Family-Based Tests of Association and Linkage That Use Unaffected Sibs, Covariates, and Interactions

Kathryn L. Lunetta,[1,2] Stephen V. Faraone,[3,4] Joseph Biederman,[3,5] and Nan M. Laird[2]

[1]Department of Biostatistical Science, Dana-Farber Cancer Institute, [2]Department of Biostatistics, Harvard School of Public Health, [3]Psychopharmacology Unit of the Child Psychiatry Service, Massachusetts General Hospital, [4]Harvard Medical School Department of Psychiatry at the Massachusetts Mental Health Center, and [5]Department of Psychiatry, Harvard Medical School, Boston

Summary

We extend the methodology for family-based tests of association and linkage to allow for both variation in the phenotypes of subjects and incorporation of covariates into general-score tests of association. We use standard association models for a phenotype and any number of predictors. We then construct a score statistic, using likelihoods for the distribution of phenotype, given genotype. The distribution of the score is computed as a function of offspring genotypes, conditional on parental genotypes and trait values for offspring and parents. This approach provides a natural extension of the transmission/disequilibrium test to any phenotype and to multiple genes or environmental factors and allows the study of gene-gene and gene-environment interaction. When the trait varies among subjects or when covariates are included in the association model, the score statistic depends on one or more nuisance parameters. We suggest two approaches for obtaining parameter estimates: (1) choosing the estimate that minimizes the variance of the test statistic and (2) maximizing the statistic over a nuisance parameter and using a corrected *P* value. We apply our methods to a sample of families with attention-deficit/hyperactivity disorder and provide examples of how covariates and gene-environment and gene-gene interactions can be incorporated.

Introduction

Rubinstein et al. (1981) and Falk and Rubinstein (1987) first suggested using the transmitted and nontransmitted parental haplotypes as cases and controls in association

tests using parent–affected-child trios. The transmitted and nontransmitted alleles can also be considered to be paired observations, leading to a McNemar's test (Terwilliger and Ott 1992), or what is commonly called the "transmission/disequilibrium test" (TDT [Spielman et al. 1993]).

Many extensions and modifications of the TDT have been developed. Here we build on a general approach developed by Self et al. (1991) and Schaid (1996) for parent–affected-child trio data. In the work of Self et al. (1991), conditional logistic regression is used to model P(genotype|disease). The affected child is the case and is matched to three controls, corresponding to the other three marker genotypes that the parents of the case could have produced. Score tests are used to test for marker-disease association. Schaid (1996) extends these methods and proposes a general framework to test for association with multiallelic markers, on the basis of log-risk models for dichotomous phenotypes. Here, we extend this methodology in several ways. First, we consider general association models for an arbitrary phenotype Y and score statistics based on likelihoods for the distribution of Y, given genotype. These score statistics provide a natural approach to incorporation of unaffected offspring or for use of measured phenotypes. Although we use an association model to obtain the score statistic, the distribution of the score is computed as a function of offspring genotypes, conditional on parental genotypes and offspring trait values, so the method does not depend on the model when no covariates are used. Second, we use association models that may include environmental factors, as well as multiple genes. These models permit us to study gene-environment and gene-gene interactions.

One distinction between our procedure and previous work discussed here is apparent when we apply the admixture correction. Self et al. (1991) and Schaid (1996) model P(genotype|disease), thus applying the correction for admixture before computing the score; we model P(disease|genotype) and apply the admixture correction after computing the score. Although these two approaches are equivalent and, in specific cases, result in the same tests, modeling P(disease|genotype) rather than

P(genotype|disease) gives us a natural way to model continuous phenotypes or both unaffected and affected offspring for dichotomous traits, and allows the seamless inclusion of covariates and gene-covariate interactions. Another important distinction between our work and that of Schaid (1996) is that Schaid considered settings with no nuisance parameters. When the trait (Y) varies among subjects in the sample, either (*a*) because we have both affected and unaffected subjects or a quantitative phenotype or (*b*) when we introduce covariates, the score statistics involve nuisance parameters, which cannot always be estimated from the data (see Rabinowitz 1997). When adequate estimates of nuisance parameters are not available, we suggest one approach to specification of score statistics that is based on optimizing the χ^2 statistic (Davies 1977) and another that is based on minimizing the variance of the score.

We illustrate the methods with a sample of 43 nuclear families, each having at least one member with attention-deficit/hyperactivity disorder (ADHD), and a total of 44 affected and 34 unaffected children. We examine associations and interactions between ADHD and the dopamine D4 (DRD4) and dopamine transporter (DAT) genes, using sex and parental affection status as covariates. For the ADHD phenotype, both a dichotomous affection status and a measured phenotype score are used. For tests without covariates, we compare results by using population and sample estimates of a single nuisance parameter and the optimized χ^2 statistics with the *P*-value correction described by Davies (1977).

Methods

Defining the test statistic involves three steps. First, we describe an association model, using standard generalized linear-regression models (McCullagh and Nelder 1989), which relate the mean phenotype to the marker or gene being tested and, possibly, to other covariates. Second, we construct a likelihood, using exponential family models, and use the likelihood to obtain a score statistic. Finally, the admixture adjustment is accomplished by calculating the mean and the variance of the score statistic by use of the distribution of genotype in offspring, conditional on parental genotypes and on offspring phenotypes. In this section, we define each of these three steps in sequence, then show how covariates, gene-environment interaction, and gene-gene interactions may be tested.

First, we introduce some notation. We assume that there are N independent families indexed by i, each having n_i offspring, indexed by $j = 1, \ldots, n_i$. Let Y_{ij} denote the phenotype of the jth offspring in the ith family, and let μ_{ij} denote $E(Y_{ij})$.

Models for the Association

There are many ways to characterize the effect of a gene on phenotype; three traditional genetic models are recessive, dominant, and additive. An advantage of these models is that they each yield a 1-df test, or, equivalently, that they each can be coded in a regression model that uses a single variable. The additive model is the basis for both the TDT (Spielman et al. 1993) and the sibling TDT (S-TDT [Spielman and Ewens 1998]) and will be used in our examples and application. Schaid and Sommer (1994) propose transmission statistics based on recessive and dominant models. Sham and Curtis (1995) consider models for multiallelic markers or genes, and Schaid (1996) proposes a general framework to test for association with multiallelic markers, on the basis of log-risk models for dichotomous phenotypes. For the purposes of this study, we will consider biallelic markers (or testing a single allele versus all others in a multiallelic setting), but the general approach extends straightforwardly when the genotype is coded by use of more than one variable.

