10

Family-Based Methods for Linkage and Association Analysis

Nan M. Laird and Christoph Lange
Department of Biostatistics, Harvard School of Public Health,
Boston, Massachusetts 02115

Advances in Genetics, Vol. 60
0065-2660/08 $35.00

DOI: 10.1016/S0065-2660(07)00410-5

ABSTRACT

Traditional epidemiological study concepts such as case-control or cohort designs can be used in the design of genetic association studies, giving them a prominent role in genetic association analysis. A different class of designs based on related individuals, typically families, uses the concept of Mendelian transmission to achieve design-independent randomization, which permits the testing of linkage and association. Family-based designs require specialized analytic methods but they have distinct advantages: They are robust to confounding and variance inflation, which can arise in standard designs in the presence of population substructure; they test for both linkage and association; and they offer a natural solution to the multiple comparison problem. This chapter focuses on family-based designs. We describe some basic study designs as well as general approaches to analysis for qualitative, quantitative, and complex traits. Finally, we review available software.

I. INTRODUCTION

Families have dominated genetic studies, dating back to Mendel's first experiments elucidating the concepts of inheritance in plants. Later, the work of Galton, Fisher, and others on familial aggregation and segregation was built on a wealth of information about inheritance patterns derived from family studies. With the progress of the Human Genome Project, genetic markers spanning the entire human genome have enabled widespread mapping efforts based on linkage analysis using families with multiple affected individuals, leading to the discovery of many genes for Mendelian diseases and traits.

Association analysis differs fundamentally from linkage in that it is not mandatory to use families, and inferences can be made about genetic association from unrelated individuals (see Chapter by Schork *et al.*, this volume). We refer to designs that use unrelated individuals as standard designs (typically case control or cohort). Family designs for studying association based on trios (two parents and an affected offspring) were introduced by Rubenstein *et al.* (1981) and Falk and Rubinstein (1987), while the analysis of such designs was discussed by Spielman *et al.* (1993), Ott (1989), and Terwilliger and Ott (1992).

A. Hypothesis testing in family designs

In testing for genetic association with a standard design, the null and alternative hypotheses are simply given as:

H0: no association between the marker and the disease
HA: association is present between the marker and the disease

A rejection simply implies that the disease trait of interest is associated with the alleles at the marker. With a family-based test (FBAT), the null and alternative hypotheses can be phrased in terms of the underlying genetics in the population. As noted in Ott (1989), family designs have no power to detect association unless linkage is present. Thus, when testing for association or linkage with family designs, the alternative hypothesis is always HA: Both linkage and association are present between the marker and a disease susceptibility locus (DSL) underlying the trait.
There are three possibilities for the null hypothesis in a family design:

(1) H0: no linkage and no association between the marker and any DSL underlying the trait.
(2) H0: linkage but no association between the marker and any DSL underlying the trait.
(3) H0: association but no linkage between the marker and any DSL underlying the trait.

When testing candidate genes, or in a whole genome scan, the appropriate null hypothesis will ordinarily be the first null, H0: no linkage and no association. However, if we are testing for association in a study that has known linkage in the region of testing, then a more appropriate null hypothesis is the second, H0: linkage but no association. This distinction is not relevant when our sample consists of parents and one offspring. However, when the sample includes multiple offspring from the same family, with or without parents, the distribution of the test statistic under the null differs, depending on whether linkage is assumed to be present. In Section III, we will show how to construct valid tests under both types of null hypotheses.

The third null hypothesis, H0: association but no linkage was proposed by Spielman *et al.* (1993), when they introduced their transmission disequilibrium test (TDT) to test for linkage in a setting where association had been demonstrated in several population studies, but conventional linkage analysis failed to find evidence of linkage. However, as the distribution of family data under the null hypothesis is the same for the first and third null hypotheses (i.e., the null distribution depends only on whether linkage and multiple sibs are present), we usually consider only the first and second nulls. We remark that some authors (Guo *et al.*, 2007) seem to prefer the null hypothesis

(4) H0: no linkage or no association between the marker and any DSL underlying the trait.

Because the distribution under the null must consider the possibility of linkage without association, the distribution of the data under the null is the same as that for the second null hypothesis, and the two are thus equivalent, in the sense that any test valid for the second is valid for the fourth.

B. The TDT test for trios

The basic family design is the trio, consisting of two parents and one offspring. The TDT test (Spielman *et al.*, 1993) is the standard approach to the analysis when the offspring is affected with the trait of interest. The analysis is similar in principle to the alleles test in a case-control analysis, in that the number of A alleles, for example, among the cases is compared to the number expected under the null hypothesis. The main difference is how the number expected is computed under the null. With a case-control (or more generally, a standard) design, the number expected is computed by assuming the distribution of alleles is the same in cases and controls under the null, and by using that common distribution to derive an expectation for the affected group. In contrast, the trio design, and family-based designs in general, relies on using Mendel's laws to compute expectations for the offspring based on their parent's genotypes.

The basic design is diagrammed in Fig. 10.1. The analysis is very intuitive. If any of the three null hypotheses mentioned above holds, then Mendel's laws dictate the transmission of alleles from parents to offspring. The mother can only transmit the A allele and is thus not informative about association of any allele with disease in the offspring. The father transmits either A or B with probability 50/50. Thus, the child is either AA or AB, with probability 50/50. The TDT test consists of using the A alleles transmitted from heterozygous parents to their offspring (n_A). Under any H0, n_A follows a binomial distribution with $p = 0.5$ and n = the number of heterozygous parents. Because parents' transmissions are independent, each heterozygote parent has probability 0.5 of transmitting

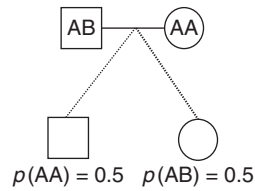


Figure 10.1. Trio Design—the TDT. Family-trios are the basis of the transmission disequilibrium test (TDT); it compares the observed number of the A alleles transmitted to the affected offspring with those expected by Mendelian transmissions. An excess of A (or B) alleles among the affected suggests that a DSL for the trait is in linkage and linkage disequilibrium (LD) with the marker locus.

the A allele. Thus, one can compute an exact test of the null hypothesis, or an asymptotic Z or χ^2 test based on this binomial distribution. The TDT usually refers to the χ^2 version of the binomial test.

As in the alleles test commonly used in standard designs (Chapter by Schork *et al.*, this volume), the potential sample size is twice the number of trios because each individual has two alleles. With the TDT test, however, that advantage is offset, as the transmissions from homozygous parents are not used. Thus, the effective sample size may be considerably less than the number of trios, depending on allele frequency. If there are multiple affected offspring, then the same test remains valid (counting n and n_A as the total number of heterozygous parent transmissions to all offspring and the number of A allele transmissions from heterozygous parents to all offspring, respectively), provided the null hypothesis tested assumes no linkage because parental transmissions to different offspring remain independent when there is no linkage between the marker and any DSL affecting the trait.

The derivation of the TDT leads to an intuitive justification for the premise that both linkage and association must be present under the alternative. If there is association, but no linkage, between the marker and the DSL in the parent population, then the marker alleles in the parents are transmitted independently of the DSL alleles, and there will be no association between the marker and any DSL in the offspring. If there is linkage of the two loci in the parents, but not association, then the two loci will be linked in the offspring, but different marker alleles will be transmitted with different DSL alleles in different families, so there will be no "population" association in the offspring. Formally, Vansteelandt *et al.* (2007) have shown that conditioning on the parents' genotypes serves to eliminate any potential confounding in the test of association, making it robust not only to population stratification and admixture but also to potential model misspecification.

C. Extensions to the TDT

Because of its great success in the analysis of trio data, there is a wealth of literature on extensions of the basic TDT. Curtis and Sham (1995a), Bickeboller and Clerget-Darpoux (1995), and Spielman and Ewens (1996) describe extensions for multiallelic tests. Schaid (1996) put the TDT test into a more general context of a score test for multinomial data, showing that the TDT is optimal for an additive alternative, and providing tests for dominant and recessive models as well. Spielman and Ewens (1998), Curtis and Sham (1995b), Schaid and Li (1997), Rabinowitz and Laird (2000), and Fulker *et al.* (1999) discuss family tests when parents are missing and/or for general pedigree designs. Martin *et al.* (2000), Horvath and Laird (1998), and Lake *et al.* (2000) describe methods for general pedigrees that are also valid when testing for association in the presence of linkage.

Fulker *et al.* (1999), Abecasis *et al.* (2000), Rabinowitz (1997), Horvath *et al.* (2001), and Laird *et al.* (2000) discuss extensions for quantitative traits. An overview of analysis methods for family designs is given by Zhao (2000). We will consider many of these extensions in Sections III and IV.

D. Design issues

There are many possible family configurations that can be used in family designs: trios, sib pairs, general nuclear families (with or without parents), and more general pedigrees. The trio design is generally the most powerful among family designs with one affected offspring per family. Although many methods have been proposed for using incomplete trios with only one parent, such methods can be biased (Curtis and Sham, 1995b) with biallelic markers, and generally it will be necessary to have at least one additional offspring to capture information from the family. Figure 10.2 shows some power comparisons for four designs: the case-control, the trio, discordant sib pairs (DSPs) (without parents), and discordant sibships (no parents, one unaffected offspring and two unaffected siblings). All four of these designs have the same number of affected cases. Panel (A) shows power for a rare disease (prevalence 0.1%) and panel (B) shows power for a common disease (14%). Both cases assume an additive disease model with allelic odds ratio of 1.3.

