferred parental genotypes can be used in a case-control type evaluation of association among the parents (see Whittemore and Tu [2000] for a description of the founder statistic and total statistic).

In conclusion, the simulations suggest that the FGAP and Rabinowitz tests have greater power than the FBAT test in some situations. The power advantage is important, because the ability of a family-based study to detect association between a trait and a gene of moderate penetrance should compare favorably to that of a case-control study. Family-based studies are more complex and involve more genotyping than case-control studies, and the possibility of false-positive
case-control findings due to confounding by population stratification remains controversial [Wacholder et al., 2000].

**ELECTRONIC DATABASE INFORMATION**

The Family-Based Association Test (FBAT) program may be found at http://www.bionstat.harvard.edu/~fbat. The Nonfounder or Partial Score Statistics programs may be found at http://www-gene.cimr.cam.ac.uk/clayton/software and at http://www.stanford.edu/dept/HRP/epidemiology/FGAP.

**ACKNOWLEDGMENTS**

The authors are grateful to Joseph B. Keller and Daniel Rabinowitz for helpful comments.

**REFERENCES**


**APPENDIX**

**PROOF THAT $\phi(0) = 0$**

Here $\phi(\theta) = \sum_i \sum_g e_i(g, t_i) P_{G=I|T_t}(g|t_i)$ is the expected value of the numerator of (1). We have

\[ P_{G=I|T_t}(g|t_i) = P_{G=I}(g) \]
\[ = \sum_{t=1}^n P_{L=I}(t) P_{G=I|L}(g|t) \]
\[ = \sum_{t=1}^n \gamma_t x_{igt}. \]  

Hence

\[ \phi(0) = \sum_{i=1}^n \sum_{g=1}^r e_i(g, t_i) P_{G=I|T_t}(g|t_i) \]
\[ = \sum_{i=1}^n \sum_{t=1}^r \sum_{g=1}^r e_i(g, t_i) x_{igt} \]
\[ = \sum_{i=1}^n \sum_{t=1}^r \gamma_t x_{igt} = 0 = 0. \]

**PROOF OF THEOREM**

a) Proof of assertion (a) follows closely the arguments of Rabinowitz [2002]. We wish to find, for fixed values $t_1, \ldots, t_n$, functions $e_i(g, t_i)$, $i = 1, \ldots, n$, to maximize (3) subject to constraints (2) and (4). We use the method of Lagrange multipliers to solve this problem. Specifically, we
replace the bracketed expression in (3) with
\[
Q = \theta^{-1} \left\{ \sum_{i=1}^{n} \sum_{g=1}^{n} e_i(g, t_i) P_{G_i|T_i}(g|t_i) \right. \\
- \sum_{i=1}^{n} \sum_{t=1}^{s_i} \beta_{it} \sum_{g=1}^{n} e_i(g, t_i) x_{igt} \right. \\
- \lambda \left. \left[ \sum_{i=1}^{n} \sum_{g=1}^{n} e_i^2(g, t_i) P_{G_i|T_i}(g|t_i) - C \right] \right\}. \tag{11}
\]

Here \(\beta_{i1}, \ldots, \beta_{in}, i = 1, \ldots, n\), and \(\lambda\) are Lagrange multipliers. The plan is to maximize \(Q\) with respect to \(e_i(g, t_i)\) for fixed \(g, t_i, \beta_i, \lambda, \theta \neq 0\), where \(\beta_i = (\beta_{i1}, \ldots, \beta_{in})\). This maximization gives the maximizing \(e_i(g, t_i)\) as functions of \(\beta_i, \lambda, \) and \(\theta\). We then take the limit as \(\theta \to 0\), and substitute the resulting \(e_i(g, t_i, \beta_i, \lambda)\) into (2) and (4) to obtain a system of \(1 + \sum_i s_i\) equations in the \(1 + \sum_i s_i\) unknowns \(\lambda, \beta_i, \) which we solve to obtain solutions \(\beta_i^*, \lambda^*, \lambda\). When the resulting \(e_i(g, t_i, \beta_i^*, \lambda^*)\) are substituted into (11), the second and third summands vanish by choice of \(\beta_i^*, \lambda^*, \lambda\), so that (11) equals the bracketed quantity in (3), and thus \(e_i(g, t_i, \beta_i^*, \lambda^*)\) maximizes (3), as required.