Let X_{ij} denote the variable that codes for genotype. Assume that we are interested in testing a particular allele, labeled "A." Then, for the additive model, X_{ij} counts the number of A alleles in the ijth individual. For the recessive model, $X_{ij} = 1$ if the ijth individual has genotype AA and is 0 otherwise, etc. The generalized linear model assumes a link function, say L_{ij}, which is some transformation of μ_{ij}, such that

$$L_{ij} = \beta_0 + \beta_1 X_{ij} \ . \tag{1}$$

With dichotomous phenotypes, the natural link function is the logit $L_{ij} = \text{logit}(\mu_{ij}) = \log[\mu_{ij}/(1 - \mu_{ij})] = \beta_0 + \beta_1 X_{ij}$, where $\mu_{ij} = E(Y_{ij})$ is the disease prevalence. For continuous responses, the natural link is the identity, meaning that $L_{ij} = \mu_{ij}$ and that the association model is simply the usual linear-regression model: $L_{ij} = \mu_{ij} = \beta_0 + \beta_1 X_{ij}$. As discussed in the next section, these models can be made more complex by adding additional terms for covariates and other genes.

Score Statistics

A general approach that has been used to test association models is the derivation of the conditional distribution of offspring genotype, given their phenotype and parental genotypes; this is used to form a likelihood. Schaid and Sommer (1994) use this approach to extend the TDT to recessive and dominant models, and Schaid (1996) develops a very general likelihood approach, using a log-risk model. These investigators use affected offspring only and derive the likelihood associated with the conditional distribution of offspring genotype, given parental genotype and offspring phenotype. Parents can

have any number of offspring; all offspring with known phenotype and genotype can be included in the likelihood. Unlike Schaid and Sommer (1994) and Schaid (1996), we compute the prospective likelihood of phenotype Y_{ij}, conditioning on genotype X_{ij} to obtain the score statistic; sibs are treated as independent, given genotype. The resulting likelihood is used to obtain the score statistic, and the adjustment for admixture is done as the final step, by computation of the distribution of the score statistic, by use of the genotype distribution of offspring, described below.

By use of exponential family models, Bernoulli for dichotomous phenotypes and normal for continuous phenotypes, the log likelihood for both models can be written as $\log L(\beta_0, \beta_1) = \Sigma_{ij}[Y_{ij}L_{ij} - a(L_{ij})]$, where $a(L_{ij})$ is a function of L_{ij}, with the property that $\partial a(L_{ij})/\partial L_{ij} = \mu_{ij}$ when L_{ij} is the canonical link function. To test the null hypothesis of no association, we obtain the first derivative of the log likelihood with respect to β_1 and set β_1 equal to 0 in the resulting equation. The first derivative of the log likelihood with respect to β_1 is $[(\partial \log L)/\partial \beta_1] = \Sigma_{ij}X_{ij}(Y_{ij} - \mu_{ij})$. Setting $\beta_1 = 0$ yields the score statistic S: $S = \Sigma_{ij}X_{ij}(Y_{ij} - \mu)$, where, under $H_0 : \beta_1 = 0$, μ is constant for all subjects. In this setting, the statistic will be the same for any link function. Notice that, in the case of dichotomous phenotypes, the use of only affected individuals means that $Y_{ij} = 1$ for all ij. In this case, $(Y_{ij} - \mu) = (1 - \mu)$ acts as a multiplicative constant and can be ignored, since it will vanish in the normalization of the test statistic. The resulting statistic, $S = \Sigma_{ij}X_{ij}$, is identical to that used in the TDT and the S-TDT.

In general, however, S depends on the nuisance parameter μ. Since μ is not a function of the genotype, misspecification of μ will not bias the test, but a good choice of μ can improve test efficiency. We discuss this in some detail in the examples.

Distribution of the Test Statistic

We evaluate the distribution of the test statistic by using the appropriate permutation distributions for the offspring allele values as described by Kaplan et al. (1997) and Rabinowitz and Laird (in press). When all parental marker data are known, under the null hypothesis of no linkage—$H_0 : \theta = \frac{1}{2}$—the permutation distribution of alleles in the offspring follows the usual Mendelian laws, conditional on known founder alleles. This case is particularly simple, since it implies that all offspring in a family are independent, conditional on the genotypes of the parents. The algorithm for nuclear families is described by Kaplan et al. (1997). Rabinowitz and Laird (in press) describe the algorithm for the case in which parental genotype data are missing.

There are two ways that one can use the permutation

distribution. If the score statistic is sufficiently simple and the number of pedigrees is sufficiently large, we can use the distribution to compute the mean and variance of each pedigree's contribution to the statistic. Then an approximate Z score is $Z = \{\Sigma_i[S_i - E(S_i)]\}/\sqrt{\Sigma_i \mathrm{var}(S_i)}$.

In the next section, we discuss vectors of statistics. In this case, for $M \geqslant 2$ score statistics, we construct an asymptotic χ^2 statistic with M df, which takes the form $\chi^2 = [\mathbf{S} - E(\mathbf{S})]^T \mathbf{V}^{-1}[\mathbf{S} - E(\mathbf{S})]$, where $\mathbf{S}$ is the vector of score statistics summed over all families and $E(\mathbf{S})$ and $\mathbf{V} = \mathrm{var}(\mathbf{S})$ are its mean and variance, respectively, computed under the permutation distribution. Alternatively, we may use the Monte Carlo method to evaluate the entire null distribution of the test statistic and thereby obtain exact P values, as do Kaplan et al. (1997).