With rare disease, the trio design, followed by the case-control, is the most powerful relative to the number of affected offspring that need to be recruited. However, more genotyping is required (three genotypes per case, as opposed to two per case in the case-control design). In addition, it can be difficult to recruit parents; notable exceptions are childhood illnesses and when using samples originally designed for linkage analysis. With more common diseases, the case-control design is more powerful, followed closely by the parent–offspring trio and the discordant sibship trio. At both levels of prevalence, the DSP design is considerably less powerful than either the trio or the case-control (Witte *et al.*, 1999), although it requires less genotyping than either of the other family designs.

We note that unaffected siblings are most commonly used in family-based designs to compensate for missing parents. However, even when parents are present, information can be gained about association by using transmissions to unaffected offspring in the case of common disorders (Lange and Laird, 2002a; Whittaker and Lewis, 1998).

Figure 10.3A shows how the power of the basic TDT can be increased (or decreased) by using information from an additional unaffected offspring when disease prevalence is 0.3. Here the dotted horizontal line indicates power for the TDT that discards the unaffected offspring. The unaffected offspring are included by using an offset μ_Y (see Section III); when the offset is zero, unaffected offspring are not included. When the offset is close to the prevalence, the

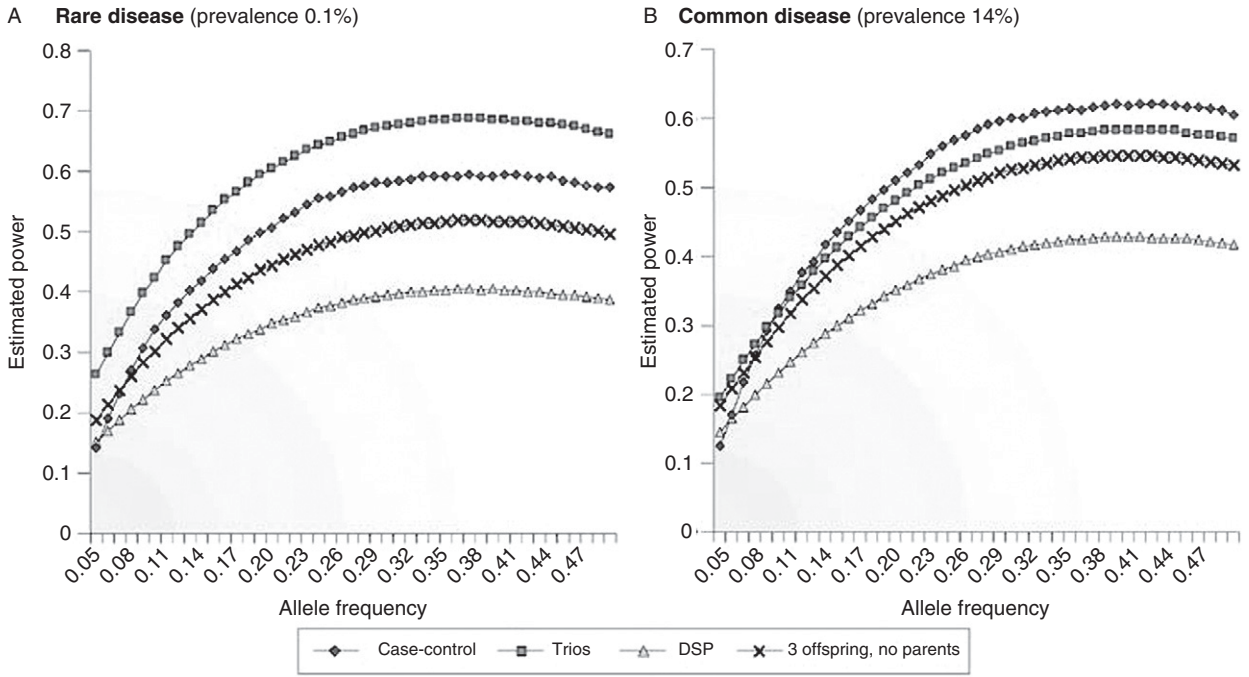


Figure 10.2. Power comparison between case-control studies and family-based designs. The estimated power levels for a case-control study with 200 cases and 200 controls are compared with those for various family-based designs: 200 trios (of an affected offspring plus parents), 200 discordant sibling (sib) pairs (DSPs; one affected and one unaffected) without parents, 200 "three discordant offspring (at least 1 affected and 1 unaffected) and no parents." Discordant sib pair designs have 50% less power than case-control designs, (Witte *et al.*, 1999). For the rare diseases (A), trio designs are more powerful than case-control designs. For common diseases (B), case-control designs are slightly more powerful than trio designs and designs with 3 discordant sibs. Although it is not shown here, for larger-effect sizes (e.g., odd ratios greater than 2), unaffected probands contain more information and the DSP design can achieve power levels that are similar to those of trios designs. The power calculations for both the family designs and the case-control designs were done in PBAT (v3.3) using Monte-Carlo simulations. These figures are reprinted from Laird and Lange (2006).

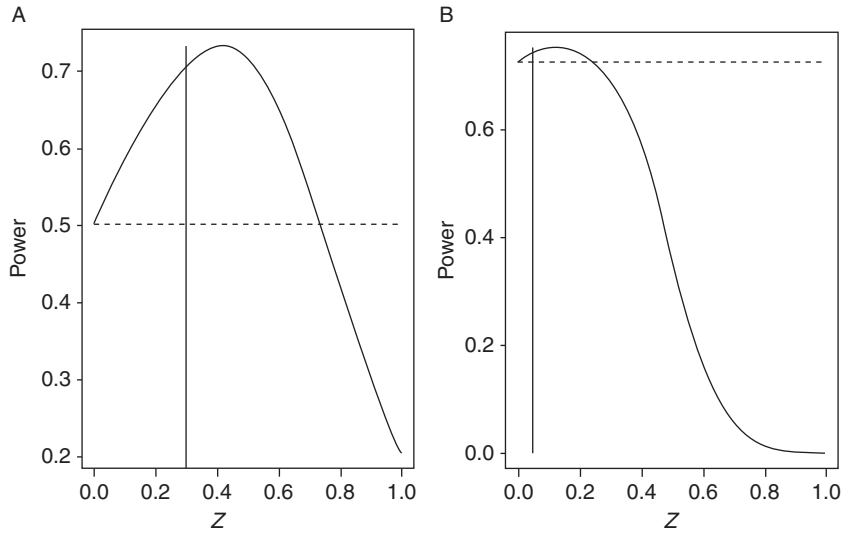


Figure 10.3. (A, B) Power increase by including unaffected offspring/dependence of the power on the offset choice. For a common and for a rare diseases, respectively, (A) and (B) show the increase in power for a sample size of 100 families if unaffected offspring are included in the FBAT statistic. The power is shown as a function of the offset μ. The solid line shows the power of the FBAT statistic and the dotted line the power of the TDT using only affected probands. The vertical line shows the prevalence of the disease. For offset $\mu = 0$, the TDT and the FBAT are identical. For a common disease including unaffected probands in the FBAT, the power increases substantially, while for rare diseases the increases is negligible. For both, the common and the rare disease scenario, the power reaches its maximum when the offset choice is approximately the disease prevalence. These figures are reprinted from Lange and Laird (2002a).

power is maximized; but if the offset is too large, too much weight is given to unaffected offspring and power is lost relative to the TDT. With rare disease, there is little to be gained from using an offset, as the maximum power is only slightly above the TDT, but again, using an offset that is too large can have negative consequences (Fig. 10.3B).

Table 10.1 shows some design and power considerations for binary traits that depend upon ascertainment conditions and family design. These power considerations assume that the optimal offset is used (see Section III) and that there is no environmental correlation between sibling phenotypes. As such, they may be slightly optimistic for designs with multiple affected offspring. For 200 families, we consider a common and a rare disease and two ascertainment conditions.

Table 10.1. Design and Power Considerations for Dichotomous Traits

No. of genotypes per family	Family type: No. of offspring	No. of parents	K = 0.3, MAF = 0.2		K = 0.05, MAF = 0.05	
			A	B	A	B
2	2	0	0.40	0.55	0.46	0.48
3	1	2	0.52	–	0.73	–
3	3	0	0.65	0.68	0.61	0.61
4	2	2	0.69	0.59	0.76	0.73
4	4	0	0.79	0.79	0.70	0.70
5	5	0	0.92	0.85	0.83	0.83

K: disease prevalence, A/B: Ascertainment condition, MAF: minor allele frequency of DSL.

Ascertainment condition A requires at least one affected offspring per family, while ascertainment condition B requires at least one affected and one unaffected offspring per family. We assume the disease prevalence is used for the offset.