To maximize \(Q\), we differentiate (11) with respect to \(e_i(g, t_i)\), and equate the derivative to zero to obtain
\[
P_{G_i|T_i}(g|t_i) = \sum_{t=1}^{s_i} \beta_{it} x_{igt} + 2\lambda e_i(g, t_i) P_{G_i|T_i}(g|t_i). \tag{12}
\]

Next we subtract \(P_{G_i|T_i}(g|t_i)\) from both sides of (12), divide by \(\theta P_{G_i|T_i}(g|t_i)\), and use relation (10) to obtain
\[
\frac{P_{G_i|T_i}(g|t_i) - P_{G_i|T_i}(g|t_i)}{\theta P_{G_i|T_i}(g|t_i)} = \left[ P_{G_i|T_i}(g|t_i) \right]^{-1} \times \sum_{t=1}^{s_i} \beta_{it} x_{igt} + \lambda' e_i(g, t_i), \tag{13}
\]

where
\[
\beta_{it}' = (\beta_{it} - \gamma_t)/\theta, \quad \text{and} \quad \lambda' = 2\lambda/\theta.
\]

and \(\gamma_t = P_{L}(\ell)\). Taking the limit of both sides of (13) as \(\theta \to 0\) gives
\[
v_g = \frac{d}{d\theta} \log P_{G_i|T_i}(g|t_i)|_{\theta=0} = \left[ P_{G_i}(g) \right]^{-1} \sum_{t=1}^{s_i} \beta_{it}' x_{igt} + \lambda' e_i(g, t_i).
\]

Solving this equation for \(e_i(g, t_i)\) yields
\[
e_i(g, t_i) = \frac{1}{\lambda'} \left[ v_g - \sum_{t=1}^{s_i} \beta_{it}' x_{igt} \left[ P_{G_i}(g) \right]^{-1} \right]
\]
or in matrix form,
\[
E_i = \frac{1}{\lambda'} \left[ V_i - X_i W_i^{-1} X_i \beta \right]. \tag{14}
\]

We now substitute the right side of (14) for \(E_i\) in the constraint (2) to obtain an equation which we solve for \(\beta_i^*:\)
\[
\beta_i^* = (X_i^TW_i^{-1}X_i)^{-1}X_i^TV_i \tag{15}
\]

where \((X_i^TW_i^{-1}X_i)^{-1}\) denotes a generalized inverse if \(X_i^TW_i^{-1}X_i\) is not of full rank. We also substitute the right side of (14) for the \(e_i(g, t_i)\) in equation (4) and choose the arbitrary constant \(C\), so that \(\lambda' = 1\). By substituting 1 for \(\lambda'\) and (15) for \(\beta_i^*\) in (14), we find that the optimal \(E_i\) is given by (5), as required.

The null variance of \(e_i(G_i, t_i)\) is \(\sigma_i^2 = E_i^TW_iE_i\). Substituting (5) for \(E_i\) in this expression gives (6).

b) The contribution of family \(i\). The contribution of family \(i\) to the FBAT statistic is the efficient score corresponding to the conditional distribution of the observed family genotype data, given the trait value \(t_i\) and given the value of the minimal sufficient statistic \(L_i\) for the set of all null parental genotype distributions. (The sufficiency of \(L_i\) is defined by the property that, given the value of \(L_i\), the distribution of \(G_i\) is invariant with respect to all possible null parental genotype distributions.) Thus the \((g, \ell)^{th}\) entry of \(X_i\) is \(P_{G_i|L_i}(g|\ell)\), where \(\ell\) indexes the \(s_i\) possible values of \(L_i\). Standard likelihood theory [Cox and Hinkley, 1974] can be used to show that the FBAT statistic satisfies (2) and that it optimizes (3). It is also straightforward to represent the FBAT statistic in the form (5).

The Rabinowitz statistic is constructed from (5), with the \((g, \ell)^{th}\) entry of \(X_i\) given by the conditional probability of family genotype \(g\), given the value \(\ell\) for the two complete parental genotypes, where \(\ell\) indexes all \(s_i\) possible values of these genotypes, as illustrated in Table I (see also Rabinowitz [2002] for derivation and discussion of this statistic).