Covariates and Interactions

We use the term "covariate" to encompass not only traditional environmental exposures but also the genotype at a gene known to affect the trait, provided that the gene is not linked to the locus being tested. We also assume that the covariate(s) are not affected by any gene linked to the tested locus. Because the adjustment for admixture also removes any bias due to confounding, it is not necessary to use covariates in the association model, but doing so may increase efficiency if the covariate is strongly predictive of phenotype. In principle, including a covariate as an additional term—say $\beta_2 Z_{ij}$, where Z_{ij} is the covariate value for the ijth subject—in the basic regression model (1) can easily do this. In the score function, then, μ becomes μ_{ij}, where now the associated link function under H_0 is

$$L_{ij} = \beta_0 + \beta_2 Z_{ij} \ . \tag{2}$$

Using a covariate requires estimates of β_0 and β_2, which are used to calculate the residuals $(Y_{ij} - \mu_{ij})$ used in the score function.

Gene-environment interactions.—The usual statistical approach to test for an interaction between two variables X_{ij} and Z_{ij} is to specify an association model, such as

$$L_{ij} = \beta_0 + \beta_1 X_{ij} + \beta_2 Z_{ij} + \beta_3 X_{ij} Z_{ij} \ , \tag{3}$$

and to test the interaction by setting $\beta_3 = 0$. In this setting, however, the reference distribution of the test statistic under H_0 is always computed under the assumption of no linkage and no linkage disequilibrium, which implies that $\beta_1 = 0$ as well. Therefore, the testable null hypothesis is $H_0 : \beta_1 = \beta_3 = 0$. One can use a vector of test statistics, obtained by differentiating the log likelihood with respect to both β_1 and β_3, $S_1 = \Sigma_{ij}X_{ij}(Y_{ij} - \mu_{ij})$ and $S_2 = \Sigma_{ij}X_{ij}Z_{ij}(Y_{ij} - \mu_{ij})$, where, under H_0, μ_{ij} is given as the antilink of equation (2). Using the vector of score statistics will yield a 2-df test. This approach is not sen-

sitive to the way in which the covariate Z_{ij} is coded, since replacing Z_{ij} by $(a + bZ_{ij})$, where a and b are arbitrary constants, will not change the value of the test statistic. A drawback of this approach is that global test statistics are not sensitive to particular alternatives. Furthermore, rejection of H_0: no linkage and no linkage disequilibrium may not imply interaction; if the test has sufficient power, the test should reject even if there is no interaction, as long as there is a main effect.

An alternative is to construct 1-df tests by use of just S_2. Such tests are valid, since they always have the required distribution under H_0, but now the coding of Z_{ij} is critical to the interpretation of the result. For example, suppose that the covariate is a dichotomous variable, the presence or absence of some environmental exposure. Then setting $Z_{ij} = 1$ if exposed and 0 if not exposed (or if exposure is missing) yields a 1-df test that follows the reference distribution, unless the marker is linked to and in linkage disequilibrium with a gene influencing the trait in exposed individuals. In this case, individuals with $Z_{ij} = 0$ do not contribute to the statistic. Reversing the coding to $Z_{ij} = 1$ if unexposed and 0 otherwise yields a test statistic that follows the reference distribution, unless the marker is linked to a gene influencing the trait among unexposed subjects. These two tests are independent, since they use different subjects, but they do not provide a direct test of interaction. Using the coding 1 or −1 for, respectively, exposed or unexposed status, yields a 1-df test that contrasts the association among exposed with that among unexposed. The power of this test, however, depends heavily on choosing the correct association model.

Gene-gene interactions.—If Z_{ij} is not an established risk factor for the phenotype but is a second locus unlinked to the locus being tested, we can use model (3) to test the effects of both genes simultaneously. The null hypothesis is $\beta_1 = \beta_2 = \beta_3 = 0$, since the reference distribution of the test statistic under H_0 is computed under the assumption of no linkage and no linkage disequilibrium for either gene. The vector of test statistics will include a third statistic, $S_3 = \Sigma_{ij} Z_{ij}(Y_{ij} - \mu_{ij})$, and $\mu_{ij} = \mu$ is a constant given by the antilink of $L_{ij} = \beta_0$. The reference distribution is computed by treating X_{ij} and Z_{ij} as random and independent. Using the vector of score statistics will yield a 3-df test. Note that a main effects–only model can be tested by use of S_1 and S_3 only.

Examples of the General Approach

We now use the general approach to derive some statistics already in the literature and to derive some new statistics. For simplicity of presentation, we assume that only one allele, denoted by "A," is of interest and that X_{ij} is the number of A alleles in the jth offspring of the ith family. Under the null hypothesis of no linkage and

no linkage disequilibrium, when the genotypes of both parents are known and we use the additive association model, all offspring are independent, and the mean and variance of X_{ij} for all offspring are $E(X_{ij}) = N_i^{AA} + N_i^A/2$ and $\mathrm{var}(X_{ij}) = N_i^A/4$, where N_i^A and N_i^{AA} are the number of parents who are, respectively, heterozygous and homozygous for allele A in the ith family.

The TDT with affected and unaffected offspring.—The TDT was proposed to test linkage via association of a particular allele A with disease, by use of affected offspring and their parents, and can be derived as a score test (e.g., see Schaid 1996). When both affected and unaffected offspring are used and the phenotype is coded $Y_{ij} = 1$ if affected and 0 if unaffected, the score statistic can be written as $S = \Sigma_{ij} X_{ij}(Y_{ij} - \mu) = (1 - \mu)S_a - \mu S_u$, where S_a is the total number of A alleles transmitted to the affected offspring and S_u is the same for the unaffected offspring. Note that transmissions from homozygous parents contribute nothing to the statistic, since their contribution to X_{ij} always equals their contribution to $E(X_{ij})$ and their contribution to $\mathrm{var}(X_{ij}) = 0$. Setting $\mu = 0$ means that $S = S_a$ and only transmissions to affected offspring are counted. When the disease is rare, $\mu \approx 0$, and most of the linkage disequilibrium can be expected to occur in the genotypes of affected individuals, since the allele frequency among unaffected individuals will be close to that of the population. With more common diseases, including the unaffected individuals by taking $0 < \mu < 1$ can increase the power of the test.