For a prevalence of 30%, DSPs without parents are as powerful as trios consisting of one affected proband and parents. When the parental genotypes are missing, discordance-ascertainment conditions can more than compensate for the power loss caused by the missing parental information. When two or more additional offspring are available, there is little effect of the ascertainment condition on power, except that if parents are available, it is advantageous to have more affected offspring, making ascertainment condition A preferable.

For rare disease/rare variant, if parents are missing, it is necessary to genotype more individuals per family to attain comparable power to those cases without missing parents. As a rule of thumb, three additional siblings compensate for the power loss caused by two missing parents. Here there is little effect of ascertainment scheme on power because with a rare disease, most siblings will be unaffected.

The situation with quantitative phenotypes is somewhat different. Although it is certainly possible to ascertain individuals into a study based on their level of a quantitative trait (Risch and Zhang, 1995), such designs are difficult to implement. More likely, individuals are ascertained according to a qualitative trait, and quantitative phenotypes are also measured, for example, asthma and FEV1, or obesity and BMI. With the availability of large cohort studies with family data, we can have family designs in which there is no ascertainment with respect to trait of interest. This can be a significant advantage for the analysis of quantitative traits, although population-based samples

will generally not be very useful for the analysis of rare qualitative traits. Ascertainment of subjects relative to the phenotype of interest is important because it can dictate how the analysis should be carried out (see Section III), and the power can depend quite heavily on the combination of ascertainment conditions and analytic method.

Figure 10.4 illustrates the effect on power of ascertainment conditions and analysis choices when dealing with a quantitative trait. The figure compares two strategies: random sampling from the population and selection of only those subjects whose trait is in the top 10% (considered affected). The analytic choices are to use a TDT with only affected offspring or to use the quantitative trait in the analysis (see Section III). When there is no ascertainment condition relative to the trait, then it is far preferable to analyze the quantitative trait, with an offset close to the population mean, which can be well estimated by the sample mean in this setting. With ascertainment of affected offspring only, using the basic TDT on the affected is always the best strategy; considerable power may be lost by analyzing the quantitative trait, unless the offset is carefully selected. Using the sample mean as an offset gives poor results because the sample mean is a biased estimate of the population mean.

II. ANALYSIS METHODS: FBAT AND PBAT

Here we discuss a very general approach to the analysis of family-based data. This approach permits any type of genetic model, multiallelic data, general family design, different null hypotheses, any phenotypic trait (binary, time-to-onset, measured, repeated measures, multivariate), haplotypes, and multiple markers. To motivate this approach, it is important to consider those aspects of the TDT that make it so robust and powerful. First, the test statistic is computed conditionally on the observed parental genotype. This conditioning serves to eliminate any assumptions about the distribution of alleles in the population, such as Hardy-Weinberg, or that allele frequencies are the same in cases and controls under the null. Second, the test statistic is computed conditional on the trait; this serves to eliminate assumptions about the distribution of the trait in the population. Finally, the random variable is the offspring genotype. Its distribution under the null is computed using Mendel's first law—thus the validity of the test statistic relies only on Mendel's law of random transmission of each parental allele with equal probability to each offspring.

The FBAT approach (Laird *et al.*, 2000) to the analysis of family data uses these same underlying principles in constructing a test statistic that generalizes the TDT to more complex situations. The general idea is the same: We condition on the traits (which can include any type and number of traits) and on parental genotypes (which can include multiple markers and haplotypes),

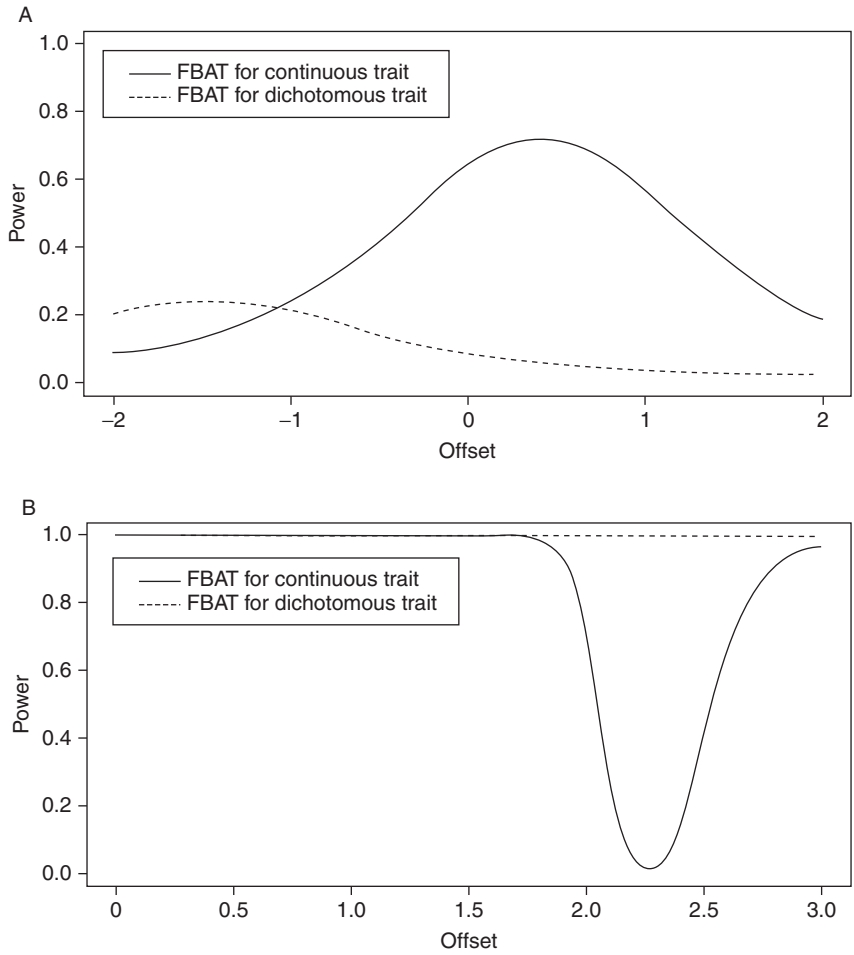


Figure 10.4. (A, B) Asymptotic power calculations for a continuous trait. The disease-allele frequency is 0.3 and the heritability, h^2, is 0.1. The dotted line shows the power of the dichotomous *FBAT* for offset choices between 0 and 1. Significance level $\alpha = 0.01$; (A) Additive model—total population sample with mean $\overline{y} = 0.39$, maximal power of FBAT 0.75, power of FBAT-o 0.74, power of PDT 0.73, and power of QTDT 0.74 ($n = 200$); (B) Additive model—affected sample with phenotypic mean $\overline{y} = 2.2$, maximal power of *FBAT* 1.00, power of FBAT-o 0.04, power of PDT 0.034, and power of QTDT 0.01 ($n = 200$). These figures are reproduced from Lange *et al.* (2002).

and we compute the distribution of the test statistic from the distribution of offspring genotypes under the null. When parents' genotypes are missing, we condition, instead, on the sufficient statistic for parental genotype (denoted S_i); see subsection on *Missing Parents or Founders* below. When analyzing haplotypes, we condition on the sufficient statistic for missing phase as well.

A. General test statistic

The test statistic uses a natural measure of association between two variables, a covariance between the traits and the genotypes. We define the covariance as

$$U = \Sigma T_{ij}(X_{ij} - E(X_{ij}|S_i)), \tag{1}$$

where i indexes family and j indexes nonfounders in the family, and summation is over all i and j. Here, T_{ij} is a coding function for the trait of interest and X_{ij} is a coding function for the genotype. The usual sample covariance centers both variables around their sample means, but with the FBAT statistic, X_{ij} is centered around its expected value, $E(X_{ij}|S_i)$, conditional on the sufficient statistic for the parental genotype, and computed under Mendel's laws. As we discuss below, T_{ij} is typically a centered phenotype. The coding function for the trait allows us to incorporate both qualitative and quantitative phenotypes, as well as time to onset. This basic formula can be used in virtually every setting; the key is the definition of the coded traits and the coded genotypes and how the distribution is computed under the null.

Note that the "centered genotype" $(X_{ij} - E(X_{ij}|S_i))$ can be thought of as the residual of the "transmission" of parental genotype to offspring. For any coded genotype, $(X_{ij} - E(X_{ij}|S_i)) = 0$ if both the parents of the ijth offspring are homozygous, regardless of what particular genotypes the parents have—that is, transmissions from homozygous parents do not contribute to the test statistic. With one homozygous parent, if we define X_{ij} as the number of A alleles, then $(X_{ij} - E(X_{ij}|S_i)) = 1/2$ if A is transmitted and $-1/2$ if the A is not transmitted (because again, transmissions from the other parent, who is homozygous, do not count). Finally, with two heterozygous parents, $(X_{ij} - E(X_{ij}|S_i))$ equals 1, 0, or -1, depending upon the number of A alleles transmitted (2, 1, or 0). Thus, in the special case where $T_{ij} = 1$ for all i and j (see below), U simply counts the total number of A transmissions from heterozygous parents, minus their expected number $(n_A - n/2)$, in the notation of Section I.B.

B. Coding the genotype

The coded genotype is chosen to reflect the selected mode of inheritance. For example, for the additive model, X_{ij} counts the number of A alleles; for the recessive, X_{ij} is 1 if the ijth offspring's genotype is AA and 0 otherwise.