For the nonfounder FGAP statistic, the \((g, \ell)^{th}\) entry of \(X_i\) is the null probability of the observed family genotype \(g\), given the value \(\ell\) of any observed parental genotypes. If both parental genotypes are missing, \(X_i\) is a column vector with entries \(x_{ig} = P_{G_i}(g), g = 1, \ldots, r_i\). When one or both parental genotypes are missing, these entries
require specifying a null parental genotype distribution (see Table 1). To describe the FGAP statistic, we write \( G_i = (K_i, H_i) \), where \( K_i \) denotes the observed portion of the parental genotype, and \( H_i \) denotes the offspring genotype. We also let \( P_i = (K_i, M_i) \) denote the complete parental genotype, so that the complete family data is \( C_i = (P_i, H_i) \). As shown by Clayton [1999] and Whittemore and Tu [2000], the efficient score can be written

\[
V_i = V_{1i} + V_{2i}
\]

where \( V_{1i} \) and \( V_{2i} \) have \( g^i \) components

\[
v_{1ig} = \sum_j P^0_{jP|G_i}(j|g)v_{1ij}, \quad \text{and}
\]

\[
v_{2ig} = \sum_j P^0_{jP|G_i}(j|g)v_{2ij}.
\]

Here \( j \) indexes the set of all possible complete parental genotypes \( P_i \), and

\[
v_{1ij} = \frac{d}{d\theta} \log P^0_{iP|T_i}(j|i_{hi}) \bigg| \theta = 0 \quad \text{and}
\]

\[
v_{2igj} = \frac{d}{d\theta} \log P^0_{G_i|P_i}(g|j,t_i) \bigg| \theta = 0
\]

denote the efficient scores for the case of complete parental genotypes. For the FGAP nonfounder statistic, \( E_i = V_{2i} \). To see that this vector satisfies the constraint \( X_i^T V_{2i} = 0 \) for the FGAP matrix \( X_i \), we note that

\[
X_i^T V_{2i} = \sum_{g=1}^{r_i} P^0_{G_i|K_i}(g|\ell) v_{2ig}
\]

\[
= \sum_{g=1}^{r_i} P^0_{G_i|K_i}(g|\ell) \left[ \sum_j P^0_{jP|G_i}(j|g) v_{2ij} \right]
\]

\[
= \sum_{g=1}^{r_i} P^0_{G_i|K_i}(g|\ell) \left[ \sum_j P^0_{M_i|K_i}(j|\ell) P^0_{H_i|P_i}(h|j,\ell) \right] v_{2ig}
\]

\[
= \sum_j P^0_{M_i|K_i}(j|\ell) \left[ \sum_{g=1}^{r_i} P^0_{H_i|P_i}(h|j,\ell) v_{2ij} \right]
\]

\[
= \sum_j P^0_{jP|G_i}(j|\ell) \cdot 0 = 0
\]

as required. However, unless both parental genotypes are observed, \( V_{21}, \ldots, V_{2n} \) is not the optimal solution (5). The latter is

\[
E_i = V_{2i} + \left[ V_{1i} - W_i^{-1} X_i (X_i^T W_i^{-1} X_i)^{-1} X_i V_{1i} \right]
\]

\[
= V_{2i} + Z_i
\]

where \( Z_i \) is the \( r_i \)-dimensional column vector whose \( g^i \) component is

\[
\sum_j \left[ P_{jP|G_i}(j|g) - P_{jP|K_i}(j|\ell) \right]
\]

\[
\times \frac{d}{d\theta} \log P^0_{jP|T_i}(\ell, j_{ti}) \bigg| \theta = 0.
\]

Note that \( Z_i = 0 \) when both parental genotypes are observed, and \( Z_i = V_{1i} \) when neither parental genotype is observed. In the latter case, the optimal statistic is constructed from the full efficient score vectors \( V_i, i = 1, \ldots, n \).

c) When both parents have been typed. It is clear from the matrices \( X_i \) for each of the three statistics that when both parents have been typed, each column of each \( X_i \) gives the null probabilities of the family data \( g \) conditional on a specific value for the parental genotypes. The assertion now follows from standard likelihood theory.

d) Column space of the Rabinowitz matrix. To show that the column space of the Rabinowitz matrix contains that of the FGAP matrix, we write the entries of the latter matrix as

\[
x_{gk}^{FG} = \sum_{g=1}^{r_i} P^0_{G_i|K_i}(g|k) x_{gk}^R = \sum_{g=1}^{r_i} P^0_{G_i|K_i}(g|k) P^0_{K_i|P_i}(k|g) = \sum_{i,j} c_{ijk} x_{gij}^R.
\]

To show that the column space of the FBAT matrix contains that of the Rabinowitz matrix, let \( L_i \) denote the FBAT minimal sufficient statistic for family \( i \). The sufficiency of \( L_i \) for the parental genotypes allows us to write

\[
x_{gij}^R = \sum_{j} P_{jP_i}(j|\ell) P^0_{G_i|P_i}(g|\ell) = \sum_{\ell} c_{ij\ell} x_{gij}^{FB}.
\]

This completes the proof.