When the genotypes of both parents are known, it is straightforward to show that, under H_0: no linkage and no association, $S_a - E(S_a) = [(b_a - c_a)/2]$ and $\mathrm{var}(S_a) = [(b_a + c_a)/4]$, where b_a and c_a are, respectively, the number of times that A is transmitted and not transmitted from heterozygous parents to the affected offspring, so $b_a + c_a$ is the total number of transmissions to affected individuals scored. Analogous calculations can be made for the unaffected individuals. Hence, under H_0,

$$S - E(S) = \frac{1}{2}[(1 - \mu)(b_a - c_a) - \mu(b_u - c_u)]$$

and

$$\mathrm{var}(S) = \frac{1}{4}[(1 - \mu)^2(b_a + c_a) + \mu^2(b_u + c_u)] . \quad (5)$$

Thus, the statistic for affected and unaffected offspring is

$$\chi_{au}^2 = \frac{[(1 - \mu)(b_a - c_a) - \mu(b_u - c_u)]^2}{(1 - \mu)^2(b_a + c_a) + \mu^2(b_u + c_u)} .$$

This 1-df χ^2 statistic can be viewed as taking a weighted average of the contributions from the affected and un-

affected sibs. When $\mu = 0$, only affected individuals are used, and χ^2_{au} is the TDT of Spielman et al. (1993). When μ is set to the population prevalence of the disease, χ^2_{au} is identical to the T_{sib} statistic derived by Whittaker and Lewis (1998) as the most powerful test under a multiplicative-genotype relative-risk model for disease.

Under the alternative hypothesis, note that the segregation proportions for affected and unaffected children,

$$\pi_a = P(\text{transmit } A \text{ to}$$

$$\text{affected child} \mid \text{heterozygote parent})$$

and

$$\pi_u = P(\text{transmit } A \text{ to}$$

$$\text{unaffected child} \mid \text{heterozygote parent})$$

should be on different sides of $\frac{1}{2}$ and, thus, that $\pi_a - \frac{1}{2}$ and $\pi_u - \frac{1}{2}$ will have different signs. With a rare disease allele in linkage disequilibrium with A, one would expect the magnitude of $\pi_a - \frac{1}{2}$ to be substantially greater than that of $\pi_u - \frac{1}{2}$. Hence, including genotyped, unaffected offspring in an association test will likely increase power substantially only when the disease allele is common or when the total number of transmissions scored to unaffected individuals is much greater than the number scored to affected individuals.

In theory, μ is the prevalence of disease in the population. If we have both affected and unaffected offspring, we can estimate μ with the proportion affected in the sample, but, with selected samples, it may be more appropriate to use an externally derived estimate of disease prevalence. An alternative approach is to minimize var(S) under H_0, as a function of μ. By use of the equation for var(S) given in equation (5), it is straightforward to show that, under H_0, var(S) is minimized by setting $\mu = [n_a/(n_a + n_u)]$, where $n_a = b_a + c_a$ and $n_u = b_u + c_u$. This is a sample estimate of the prevalence, where offspring are weighted by the number of heterozygous parents. We call this the "heterozygous weighted-sample estimate" (HWSE).

Yet another approach is to maximize the statistic over $0 \leq \mu \leq 1$ and then adjust the P value of the test to reflect the maximization. Davies's (1977) method of adjustment provides an upper bound on the significance level of the test. To get the upper bound, we first identify a statistic $Z(\mu)$ that is dependent on the weight μ and that has a standard normal distribution under H_0. In our case, $Z(\mu) = \sqrt{\chi^2_{au}}$. The upper bound is

$$P\left[\sup_{0 \leq \mu \leq 1} |Z(\mu)| > x\right]$$

$$\leq 2\left\{\Phi(-x) + \frac{\exp\left(-\frac{1}{2}x^2\right)}{2\pi}\int_0^1 [-\rho_{11}(\mu)]^{1/2}\,d\mu\right\},$$

where, under H_0, $\rho_{11}(\mu) = \{[\partial^2\rho(\mu_1,\mu_2)]/\partial\mu_1^2\}_{|\mu_1=\mu_2}$ and $\rho(\mu_1,\mu_2)$ is the covariance between $Z(\mu_1)$ and $Z(\mu_2)$ (Davies 1977). Writing var(S_a) $= \sigma_a^2$ and var(S_u) $= \sigma_u^2$, one can show that $\rho_{11}(\mu) = \{(-\sigma_a^2\sigma_u^2)/[(1-\mu)^2\sigma_a^2 + \mu^2\sigma_u^2]^2\}$ and that

$$\int_0^1 [-\rho_{11}(\mu)]^{1/2}\,d\mu = \int_0^1 \frac{\sigma_a\sigma_u}{(1-\mu)^2\sigma_a^2 + \mu^2\sigma_u^2}\,d\mu = \frac{\pi}{2}\,.$$

Therefore, the upper bound for a two-sided test is

$$P\left(\sup_{0 \leq \mu \leq 1} |Z(\mu)| > x\right) \leq 2\left[\Phi(-x) + \frac{1}{4}\exp\left(-\frac{1}{2}x^2\right)\right]\,. \quad (6)$$

It is also possible to compute an upper bound for the statistic for quantitative traits. However, that bound is more difficult to compute and depends on the trait values.

Segregation-distortion tests.—χ^2_{au} differs from the 1-df χ^2 test for segregation distortion, "χ^2_s," proposed by Spielman et al. (1993). χ^2_s is a Pearson χ^2 test of independence and is designed to test $H_0: \pi_a = \pi_u$, whereas χ^2_{au} is designed to test $H_0: \pi_a = \pi_u = \frac{1}{2}$. If we choose μ to be $\hat{\mu}_0 = n_a/(n_a + n_u)$, then $S - E(S)$ also has expectation 0 under the segregation-distortion hypothesis $\pi_a = \pi_u$. However, as a general test of segregation distortion, χ^2_{au} is conservative, since the variance is calculated under the hypothesis that the common π equals $\frac{1}{2}$. We can easily derive an alternative to χ^2_s to test $H_0: \pi_a = \pi_u$, using the score statistic S but evaluating its distribution under a different null hypothesis. To find $E(S)$ and var(S) under $H_0: \pi_a = \pi_u$, note that, for a heterozygous parent, $E(T_{ij}) = \pi$, and var($T_{ij}) = \pi(1-\pi)$, where T_{ij} is 1 or 0, depending on whether A is or is not transmitted to the child by a heterozygous parent, and where π is the common segregation proportion. Under $H_0: \pi_a = \pi_u$, all transmissions are independent, and π can be estimated, from all of the offspring transmissions, as $\hat{\pi} = (b_a + b_u)/(n_a + n_u)$. The test statistic is therefore

$$\chi^2_{sd} = \frac{[(1-\mu)(b_a - n_a\hat{\pi}) - \mu(b_u - n_u\hat{\pi})]^2}{\hat{\pi}(1-\hat{\pi})[(1-\mu)^2 n_a + \mu^2 n_u]}\,.$$

Here, if we substitute $\hat{\mu}_0 = n_a/(n_a + n_u)$ for μ, then χ^2_{sd} is identical to χ^2_s.