In a multiallelic setting with p alleles, X_{ij} is a p-dimensional vector, each element of the vector coding for a different allele. In this case, U will also be a p-dimensional vector. Other specifications for X_{ij} will be discussed under multiple markers and haplotypes.

C. Coding the trait: Dichotomous outcomes

Consider first the case where the phenotype of interest is affection status. Setting $T_{ij} = 1$ if affected and $T_{ij} = 0$ otherwise means the test statistic will not incorporate information about transmissions to unaffected offspring. Note that this is equivalent to including only affected individuals in the test statistic. To incorporate unaffected individuals into the test statistic, we use an offset, letting $T_{ij} = (Y_{ij} - \mu)$, where Y_{ij} is the original 1/0 phenotype and μ is a user-defined offset parameter. Thus, for a dichotomous trait, with $Y_{ij} = 1$ if affected and 0 otherwise, $T_{ij} = 1 - \mu$ for affected individuals and $-\mu$ for unaffected. The optimal offset (Lange and Laird, 2002a; Whittaker and Lewis, 1998) is approximately the prevalence of the disorder, $\mu = E(Y)$. Note that the U statistic can now be thought of as a contrast between transmissions to affected offspring weighted by $(1 - \mu)$ and unaffected offspring, weighted by μ. A contrast is used because an overtransmission of the A allele to affected offspring should correspond to an undertransmission to the unaffected. As noted in Section I.D however, assigning too much weight to the unaffected (μ much larger than the prevalence) will result in a loss of power relative to using affected offspring only. When $\mu = 0$, only affected individuals are included.

Of course, a general problem is lack of knowledge about population prevalence. In addition, with dichotomous traits, ascertainment on the trait usually means that it is not possible to estimate μ from the data in hand. Fortunately, power of the test is reasonably good in a neighborhood around μ (Lange and Laird, 2002a). An alternative approach to choosing the offset is to choose μ to minimize var(U). This gives an easily computed offset (Lunetta *et al.*, 2000) that is close to the sample prevalence. However, the offset can be very large if a large number of unaffected individuals are included in the sample. Because minimizing the variance does not maximize power, we suggest limiting the offset size to a maximum of 0.5.

Quantitative and time-to-onset traits, as well as adjustment for covariates, will be discussed in Section II.H.

D. The test statistic: Large sample distribution under the null

The distribution of the FBAT statistic under the null hypothesis is obtained by treating the X_{ij} as random, but conditioning on the trait, T_{ij}, and the sufficient statistic. As $E(U) = 0$ by construction under H_0, it remains to normalize U by

its standard deviation, again computed under the conditional distribution of offspring genotype, given offspring trait and S_i. For univariate X or T

$$Z = \frac{U}{\sqrt{\text{var}(U)}}, \text{ or equivalently, } \chi^2_{\text{FBAT}} = \frac{U^2}{\text{var}(U)},$$

where

$$\text{var}(U) = \sum_i \sum_{jj'} T_{ij} T_{ij'} \text{cov}(X_{ij}, X_{ij'}, S_i, T_{ij}, T_{ij'}) \tag{2}$$

and $\text{cov}(X_{ij}, X_{ij'} S_i, T_{ij}, T_{ij'})$ is computed conditional on the traits and the sufficient statistics, assuming the null hypothesis is true. Note that this covariance only depends on S_i and not the traits when no linkage is part of the null hypothesis. For testing no association in the presence of linkage, an empirical variance can be used to estimate var(U) (Lake *et al.*, 2000). For large samples, Z is approximately distributed as $N(0,1)$, and χ^2_{FBAT} is distributed as approximately χ^2 on one degree of freedom. In the setting where U is a vector because of either multiple alleles or multiple traits, var(U) is a variance/covariance matrix, and the test statistic is the quadratic form $U^T \text{var}(U)^- U$, which is distributed as χ^2 with degrees of freedom equal to the rank of var(U) (Laird *et al.*, 2000; Lange and Laird, 2002b).

E. The TDT and χ^2_{FBAT}

When we include only affected offspring ($T_{ij} = 1$), and all parents are known, then as previously noted, $U = (n_A - n/2)$, where n_A is the number of heterozygous transmissions of A to all affected children, and n is the number of heterozygous parent–child pairs. Under a null hypothesis that includes no linkage, multiple offspring are independent, and var(U) reduces to

$$\text{var}(U) = \Sigma ij \text{ var}(X_{ij}|S_i) = n\left(\frac{1}{2}\right)^2$$

because each transmission has variance equal to $(1/2)^2$. As a result

$$\chi^2_{\text{FBAT}} = \frac{(n_A - n/2)^2}{n(1/2)^2} = \frac{(n_A - n_B)^2}{n}$$

where n_B is the number of transmissions of the B allele from heterozygous parents to affected offspring. Thus, χ^2_{FBAT} is identical to the TDT when (1) only affected

offspring are included, (2) parents' genotypes are known, and (3) the null hypothesis assumes no linkage. In other cases, χ^2_{FBAT} can be viewed as generalizing the TDT.

F. Computing the distribution with general pedigrees and/or missing founders

As defined above, the test statistic is very general. It applies to trios, parents with multiple offspring, families without parents, and general pedigrees, with or without founder genotypes. The summation over i denotes the independent pedigrees or families, and the summation over j denotes summation over all offspring in the pedigree. However, exactly how the distribution of each X_{ij} is computed depends upon the family structure and whether founder genotypes are known. For trios, it is straightforward to compute the distribution of X_{ij} given parents with known genotype, using Mendel's first law. When there are multiple offspring and no linkage, transmissions to all offspring in the family are independent, and one can treat each offspring as if it comes from a separate family. When linkage is present, the covariance between multiple offspring in the same family depends upon the unknown recombination fraction. To remove dependence of the joint distribution on the unknown recombination fraction, we can condition the distribution on patterns of identity by descent observed among the offspring (Rabinowitz and Laird, 2000). This approach to computing the conditional distribution of X_{ij} leads to discarding many families as noninformative, especially when parental genotypes are unknown. Thus, we generally use an empirical variance, as described above.

These basic ideas extend easily to general pedigrees, where the genotypes of all founders are known, and instead of conditioning on parents, we condition on the founders of the pedigrees. There is potential for a considerable gain in power in this setting when we analyze pedigrees, rather than treating all families within the pedigree as separate families (Laird and Lange, 2006; Rabinowitz and Laird, 2000).

When parents or founder's genotypes are unknown, the situation is slightly more complex, but the joint distribution of offspring outcomes can be calculated using the conditioning algorithm described in Rabinowitz and Laird (2000). The basic idea of this algorithm is to condition the distribution of observed offspring genotypes on the sufficient statistics for the unobserved parental genotypes. In this way, the distribution will not require making any assumptions about the distribution of the unobserved parental genotypes, and robustness to population substructure is maintained. To give an example, suppose we have a family with two offspring and no parents. The conditional distribution of the two offspring genotypes depends upon what genotypes are observed in the two offspring. If we observe that both genotypes are AA [or BB], then nothing

can be inferred about the parents except that each has an A [or a B]. The probability of other possible outcomes for the two offspring will depend upon the unknown parental alleles. Thus, such pairs of offspring are not informative.

If instead, we observe one AA and one BB offspring, then we know that the parents are both AB. To condition on the sufficient statistic for missing parents, we require that all possible outcomes contain one AA and one BB offspring. Any other possible outcome—for example, (AA,AB)—would not allow us to infer that the parents are both AB. Thus the two possible outcomes conditioning on the sufficient statistics are (AA,BB) and (BB,AA), assuming order matters. For an additive coding with X denoting the number of A alleles, it is straightforward to show that $E(X_{i1}|S_i) = E(X_{i2}|S_i) = 1$, $\text{var}(X_{i1}|S_i) = \text{var}(X_{i2}|S_i) = 1$ and $\text{cov}(X_{i1}, X_{i2}) = -1$. Now consider this family's contribution to U and $\text{var}(U)$. Using Eqs. (1) and (2) above, we can show that the contribution to U is $(T_{i1} - T_{i2})$ and the contribution to $\text{var}(U)$ is $(T_{i1} - T_{i2})^2$, assuming the first child is AA and the second is BB. Thus, families in which the trait is constant (i.e., both affected or both unaffected) will make no contribution to the test statistic, but DSPs will be informative.

Finally, if we observe (AA,AB), we know that one parent must be AB and the other must have an A, but otherwise we cannot distinguish between [AA,AB] parents and [AB,AB] parents. Because transmission probabilities differ for the two possible parents, we must therefore keep fixed the set of observed genotypes, and permute them between the two offspring, as above. Again, only DSPs will be informative.

The conditioning algorithm determines the joint distribution of all offspring genotypes in a family or a pedigree; hence, it is also possible to carry out exact tests of H0 by computing the exact probability that the random χ^2_{FBAT} test statistic exceeds the observed statistic under the null (Schneiter *et al.*, 2005). The exact p value can also be estimated via Monte Carlo by drawing from the conditional distribution of offspring genotypes.

G. Haplotypes and multiple markers

A common scenario in association studies uses multiple, closely spaced markers, usually single nucleotide polymorphisms (SNPs), often within the same gene. Here, the null hypothesis of interest may be whether any marker is associated with a DSL underlying the trait. Testing each marker separately and then using FDR or Bonferonni to adjust the p values for multiple testing is one strategy, but may be quite inefficient when SNPs are in high linkage disequilibrium with each other (Chapter 10). Two approaches we discuss here use haplotypes (Chapter by Liu *et al.*, this volume) or multimarker tests.

A multilocus haplotype refers to the set of alleles, one from each marker, that are inherited from a single parent, either from the mother or from the father. There are several circumstances when using haplotypes for testing may be preferable to single-marker testing. If a disease locus is present in the region spanned by the markers, but does not correspond exactly to any of the markers tested, the DSL may not be in sufficiently high disequilibrium with any one marker to be detected by one-at-a-time testing of the markers. If we use enough markers to capture the haplotype diversity in the population, then the haplotypes should capture the variation directly at the disease locus. Alternately, if the DSL is a series of changes in base pairs at two or more of the observed markers, then using haplotypes should again be more powerful than one-at-a-time testing. However, if a single marker corresponds to only DSL in the region, or if many markers are included in the haplotype construction that are not in high LD with the putative DSL, then using haplotypes can be a poor testing strategy.

If each person's haplotype is observed, then the set of markers forming the haplotypes can be considered as a single marker with many alleles, and methods used for testing association with multiallelic markers apply. Although more can be inferred about haplotypes in the family-based setting than in population-based studies because of knowledge of parents or sibling genotypes, phase (i.e., which parent transmitted which allele) cannot always be resolved even in families, especially if parents' genotypes are missing. The principle of conditioning on the sufficient statistics for missing parental genotypes extends quite straightforwardly to handle missing parental phase (Horvath *et al.*, 2004). One now obtains a distribution for the phased offspring genotype, conditioning on the sufficient statistic for both parental genotypes and possibly missing phase. The FBAT test statistic can be computed in the same way, recognizing that the set of markers forming the haplotypes is treated as one multiallelic marker with each haplotype forming an allele. In principle, with n SNPs, there can be 2^n haplotypes, but in practice the number of haplotypes observed in a family-based analysis is usually quite a bit less than 2^n. The availability of family data enables one to eliminate many possibilities as not compatible with observed family data.

The haplotype analysis implemented in FBAT uses the principle of conditioning on the sufficient statistics for both phase and any missing parental data, and on offspring traits, and hence is not biased by population stratification and/or admixture. Either biallelic or multiallelic tests are computed in the usual way, and the empirical variance option can be used to account for the presence of linkage. Any trait can be used. A feature of the implementation of the conditioning algorithm is that it also allows one to recover information from only partially phase known families by using weights (Horvath *et al.*, 2004).

As the number of SNPs increases substantially (more than 5–10), the attractiveness of a haplotype analysis diminishes because of increasing difficulty in resolving phase. In addition, if founder information is missing, it may take

considerable computer time to determine the conditional distribution. In such cases, an alternative approach uses multimarker tests. The basic idea of a multimarker approach is very straightforward; in the context of a case-control study, it is similar to a Hotelling's T^2 test where the vector of marker values is tested for equality in the cases and controls. In the family context, multimarker tests are constructed by letting X_{ij} be a vector, where each element of the vector corresponds to a different coded marker value. The full-joint distribution of the different elements of X_{ij} would require knowledge of the haplotype distribution, as discussed above. To circumvent conditioning on unknown phase, we instead use an empirical estimator of var(X_{ij}) in calculating var(U). The resulting χ^2_{FBAT} has the same quadratic form as in the multiallelic setting, with degrees of freedom again equal to the rank of var(U). The rank will be generally equal to the number of markers, unless two or more of the markers are in near perfect LD (Rakovski *et al.*, 2007). Another approach to the multimarker testing with families is discussed by Xu *et al.* (2006).

H. Coding the trait for complex phenotypes: Age-to-onset phenotypes, quantitative outcomes, and FBAT-GEE

In principle, T_{ij} can be any function of the phenotype and other individual characteristics, as long as the trait does not depend on the genotype being tested. When the phenotype of interest is time to onset, a strategy similar to the log-rank test may be used, letting T_{ij} be log-rank residuals, computed at each failure time (Lange *et al.*, 2004a).

For quantitative phenotypes, typically, a phenotypic residual is used for the coded trait, that is $T_{ij} = (Y_{ij} - \mu)$, where Y_{ij} is the original phenotype and μ is a user-defined offset parameter. Unless subjects have been ascertained into the study on the basis of their quantitative trait, the optimal choice is the phenotypic sample mean. In such situations, the FBAT statistic for quantitative phenotypes has higher statistical power than for dichotomous traits (Lange *et al.*, 2002). However, to benefit from the advantages of quantitative traits in the analysis, a few hurdles have to be overcome. Quantitative traits such as BMI or lung volume also depend on many other nongenetic factors. These can be proband characteristics but also include environmental influences. For example, BMI will depend on age and gender as well as on smoking history and dietary habits. The unadjusted, raw measurements of such traits will not accurately reflect affection status (e.g., obesity) or the severity of the disease. In such situations, it is preferable to adjust the raw measurements for all known covariates because this decreases the outcome variability, and to compute the FBAT statistic based on the phenotypic residuals.

For lung-function phenotypes, such as FEV, standard adjustment formulas have been derived (Ware and Weiss, 1996). However, such adjustment formulas typically are based on unaffected probands and, consequently, do not

always incorporate all disease-specific confounding variables. For example, for FEV1, the standard adjustment includes gender and height but not smoking history, which is an important factor when looking at diseases such as chronic obstructive pulmonary disease. A detailed discussion of the limitations is given by Naylor *et al.* (2005).

In such situations, an alternative approach is to adjust the raw phenotypic measurements, by regressing the phenotype on such confounding variables, and use the phenotypic residuals in the analysis. Such an adjustment will be study specific, requires statistical model building, and might not be reproducible in other studies that may not have recorded the same confounding variables. The goal of such a within-study adjustment is to measure and incorporate all environmental factors and other covariates into the analysis. However, for many phenotypes, the confounding variables are not necessarily known prior to the study design or can be difficult to measure and model.

A second analysis issue for quantitative traits is the multiple testing problem. While affection status is defined by only one variable, the disease and its severity are often described and characterized by a set of quantitative phenotypes. Such quantitative phenotypes typically cluster together into symptom groups. For example, in asthma studies, multiple quantitative phenotypes are recorded that describe lung function (FEV, FVC) (DeMeo and Silverman, 2003). For most complex diseases, such symptom groups of quantitative phenotypes can be defined based on clinical knowledge, knowledge about the underlying biological processes, or simply based on phenotypic correlation. In such situations, it is not desirable to test all quantitative phenotypes individually and adjust for multiple testing. There are two reasons for this: first, the association tests for one symptom group will be correlated, and adjustments for multiple testing tend to be conservative in such situations. Second, if the assumption holds that the phenotypes in a symptom group are influenced by the same pathway or share similar environmental confounding, it will have more power to look at the entire symptom group and test all phenotypes jointly in a single multivariate test, without having to adjust for multiple comparisons.

In the FBAT context, such a multivariate test was introduced by Lange *et al.* (2003b), the so-called FBAT-GEE statistic. Like the original FBAT statistic, FBAT-GEE is easy to compute, tests all phenotypes simultaneously, and does not need distributional assumptions about the phenotypes that will be tested, even if the tested phenotypes are of different variable types (e.g., normally distributed phenotypes, count variables). Assuming that m traits have been recorded for each proband and they form a symptom group that we want to test simultaneously, we denote the vector containing all m observations for each proband by $Y_{ij} = (Y_{ij1}, \ldots, Y_{ijm})$ where Y_{ijk} is the kth phenotype for the jth offspring in the ith family. The multivariate FBAT-GEE statistic is then derived by defining a coding vector T_{ij}.

$$T_{ij} = Y_{ij} - \hat{Y}_{ij} = \begin{pmatrix} Y_{ij1} \\ \vdots \\ Y_{ijk} \\ \vdots \\ Y_{ijm} \end{pmatrix} - \begin{pmatrix} \hat{Y}_{ij1} \\ \vdots \\ \hat{Y}_{ijk} \\ \vdots \\ \hat{Y}_{ijm} \end{pmatrix}$$

where the $\hat{Y}_{ijk}$'s are either the observed sample means for the kth trait or the predicted trait values based on a regression model for covariates. Then the univariate coding variable T_{ij} in the FBAT statistic is replaced by the vector T_{ij} and the FBAT-GEE statistic is given by,

$$T_{\text{FBAT}-\text{GEE}} = C^T V^{-1} C.$$

Under the null hypothesis, the FBAT-GEE statistic has a χ^2 distribution with m degrees of freedom. The FBAT-GEE can also be derived based on a generalized estimating equation model with appropriated assumption about the link functions for each phenotype and the covariance structure. However, because the FBAT-GEE test is a score test, all these assumptions cancel out in the derivation of the test statistic and the FBAT-GEE statistic is obtained, making the multivariate FBAT-GEE robust against distribution assumption about the phenotype.