TDT for quantitative traits.—Rabinowitz (1997) proposed a score test of H_0: no linkage and no linkage disequilibrium for quantitative traits given by

$$Z = \frac{\sum_{ij}(Y_{ij} - \overline{Y})(X_{ij} - E(X_{ij}))}{\sqrt{\sum_{ij}(Y_{ij} - \overline{Y})^2 \, \text{var}(X_{ij})}} \quad .$$

Here again, it is straightforward to minimize the variance of the score test under H_0 as a function of μ, and the resulting statistic is the same as that proposed by Rabinowitz, except that $\overline{Y}$ is replaced by the HWSE—that is, the sample average Y_{ij}—weighting each offspring by the number of heterozygous parents.

Inclusion of covariates.—As discussed in the Methods section, we may be able to increase efficiency when testing an allele for association, by including in our model one or more covariates as main effects. In this case, μ is replaced by μ_{ij}, which depends on the regression model for the mean as a function of covariates. For example, if the covariate Z_{ij} is a dichotomous indicator of some environmental exposure, then there will be two values for μ_{ij}—the prevalence of the disorder, or the mean phenotype—in the presence and absence of the exposure. Heterozygote weighted-sample averages can be used to estimate the two values of μ_{ij} by use of the exposed and unexposed groups. Population parameters can be used when available, and, in principle, Davies's (1977) approach could also be generalized to more than one parameter.

Testing strategies and interpretation for gene-environment and gene-gene interaction tests.—One of the drawbacks of the global 2-df test that includes a gene-covariate interaction is that it is not sensitive to particular alternatives. The global 2-df test provides the most information in instances in which the 1-df test of the gene's main effect does not reject H_0 but the 2-df test does. This case strongly suggests that an interaction is present.

If the 1-df main-effect test rejects H_0 and there is sufficient power, then the 2-df test that includes the interaction will reject it as well. In this case, we have not gained much information about gene-covariate interactions. As an ad hoc procedure, it is logical to compare the P values of the 1-df main-effect–only and 2-df interaction test. If there is no interaction, we would expect the P value of the 2-df test to be similar to or greater than the P value for the 1-df test; if an interaction exists, we would expect the P value to be smaller for the 2-df test. Alternatively, one might look at the test based on S_2 alone, for an appropriate choice of Z.

Application to Families with ADHD

Researchers have examined candidate genes in dopamine pathways because animal models, theoretical considerations, and the effectiveness of stimulant treatment implicate dopaminergic dysfunction in the pathophysiology of ADHD (Faraone and Biederman 1998).

A population-based association study implicated the A1 allele of the dopamine D2–receptor gene (Comings et al. 1991) for ADHD, but no attempts to replicate this have been reported. Cook et al. (1995) reported an association between ADHD and the 480-bp allele of the DAT gene (DAT-480), using a family-based association study. This finding was replicated in family-based studies of ADHD by Gill et al. (1997), Waldman et al. (1998), and Daly et al. (1999) but not in other studies (Asherson et al. 1998; Poulton et al. 1998). Several groups have reported an association between ADHD and the seven-repeat allele of DRD4 (DRD4-7 [LaHoste et al. 1996; Rowe et al. 1998; Smalley et al. 1998; Swanson et al. 1998; Faraone et al. 1999; Comings et al. 1999]), but other groups could not replicate this association (Asherson et al. 1998; Castellanos et al. 1998; Daly et al. 1998). The positive DRD4 results have generated much interest, because the implicated allele is known to mediate a blunted response to dopamine (Asghari et al. 1995).

As part of an ongoing study, we recruited nuclear families in which at least one family member (child or parent) was believed to have ADHD. All individuals were assessed for DSM-IV ADHD, and DNA was genotyped for the DRD4 and DAT genes. Sixty parent-child trios (33 affected and 27 unaffected children, in 35 nuclear families) were genotyped for DAT, and an additional 12 trios were genotyped for DRD4, resulting in a total of 72 trios, with 42 affected and 30 unaffected children, in 39 nuclear families. For DAT-480, a total of 17 unaffected children had at least one heterozygous parent; 2 of these 17 had two heterozygous parents. A total of 21 affected children had at least one heterozygous parent; 6 of these 21 had a second heterozygous parent. For DRD4-7, 13 unaffected children had at least one heterozygous parent; and 2 of these 13 had two heterozygous parents. A total of 17 affected children had at least one heterozygous parent; 4 of these 17 had two heterozygous parents. For all analyses, we treated each gene as biallelic, testing DRD4-7 against all other DRD4 alleles and testing DAT-480 against all other DAT alleles.

In table 1, we summarize the results of our analyses for each allele separately, without incorporating covariates into the weights. For each allele, we display the P values for statistics computed by use of six different values for the weight (nuisance parameter) μ. Weight $\mu = 0$ corresponds to the traditional TDT test of Spielman et al. (1993), which uses only affected offspring. Weights $\mu = .05$ and $\mu = .10$ correspond to the range of ADHD prevalence rates reported in the literature. The HWSE is the sample estimate of the prevalence, in which offspring are weighted by the number of heterozygous parents. This is the μ estimate, which minimizes var(S). Setting $\mu = $ max corresponds to the weight μ at which the χ^2_{au} statistic is maximized; the P value is adjusted by use

Table 1

Tests of Association between ADHD Status and DAT-480 or DRD4-7 Alleles

	N^a					
ALLELE	b_a	c_a	b_u	c_u	μ	P
DAT-480	17	10	6	13	.00	.18
					.05	.16
					.10	.14
					.59 (HWSE)	.04
					.59 (max)	.04 $(.09)^b$
					1.00	.11
DRD4-7	15	6	5	10	.00	.05
					.05	.04
					.10	.04
					.44 (max)	.02 $(.05)^b$
					.58 (HWSE)	.02
					1.00	.20

^{a} b_a and c_a = number of transmissions and nontransmissions from heterozygous parents to affected offspring; b_u and c_u = same for unaffected offspring.

^{b} Corrected P value for statistic, maximized over $0 \leqslant \mu \leqslant 1$ by use of Davies's (1977) method.

of the method of Davies (1977), described earlier. Because we are testing candidate loci for linkage or association, a type 1 error rate of $\alpha = .05$ is appropriate. By use of $\alpha = .05$ and Davies's (1977) correction, the DRD4-7 allele is significantly associated with ADHD in our sample. Only when affected individuals are completely excluded ($\mu = 1$) is the test not significant; an unaffected individuals–only test is expected to have low power for many underlying models. The association with the DAT-480 allele is not significant when population prevalence is used as the weight μ, but the Davies P value is small enough (.09) that further study with larger samples is warranted. In both cases, use of the HWSE weight gives an unadjusted P value close to the unadjusted P value for the maximized statistic.

A main interest of investigators is looking at gene-gene interaction. Table 2 shows the results of tests involving the alleles at the DAT and DRD4 loci simultaneously. The first model has main effects only (2-df test) and tests for effects of both genes. Results for five different weights are displayed. The weight .59 was chosen so that comparisons to the tests in table 1 could be made. Given the single-locus results (table 1), it is not surprising that there is some evidence for association. Whether the two-locus model provides more evidence for association than do the single-locus models depends on the weight. At $\mu = .59$, the P value for the main-effects model is lower than that for either of the two single-locus models. This is not the case for either the affected individuals–only ($\mu = 0$) analyses or the population-prevalence ($\mu = .05$ or .10) weights. The second model incorporates an interaction effect for the two alleles as well. These

tests do not show stronger evidence for association than do the main-effects-model tests.

Because the rates of ADHD diagnosis are quite different in males and females, sex may be an important covariate to consider when one is testing for associations. Parental affection status may also influence association. Table 3 is a summary of tests of the alleles with these covariates. In all cases, the HWSE of the nuisance parameter(s) was used. Comparing the P values of tests of the allele alone, using the sample estimates of μ in table 1, with the P values of the tests in table 3, we see that the inclusion of sex or number of affected parents as a main effect in the model does not change the significance level of the tests for DAT-480. Hence, providing different weights based on sex or parental affection status does not substantively change the evidence for this tentative association. The lack of change may be due to the fact that the HWSE weights used for these tests were quite similar: $\mu_m = .5454$ for males and $\mu_f = .6923$ for females, $\mu_0 = .50$ for children with no affected parents and $\mu_1 = .60$ for children with at least one affected parent. There is no evidence of DAT-covariate interactions under the model that we employed: the global 2-df tests that include interaction terms do not have lower P values than do the tests that use models without the interaction terms. However, when the sample is subdivided into males and females, we find that it is only the females who give evidence for an association with DAT-480, thus suggesting a sex–DAT-480 interaction. Since the test restricted to females is based on only 13 transmissions, these results need replication in larger samples. Parental affection status seems to contribute little to the analysis. Offspring in families with an affected parent appear to show somewhat more evidence of association than do offspring in families without an affected parent, but there are only six transmissions scored in families with no affected parents.

Similarly, for DRD4-7, inclusion of either covariate as

Table 2

Joint Tests of Association between ADHD Status and DAT-480 and DRD4-7 Alleles

Model and μ	P
Main effects only (2-df):	
.00	.06
.05	.05
.10	.04
.59	.01
1.00	.12
Main effects and interaction (3-df):	
.00	.07
.05	.06
.10	.05
.59	.01
1.00	.20

Table 3

Tests of Association for DAT-480 and DRD4-7 Alleles, with covariates for ADHD Phenotype and Quantitative Sum of Summary-Scores Pheonotype (QT)

	N^b				COEFFICIENT		ADHD	QT
TERMS IN MODEL[a]	b_a	c_a	b_u	c_u	TESTED	df	P	P
DAT,sex	17	10	6	13	DAT	1	.03	.06
DAT,sex,DAT*sex	17	10	6	13	DAT,DAT*sex	2	.04	.16
DAT (males only)	11	7	6	9	DAT	1	.23	.12
DAT (females only)	6	3	0	4	DAT	1	.03	.26
DAT,paraff	17	10	6	13	DAT	1	.04	.06
DAT,paraff,DAT*paraff	17	10	6	13	DAT,DAT*paraff	2	.11	.13
DAT (paraff=1)	15	9	5	11	DAT	1	.05	.08
DAT (paraff=0)	2	1	1	2	DAT	1	.41	.34
DRD4,sex	15	6	5	10	DRD4	1	.03	.34
DRD4,sex,DRD4*sex	15	6	5	10	DRD4,DRD4*sex	2	.03	.44
DRD4 (males only)	9	3	5	3	DRD4	1	.58	.23
DRD4 (females only)	6	3	0	7	DRD4	1	.01	.68
DRD4,paraff	15	6	5	10	DRD4	1	.08	.46
DRD4,paraff,DRD4*paraff	15	6	5	10	DRD4,DRD4*paraff	2	.16	.47
DRD4 (paraff=1)	10	3	3	2	DRD4	1	.52	.92
DRD4 (paraff=0)	5	3	2	8	DRD4	1	.07	.22

[a] sex = sex of individual, paraff = indicator for more than one parent affected with ADHD.

[b] Variables are as in table 1.

a main effect does not improve the evidence for association by decreasing the P value, and there is no evidence of interaction effects by use of the global 2-df test. For this allele, as for DAT-480, the HWSE weights for males and females were similar: $\mu_m = .6000$ for males and $\mu_f = .5625$ for females. For parental affection status, the weights differed more: $\mu_0 = .4444$ for children with no affected parents and $\mu_1 = .7222$ for children with at least one affected parent. Interestingly, in the subdivided sample, it is again only the 17 transmissions to females that produce evidence for an association with DRD4-7, again suggesting an interaction with sex.