I. A General approach to complex phenotypes: Separating the population and family information in family data

The conditioning of the FBAT statistic on the traits and parental genotypes means that the FBAT statistic does not use all of the information about linkage and association that is available in the sample. While this means we retain robustness, the test is generally not the most efficient. Here we discuss how the extra information in the data that is not used by the FBAT statistic can be used for enhancing power of the test statistics. In particular, with the multiple comparison and model selection issues that arise in modeling complex traits and in large-scale association studies, this extra information can be used to guide the testing strategy. Here, we consider a general approach that is based on separating family data into two independent partitions corresponding to the population information, and the within family information. The population information that is subject to bias by population substructure is used for screening,

or model development, and the within subject information is used for confirmatory testing. The idea is similar to cross-validation, except that the partition into the two components is not random.

Consider a simple case of offspring–parent trios. The full distribution for the data consists of a joint distribution for the offspring phenotype, Y, the offspring genotype, X, and the parental genotype, P (or more generally the sufficient statistics for parental or founder genotypes, S). We partition the joint distribution into two independent parts:

$$P(Y, X, S) = P(X|Y, S)P(S, Y). \quad (3)$$

Model building, hypothesis generation, and screening can be based on S and Y, so that subsequent hypothesis testing using any test statistic whose distribution is based on $P(X|S,Y)$ will be independent of the selected model. Note that Eq. (3) simplifies further under a null hypothesis assuming no linkage because $P(X|Y,S)$ can then be replaced by $P(X|S)$.

There are numerous ways to model $P(S,Y)$, in order to obtain information about association. In general, the approach will depend on the specific design, for example, is Y quantitative or qualitative. To illustrate, consider testing a quantitative phenotype with a single marker. To utilize a population-based approach, we (Lange *et al.*, 2002, 2003a) proposed a "conditional mean model":

$$E(Y) = m + aE(X|S) \quad (4)$$

The parameter a which determines the effect size can be fit using ordinary regression of the phenotype Y on $E(X|S)$. Note that for doubly homozygous parents, $X = E(X|S)$; otherwise we can think of X as missing if parents are informative, and $E(X|S)$ replaces the missing X. In effect, Eq. (4) defines a population regression where some Xs are imputed using parental information (or the sufficient statistics for parental information if parents are missing).

As the regression uses only (Y,S), all the information from the regression will be independent of the FBAT statistic by Eq. (3). Model (4) can be fit repeatedly for any choice of genetic model, any number of phenotypes, and any number of markers. The results of the regression can be used for generating p-values for testing the H0; $a = 0$ or by computing the conditional power of the FBAT statistic for an effect size of a. The power calculation also depends upon the observed parental genotypes and traits (Lange *et al.*, 2002, 2003a). In general, selection based on the conditional power is preferable (Van Steen *et al.*, 2005).

This basic approach has been extended to handle longitudinal and repeated measures (FBAT-PC) (Lange *et al.*, 2004b) and multivariate data (Su *et al.*, 2006) by using the screening stage to select optimal linear combinations of traits for subsequent testing and testing for multiple markers (Xu *et al.*, 2006). Jiang *et al.* (2006) proposed a method to determine the genetically relevant age range for time to onset that is particularly useful for diseases in

which an early onset suggests a strong genetic component, while a late onset is mainly attributable to environmental effects, for example Alzheimer or childhood asthma.

J. Testing strategies for large-scale association studies

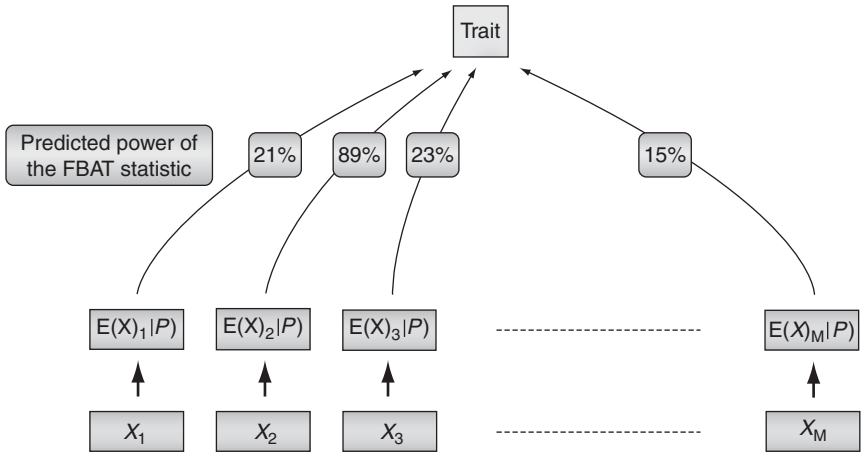
A major scientific obstacle in genome-wide association studies is the hundreds of thousands of SNPs and potential statistical tests that may be computed, resulting in multiple testing issues. Multistage designs have been proposed for case/control studies (Hirschhorn and Daly, 2005; Thomas *et al.*, 2004) as one way of handling this problem. The number of genotyped SNPs is reduced in each stage of the design, so that genome-wide significance is achieved step-by-step. The screening approach for family studies described above is well suited to a genome-wide association study with quantitative traits (Van Steen *et al.*, 2005). The approach is illustrated in Fig. 10.5.

With family-based designs, the screening procedure uses all families, even the "noninformative ones." Assuming moderate to low effect sizes, simulation studies suggest that if a true DSL or a SNP in LD with a DSL is included in the data set, it is sufficient to select the highest 10 or 20 SNPs for further testing and retain high power. The advantage of family-based screening is that the same data set is used for the screening step and the testing step. This means only one sample needs to be recruited, and replication in other studies serves the purpose of generalizing a significant finding to other populations. The strategy has been successful applied to a 100 k SNP scan for obesity in families from the Framingham Heart Study. Table 10.2 displays the top 10 SNPs from that study, as determined by estimated power, selected from the 100 k scan, along with their p values. A novel SNP for BMI was discovered (FBAT p value 0.0026) that would have been missed by standard approaches (e.g., the Bonferroni or Hochberg corrections for multiple testing). Using the same genetic model, the finding was replicated in four independent studies, including cohort, case-control, and family-based samples (Herbert *et al.*, 2006).

III. OTHER APPROACHES TO FAMILY-BASED ANALYSES, INCLUDING THE PDT AND THE QTDT

Many methods have been suggested to handle specific issues that arise with family designs, such as quantitative traits, multiple siblings, or missing parents. It is beyond the scope of this chapter to provide a review of all such methods, but here we mention a few of the more popular methods used to handle general family data, with either quantitative or qualitative outcomes. We first make some

A



B

Power rank	Estimated power of FBAT statistic	SNP	*p* value of FBAT statistic
1	0.92	3	0.90
2	0.89	100	0.20
3	0.85	25	0.00001
...	...	...	...
K	0.70	53	0.20

Figure 10.5. The screening technique. The conditional mean model approach is used to minimize the multiple testing problem. In this example, we look at 1 quantitative trait and M SNPs. In the first step, the marker information in the offspring is assumed to be missing and imputed, using the expected markers scores conditional upon the parental genotypes/sufficient statistic. Then the conditional mean model is used to estimate the power of the FBAT statistic for each SNP. The power depends on observed parental genotypes and the effect size estimated from the conditional mean model. In the final step, the *K* SNPs with the highest power estimates are tested for association with the FBAT statistic at a Bonferroni-adjusted significance level of α'/K. Since only *K* SNPs have been selected for testing, it is only necessary to adjust for *K* comparisons instead of M. (A) *Step 1*: Screen SNPs with conditional mean model for testing via the FBAT statistic. (B) *Step 2*: Select the top *K* SNPs for testing via the FBAT statistic. The *p* value of the FBAT statistic has to be significant at α'/K in order to achieve overall significance. These figures are reproduced from Laird and Lange (2006).

general remarks about likelihood approaches to the analysis and how they connect with the FBAT approach. Then we consider the PDT and QTDT in more detail.

Table 10.2. Screening and Testing of SNPs for Association with BMI

Ranking from screen	SNP	Chromosome	Frequency	Informative families	*p* value FBAT
1	rs3897510	20p12.3	0.36	30	0.2934
2	rs722385	2q32.1	0.16	15	0.1520
3	rs3852352	8p12	0.33	34	0.7970
4	**rs7566605**	**2q14.1**	**0.37**	**39**	**0.0026**
5	rs4141822	13q33.3	0.29	27	0.0526
6	rs7149994	14q21.1	0.35	31	0.0695
7	rs1909459	14q21.1	0.39	38	0.2231
8	rs10520154	15q15.1	0.36	38	0.9256
9	rs440383	15q15.1	0.36	38	0.8860
10	rs9296117	6p24.1	0.40	44	0.3652

Genome-wide SNPs (86,604) were screened using parental genotypes to find those likely to affect offspring BMI. The top 10 SNPs from the screening step (ranked by power from most likely to least likely) are shown. These SNPs were tested using offspring genotypes for association with BMI using the FBAT. The rs7566605 SNP is highlighted in bold.