Parental affection status seems to contribute little to the analysis of DRD4-7. There is some evidence for association with DRD4-7 among offspring with no affected parents, but there is none among offspring with affected parents.

Quantitative Phenotype

In addition to DSM-IV ADHD phenotypes, we also have hyperactivity-impulsivity and inattention summary scores, which provide a quantitative phenotype related to the ADHD-affection-status phenotype. To illustrate the use of a quantitative phenotype by our methods, we repeated our analyses, using as the phenotype (Y_{ij}) the sum of the hyperactivity-impulsivity and inattention summary scores. As before, HWSEs of the nuisance parameters were used, and the same test statistics were employed to test for associations. The last column of table 3 is a summary of the results. Interestingly, there is no evidence for a DRD4-7 association with this phe-

notype. Evidence for a DAT-480 association is slightly weaker than that for the binary ADHD phenotype.

Discussion

Our approach shares some features of the Monte Carlo methods proposed by Kaplan et al. (1997) and Martin et al. (1997) and of the score-test methods proposed by Self et al. (1991) and Schaid (1996). Kaplan et al. (1997) and Martin et al. (1997) show how the P value of any of several test statistics can be evaluated by Monte Carlo sampling from the distribution of the transmitted alleles, conditional on parental genotypes at the marker. Self et al. (1991) and Schaid (1996) use conditional-likelihood methods to compute score statistics for standard association models. All of these approaches use only affected sibs with genotyped parents. One distinction between our procedure and that of Schaid (1996) is in our application of the admixture correction. Schaid (1996) applies the correction before computing the score, whereas we do it afterward. Although these two approaches are equivalent and, in specific cases, result in the same tests, choosing to model P(disease|genotype) rather than P(genotype|disease) has several advantages. First, this formulation provides an easy way to include siblings of all phenotypes—both affected and unaffected siblings for binary traits, or any quantitative phenotype—into a test statistic. Second, it allows us to seamlessly include relevant environmental (or genetic) covariates, reducing variability and potentially increasing the efficiency of tests. Third, derivation of the mean and variance of score statistics by use of

the conditional distributions of genotype under H_0 are generally much simpler than derivations of the conditional likelihoods.

Several investigators have recently described family-based tests of association for quantitative traits. Allison's (1997) TDT_{Q5} is an F-ratio test that also allows the incorporation of covariates but that makes assumptions about the trait distribution. Clayton and Jones (1999) describe an extension of the Self et al. (1991) and Schaid (1996) score-test approach for marker haplotypes that uses either discrete or quantitative traits. Their focus is on multiallelic tests, and they do not consider the use of covariates or unaffected individuals in tests of discrete traits. Fulker et al. (1999) extend maximum-likelihood variance-components procedures to allow a test for allelic association, as well as for linkage by use of sib-pair data, allowing the incorporation of covariates and interactions.

We have used standard logistic and linear association models for binary and quantitative traits, but the general approach can be extended to time-until-onset, categorical, ordinal, or multivariate phenotypes. Treating the phenotype as outcome allows for flexibility in modeling, but, because we treat it as fixed in the computation of the distribution of the test statistic, the approach is valid for any type of ascertainment scheme that depends on the phenotypes of any of the family members. Furthermore, because we compute the distribution of the test statistic under the correct conditional distribution of the transmitted alleles, the tests are unbiased even when the associated model or phenotype distribution is misspecified and the population is heterogeneous.

When there is variation in the phenotype Y_{ij}, or when covariates are included in the model, our approach requires the specification of nuisance parameters. Misspecifying the parameters has no effect on the validity of the statistic and, within a reasonable range, should have little effect on the power. For our data example, specifying the nuisance parameter as the proportion of parental transmissions to affected sibs, relative to all transmissions, worked well. This is not surprising, since the HWSE of prevalence minimizes the variance of the test statistic. Intuitively, this estimate seems to be reasonable if all sibs in each family are phenotyped: if a high proportion are affected, then we expect unaffected individuals to carry information about linkage and association. If a low proportion of offspring are affected, then we expect affected individuals to carry the majority of linkage information. For applications with only one nuisance parameter, maximizing the test statistic over the parameter and then adjusting the P value of the test appropriately is also a reasonable choice. Future work will examine the power of this strategy, compared with the use of sample and population estimates. In our data example, the P values derived from the sample-based estimates and the uncorrected P values for the maximized statistic were similar.

We provide an alternative derivation of the T_{sib} statistic derived by Whittaker and Lewis (1998), in which the unaffected individuals are weighted by the population prevalence. T_{sib} was derived as the most powerful test under a multiplicative-genotype relative-risk model. Our derivation can be generalized to other types of phenotypes and association models.

The application to families with ADHD illustrates how the methods may be used to investigate associations by use of data from both affected and unaffected offspring. In this data set, including the unaffected offspring with an appropriate weighting scheme increased the evidence for association. Whether this is true in general, and whether the typing of unaffected individuals is efficient, is likely to depend on the underlying disease model and ascertainment scheme. Whittaker and Lewis (1998) found that unaffected children contribute very little additional information for the simple genetic models and ascertainment scheme that they considered. Whether this is a more general result—for example, for diseases with higher prevalence and alleles with higher frequency or for models that involve a shared environmental and genetic risk beyond the allele being tested—is yet to be determined.