The general likelihood method specifies a probability density for the observed data, along with a model for how the genotype affects the phenotype. Either likelihood ratio or score tests can be used to test the hypothesis of no association. Self (1991) proposed a likelihood method for case-parent trios based on creating "pseudocontrols," using pairs of the nontransmitted alleles as unaffected siblings. Assuming a relative risk model for the genetic effect yields the conditional logistic regression likelihood; using a log-additive relative risk model yields a likelihood ratio test equivalent to the TDT. As such, this approach can be implemented using standard software packages for conditional logistic regression. This popular approach has been extended to haplotypes (Dudbridge, 2003), gene–environment interactions, and gene–gene interactions (Cordell, 2004). It can be easily generalized to multiple affected offspring in the context of testing a null hypothesis that includes no linkage. When parents are missing and unaffected siblings are available, the conditional logistic regression approach still applies, conditioning on family (Witte *et al.*, 1999).

Another approach for the analysis of trios with dichotomous phenotype data is based on multinomial likelihoods. The contribution of a case parent trio to the likelihood of the data can be factored as:

$$L_t = L_c L_p,$$

where L_c is proportional to the probability density of the child's genotype conditional on parents genotype and the child's disease status and L_p is

proportional to the probability density of the parental mating type, given the child's disease status. Note that L_c depends only on Mendel's laws plus the unknown penetrance functions, that is $P(\text{disease}|X)$, whereas L_p depends on those factors as well as on the mating type frequencies. Thus, robust and efficient tests of association can be constructed from L_c alone that do not require making any assumptions about parental mating type frequencies. Likelihood ratio tests based on L_c are easily constructed for testing genetic association, using different genetic models (e.g., additive, dominant) (Clayton, 1999; Schaid and Li, 1997; Whittemore and Tu, 2000). However, score tests are generally more popular than likelihood ratio tests, partly because they yield the TDT when an additive genetic model is used; these score tests are also equivalent to FBAT tests. More importantly, score tests provide a simple way to extend the model to accommodate multiple offspring, including unaffected, without the need to specify the joint distribution of offspring under the alternative. Score tests only require specifying the distribution of the data under the null. This approach to factoring the likelihood and constructing score tests has been extended to encompass quantitative and time-to-onset phenotypes, as well as multiple offspring (Shih and Whittemore, 2002).

The FBAT statistic is a score test under more general assumptions about the distribution of the offspring phenotypes (Laird *et al.*, 2000). When parents are observed, a score test from L_c as defined above will be equivalent to the FBAT statistic for dichotomous traits. The FBAT and the likelihood approaches diverge in the treatment of missing parental data, or in the case of haplotypes, when parental phase is unknown. When parents are missing, the FBAT approach replaces conditioning on parents by conditioning on the sufficient statistics for parental genotypes, S. However, the likelihood approach estimates the probabilities of parental mating types from the likelihood, L_p, of the observed parents and averages $E(X|P)$ over the estimated distribution of parental mating types. It is thus easy to see that likelihood approaches are generally more efficient: Single cases without any parents do not have to be discarded as they are in the FBAT conditioning approach, but their inclusion relies on the assumption that their parental mating type can be estimated from the data on other parents. This is a strong assumption, and one that is unrealistic in the presence of population substructure.

A series of papers (Kistner and Weinberg, 2004; Kistner *et al.*, 2006; Umbach and Weinberg, 2000; Weinberg, 1999; Weinberg *et al.*, 1998) describes an extension of the multinomial model to the Poisson that allows the incorporation of methods for testing for parental imprinting, gene–environment interaction, and quantitative phenotypes. Rather than score tests, these authors use likelihood ratio tests. To incorporate multiple siblings, they use Wald tests computed with an empirical variance to avoid specifying the joint distribution under the alternative. Missing parents are again handled by estimating a distribution for mating types from the observed parents.

Likelihood-based approaches to the analysis of family data with quantitative traits assume that the trait follows a normal distribution, with mean depending linearly on X (Abecasis *et al.*, 2000; Fulker *et al.*, 1999; Gauderman, 2003). Unlike the setting described above, where inference about association is based on L_c, inferences about the genetic parameters are based directly on the normal likelihood for phenotype given genotype. Note that L_t can alternatively be factored as

$$L_t = L_{yx}\, L_x$$

Where as before, L_t is the likelihood for all the trio data, L_{yx} is the likelihood associated with the phenotype distribution, and L_x is the likelihood of the genotypes, both offspring and parents. Note that basing inferences on L_{yx} is potentially biased but fully efficient because there is no information in L_x about association. A correction for population substructure is made by incorporating $(X_{ij} - E(X_{ij}|P_i))$ into the model for the mean, as described below. We refer to this as a model-based approach because its validity will generally depend upon the correctness of the model for the distribution of the phenotypes. Likelihood ratio tests based on these models can be sensitive to distributional assumptions and ascertainment conditions.

In general, likelihood-based approaches will often be more efficient than score tests or other nonparametric approaches, and offer the possibility of testing nested models, but their validity generally depends upon the correctness of the assumed likelihood model, that is the distributional assumption about the phenotype or the appropriateness of the model for parental mating types when some parents are missing. Simulation studies are a common way of attempting to validate likelihood-based approaches in the presence of model misspecification, but in the absence of theory, simulations offer only limited assurance. In contrast, the validity of the conditioning approach depends on correctly identifying the sufficient statistics and specifying the conditional distribution under the null hypothesis.

A. The PDT and APL

Another family-based association test that is conceptionally very similar to the FBAT-approach is the PDT, the pedigree disequilibrium test (Martin *et al.*, 1997). Both the PDT and the FBAT are score tests that share a similar numerator, but instead of using $E(X|S)$ with missing parents, these tests use contrasts between affected and unaffected offspring in the same family. Additionally, the PDT approach relies on empirical variance estimators instead of variances that are analytically derived. The PDT approach is a valid test for all three null hypotheses. To obtain a more powerful test for association in the presence of linkage (second

null hypothesis), the PDT approach has been extended to incorporate linkage information, the so-called "test for association in the presence of linkage" (Martin *et al.*, 2000). In contrast to the FBAT approach, which does not attempt to estimate the joint transmission probabilities to the offspring, but estimates directly the variance of the test statistic using an empirical variance/covariance estimator, the APL approach models the joint transmission probabilities to multiple offspring based on the identity by descent status. The transmission probabilities are estimated using the EM-algorithm and are used in the computation of the test statistic directly. The approach has been extended to general nuclear families with missing parental information, to haplotype analysis, to the analysis of the X-chromosome, and to handle ordinal/rank-based phenotypes such as age of onset (Chung *et al.*, 2006, 2007a,b; Martin *et al.*, 2003).

B. Quantitative traits: The QTDT

The QTDT is a widely used method for testing association with family data and quantitative traits. The essential idea relies on assuming a standard QTL model for the phenotype, that is,

$$E(Y_{ij}) = \mu + \beta X_{ij}$$

To make a correction for population substructure admixture, we may add and subtract $\beta\, E(X_{ij}|P_i)$ to the mean to obtain:

$$E(Y_{ij}) = \mu + \beta[X_{ij} - E(X_{ij}|P_i)] + \beta E(X_{ij}|P_i)$$

Subscripting the two βs by w and b, to denote within subject and between subject effects, we write:

$$E(Y_{ij}) = \mu + \beta_w[X_{ij} - E(X_{ij}|P_i)] + \beta_b E(X_{ij}|P_i) \quad (4)$$

As shown in Abecasis *et al.* (2000), OLS estimates of β_w remain unbiased for the true genetic effect in the presence of population stratification, while estimates of β_b are contaminated by population stratification. The model 4 is the basis for the QTDT test that further assumes normality of the error terms and uses a likelihood ratio test for the null hypothesis H0: $\beta_w = 0$, leaving β_b unspecified. A score test for H0: $\beta_w = 0$ that is derived from the same likelihood function will be equivalent to the FBAT statistic for quantitative traits (Lange *et al.*, 2003a).

The original model proposed by Fulker *et al.* (1999) was designed for sib pairs, or more generally sibships, with missing parental genotypes. Here, assuming an additive model, $E(X_{ij}|P_i)$ is replaced by the mean of the offspring genotypes in the ith family that is equivalent to $E(X_{ij}|S_i)$. In this case, the model can be expanded to include variance and covariance terms between siblings, which

include components due to the genetic variance, and residual sibling resemblance. With siblings, likelihood ratio tests can also be constructed to test for linkage. However, when the observed marker is the DSL (perfect LD between the marker and the DSL), there is no power to test for linkage. As the LD between the marker and DSL decreases, the power to detect association of the trait and the marker decreases, but the power to detect linkage increases (Sham *et al.*, 2000).

An advantage of the approach is that it provides estimates of the association parameter, β_w, and the recombination parameter. However, the method does not lend itself to extension to haplotypes or multiple endpoints. Although the estimate of β_w is robust for population substructure, the contaminated estimate for the between-component β_b does not cancel out in the construction of the likelihood ratio test, causing an inflated type-1 error (Yu *et al.*, 2006).

IV. SOFTWARE

With family-based designs, there is generally a need for special software to analyze the data. Fortunately there is now a wide variety of software packages available. Most of the packages were developed by the original authors of the methods and are home-grown. Despite the lack of general support for such software packages in academia, the packages have proven to be reliable and user-friendly tools. Recently, commercial packages with professional user-support and documentation have become available that are particularly suited for less statistical-oriented users and for large-scale projects. Table 10.3 shows an overview of the most popular packages and their functions.