Acknowledgments

References

Allison, DB (1997) Transmission-disequilibrium tests for quantitative traits. Am J Hum Genet 60:676–690

Asghari V, Sanyal S, Buchwaldt S, Paterson A, Jovanovic V, Van Tol HH (1995) Modulation of intracellular cyclic AMP levels by different human dopamine D4 receptor variants. J Neurochem 65:1157–1165

Asherson P, Virdee V, Curran S, Ebersole S, Freeman B, Craig I, Simonoff E, et al (1998) Association of DSM-IV attention deficit hyperactivity disorder and monoamine pathway genes. Am J Med Genet 81:549

Castellanos FX, Lau E, Tayebi N, Lee P, Long RE, Giedd JN, Sharp W, et al (1998) Lack of an association between a dopamine-4 receptor polymorphism and attention-deficit/hyperactivity disorder: genetic and brain morphometric analyses. Mol Psychiatry 3:431–434

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

Comings DE, Comings BG, Muhleman D, Dietz G, Shahbahrami B, Tast D, Knell E, et al (1991) The dopamine D2

receptor locus as a modifying gene in neuropsychiatric disorders. JAMA 266:1793–1800

Comings DE, Gonzalez N, Wu S, Gade R, Muhleman D, Saucier G, Johnson P, et al (1999) Studies of the 48 bp repeat polymorphism of the DRD4 gene in impulsive, compulsive, addictive behaviors: Tourette syndrome, ADHD, pathological gambling, and substance abuse. Am J Med Genet 88: 358–368

Cook EH Jr, Stein MA, Krasowski MD, Cox NJ, Olkon DM, Kieffer JE, Leventhal BL (1995) Association of attention-deficit disorder and the dopamine transporter gene. Am J Hum Genet 56:993–998

Daly G, Hawi Z, Fitzgerald M, Gill M (1998) Attention deficit hyperactivity disorder: association with the dopamine transporter (DAT1) but not with the dopamine D4 receptor (DRD4). Am J Med Genet 81:501

——— (1999) Mapping susceptibility loci in attention deficit hyperactivity disorder: preferential transmission of parental alleles at DAT1, DBH and DRD5 to affected children. Mol Psychiatry 4:192–196

Davies RB (1977) Hypothesis testing when a nuisance parameter is present only under the alternative. Biometrika 64: 247–254

Falk CT, Rubinstein P (1987) Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. Ann Hum Genet 51:227–233

Faraone SV, Biederman J (1998) Neurobiology of attention deficit hyperactivity disorder. Biol Psychiatry 44:951–958

Faraone S, Biederman J, Weiffenbach B, Keith T, Chu M, Weaver A, Spencer T, et al (1999) Dopamine D4 gene 7-repeat allele and attention-deficit hyperactivity disorder. Am J Psychiatry 156:768–770

Fulker DW, Cherny SS, Sham PC, Hewitt JK (1999) Combined linkage and association sib-pair analysis for quantitative traits. Am J Hum Genet 64:259–267

Gill M, Daly G, Heron S, Hawi Z, Fitzgerald M (1997) Confirmation of association between attention deficit hyperactivity disorder and a dopamine transporter polymorphism. Mol Psychiatry 2:311–313

Kaplan NL, Martin ER, Weir BS (1997) Power studies for the transmission/disequilibrium tests with multiple alleles. Am J Hum Genet 60:691–702

LaHoste GJ, Swanson JM, Wigal SB, Glabe C, Wigal T, King N, Kennedy JL (1996) Dopamine D4 receptor gene polymorphism is associated with attention deficit hyperactivity disorder. Mol Psychiatry 1:121–124

Martin ER, Kaplan NL, Weir BS (1997) Tests for linkage and association in nuclear families. Am J Hum Genet 61:439–448

McCullagh P and Nelder JA (1989) Generalized linear models, 2d ed. Chapman & Hall, New York

Poulton K, Holmes J, Hever T, Trumper A, Fitzpatrick H, McGuffin P, Owen M, et al (1998) A molecular genetic study of hyperkinetic disorder/attention deficit hyperactivity disorder. Am J Med Genet 81:458

Rabinowitz D (1997) A transmission disequilibrium test for quantitative trait loci. Hum Hered 47:342–350

Rabinowitz D, Laird NM. Adjusting association tests for population admixture with arbitrary pedigree structure and arbitrary missing marker information. Hum Hered (in press)

Rowe DC, Stever C, Giedinghagen LN, Gard JM, Cleveland HH, Terris ST, Mohr JH, et al (1998) Dopamine DRD4 receptor polymorphism and attention deficit hyperactivity disorder. Mol Psychiatry 3:419–426

Rubinstein P, Walker M, Carpenter C, Carrier C, Krassner J, Falk C, Ginsberg F (1981) Genetics of HLA disease associations: the use of haplotype relative risk (HRR) and the "haplo-delta" (Dh) estimates in juvenile diabetes from the three racial groups. Hum Immunol 3:384

Schaid DJ (1996) General score tests for association of genetic markers with disease using cases and their parents. Genet Epidemiol 13:423–449

Schaid DJ, Sommer SS (1994) Comparison of statistics for candidate-gene association studies using cases and parents. Am J Hum Genet 55:402–409

Self SG, Longton G, Kopecky KJ, Liang K-Y (1991) On estimating HLA/disease association with application to a study of aplastic anemia. Biometrics 47:53–62

Sham PC, Curtis D (1995) An extended transmission/disequilibrium test (TDT) for multi-allele marker loci. Ann Hum Genet 59:323–336

Smalley SL, Bailey JN, Palmer CG, Cantwell DP, McGough JJ, Del'Homme MA, Asarnow JR, et al (1998) Evidence that the dopamine D4 receptor is a susceptibility gene in attention deficit hyperactivity disorder. Mol Psychiatry 3:427–430

Spielman RS, Ewens WJ (1998) A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test. Am J Hum Genet 62:450–458

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

Swanson JM, Sunohara GA, Kennedy JL, Regino R, Fineberg E, Wigal T, Lerner M, et al (1998) Association of the dopamine receptor D4 (DRD4) gene with a refined phenotype of attention deficit hyperactivity disorder (ADHD): a family-based approach. Mol Psychiatry 3:38–41

Terwilliger JD, Ott J (1992) A haplotype-based "haplotype relative risk" approach to detecting allelic associations. Hum Hered 42:337–346

Waldman ID, Rowe DC, Abramowitz A, Kozel ST, Mohr JH, Sherman SL, Cleveland HH, et al (1998) Association and linkage of the dopamine transporter gene and attention-deficit hyperactivity disorder in children: heterogeneity owing to diagnostic subtype and severity. Am J Hum Genet 63: 1767–1776

Whittaker JC Lewis CM (1998) The effect of family structure on linkage tests using allelic association. Am J Hum Genet 63:889–897