V. DISCUSSION

The advent of whole-genome association scans offers great promise for genetic association studies. Most projections agree that large samples of individuals will be necessary to disentangle the wheat from the chaff in these large genome scans, no matter what the design (Clayton *et al.*, 2005; Hirschhorn and Daly, 2005; Van Steen *et al.*, 2005). While it is inescapable that large samples from existing cohort or case-control studies that do not include data on relatives are generally much easier to obtain than large numbers of suitable families, we believe that the creative use of the population information contained in family data for screening and hypothesis generation, coupled with their robustness to population substructure, make these family studies competitive. In addition, with technology available for handling pedigrees with missing founders, family data already collected for linkage studies can, in many cases, be recycled for association.

Table 10.3. Software Programs for the Analysis of Family-Based Association Tests

Package	Genetic analysis capability	Phenotypic analysis capability	Special features
APL/PDT	Single marker, haplotype	Binary traits, quantitative traits, ranked traits, time-to-onset	X-chromosome
FBAT	Single marker, haplotype, multimarker	Binary traits, quantitative/multivariate traits, ranked traits, time-to-onset	X-chromosome, permutation tests
P2BAT/PBAT, PBAT GoldenHelix	Single marker, haplotype, multimarker	Binary traits, quantitative traits/multivariate, ranked traits, time-to-onset, gene–environment interaction	Covariate adjustment, VanSteen algorithm for multiple testing, X-chromosome, permutation tests
QTDT	Single marker	Quantitative traits	Permutation tests

There are several features of family-based designs that make them less attractive than their population-based counterparts. One feature is the considerable sensitivity to genotyping errors (Gordon and Ott, 2001; Gordon *et al.*, 2001, 2004). It is clear that genotyping errors can lead to false inferences because the test distribution depends crucially on the assumption that parental genotypes are correct. In the population-based setting, nondifferential genotyping errors will only make tests conservative under the null, but with FBATs, random genotyping errors can inflate the false positive rate, sometimes substantially (Hirschhorn and Daly, 2005). This underscores the importance of validating the genotyping for any significant finding.

There is widespread belief that gene–environment (and gene–drug) interactions as well as gene–gene interactions play an important role in many complex diseases. For example, genetic interactions with smoking status and/or smoking history are believed to be determinants of the severity of chronic obstructive pulmonary disease (Celedon *et al.*, 2004; DeMeo *et al.*, 2005).

Possible unmeasured gene–environment or gene–gene interactions are often thought to be responsible for lack of reproducibility of many genetic associations. Although such hypotheses about interactions between genes and exposure variables are widely accepted in the field, adequate development of statistical methodology to test such hypotheses has lagged behind other technical developments.

While some family-based approaches have been proposed for special settings (Cordell *et al.*, 2004; Gauderman, 2002; Lake and Laird, 2004; Umbach and Weinberg, 2000; Witte *et al.*, 1999), the general problem is hampered by reliance on statistical models for main and interaction effects (which may not correspond to biological models) (Cordell *et al.*, 2004) and the difficulty of testing interactions when main effects are assumed to be present.

Much work remains to develop better statistical methods for complex disorders that are characterized by multiple, possibly interacting genes and environmental factors, and are characterized by numerous interrelated traits. The focus here has been on family-based designs for testing this is appropriate when mapping via linkage disequilibrium is the major objective. However, as we move from discovery to verification and characterization, the focus will appropriately shift to effect estimation. The development of models for complex disease phenotypes with multiple covariates, genes, and interactions will be crucial for characterizing the role of individual polymorphisms in complex disease. The challenge will be to model pathways incorporating several genes at the same time, in combination with gene–environment interactions and endophenotypes, where modest effects may add up to a substantial impact on disease.

References

Abecasis, G. R., Cardon, L. R., and Cookson, W. O. C. (2000). A general test of association for quantitative traits in nuclear families. *Am. J. Hum. Genet.* **66**(1), 279–292.

Bickeboller, H., and Clerget-Darpoux, F. (1995). Statistical properties of the allelic and genotypic transmission/disequilibrium test for multiallelic markers. *Genet. Epidemiol.* **12**(6), 865–870.

Celedon, J. C., Lange, C., Raby, B. A., Litonjua, A. A., Palmer, L. J., DeMeo, D. L., Reilly, J. J., Kwiatkowski, D. J., Chapman, H. A., Laird, N., Sylvia, J. S., Hernandez, M., *et al.* (2004). The transforming growth factor-beta1 (TGFB1) gene is associated with chronic obstructive pulmonary disease (COPD). *Hum. Mol. Genet.* **13**(15), 1649–1656.

Chung, R. H., Hauser, E. R., and Martin, E. R. (2006). The APL test: Extension to general nuclear families and haplotypes and examination of its robustness. *Hum. Hered.* **61**(4), 189–199.

Chung, R. H., Hauser, E. R., and Martin, E. R. (2007a). Interpretation of simultaneous linkage and family-based association tests in genome screens. *Genet. Epidemiol.* **31**(2), 134–142.

Chung, R. H., Morris, R. W., Zhang, L., Li, Y. J., and Martin, E. R. (2007b). X-APL: An improved family-based test of association in the presence of linkage for the X chromosome. *Am. J. Hum. Genet.* **80**(1), 59–68.

Clayton, D. (1999). A generalization of the transmission/disequilibrium test for uncertain-haplotype transmission. *Am. J. Hum. Genet.* **65**(4), 1170–1177.

Clayton, D. G., Walker, N. M., Smyth, D. J., Pask, R., Cooper, J. D., Maier, L. M., Smink, L. J., Lam, A. C., Ovington, N. R., Stevens, H. E., Nutland, S., Howson, J. M., *et al.* (2005). Population structure, differential bias and genomic control in a large-scale, case-control association study. *Nat. Genet.* **37**(11), 1243–1246.

Cordell, H. J. (2004). Properties of case/pseudocontrol analysis for genetic association studies: Effects of recombination/ascertainment, and multiple affected offspring. *Genet. Epidemiol.* **26**(3), 186–205.

Cordell, H. J., Barratt, B. J., and Clayton, D. G. (2004). Case/pseudocontrol analysis in genetic association studies: A unified framework for detection of genotype and haplotype associations, gene-gene and gene-environment interactions, and parent-of-origin effects. *Genet. Epidemiol.* **26**(3), 167–185.

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

Curtis, D., and Sham, P. C. (1995b). A note on the application of the transmission disequilibrium test when a parent is missing. *Am. J. Hum. Genet.* **56**(3), 811–812.

DeMeo, D. L., and Silverman, E. K. (2003). Genetics of chronic obstructive pulmonary disease. *Semin. Respir. Crit. Care Med.* **24**(2), 151–160.

DeMeo, D. L., Mariani, T. J., Lange, C., Srisuma, S., Litonjua, A. A., Celedon, J. C., Lake, S. L., Reilly, J. J., Chapman, H. A., Mecham, B. H., Haley, K. J., Sylvia, J. S., *et al.* (2005). The SERPINE2 gene is associated with chronic obstructive pulmonary disease. *Am. J. Hum. Genet.* **78**(2), 253–264.

Dudbridge, F. (2003). Pedigree disequilibrium tests for multilocus haplotypes. *Genet. Epidemiol.* **25**(2), 115–121.

Falk, C. T., and 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.

Fulker, D. W., Cherny, S. S., Sham, P. C., and Hewitt, J. K. (1999). Combined linkage and association sib-pair analysis for quantitative traits. *Am. J. Hum. Genet.* **64**(1), 259–267.

Gauderman, W. J. (2002). Sample size requirements for matched case-control studies of gene-environment interaction. *Stat. Med.* **21**(1), 35–50.

Gauderman, W. J. (2003). Candidate gene association analysis for a quantitative trait, using parent-offspring trios. *Genet. Epidemiol.* **25**(4), 327–338.

Gordon, D., and Ott, J. (2001). Assessment and management of single nucleotide polymorphism genotype errors in genetic association analysis. *Pac. Symp. Biocomput.* 18–29.

Gordon, D., Heath, S. C., Liu, X., and Ott, J. (2001). A transmission disequilibrium test that allows for genotyping errors in the analysis of single nucleotide polymorphism data. *Am. J. Hum. Genet.* **69**(4), 507–507.

Gordon, D., Haynes, C., Johnnidis, C., Patel, S. B., Bowcock, A. M., and Ott, J. (2004). A transmission disequilibrium test for general pedigrees that is robust to the presence of random genotyping errors and any number of untyped parents. *Eur. J. Hum. Genet.* **12**(9), 752–761.

Guo, C.-Y., Lunetta, K. L., DeStefano, A. L., Ordovas, J. M., and Cupples, L. A. (2007). Informative-transimission disequilibrium test (i-TDT): Combined linkage and association mapping that includes unaffected offspring as well as affected offspring. *Genet. Epidemiol.* **31,** 115–133.

Herbert, A., Gerry, N. P., McQueen, M. B., Heid, I. M., Pfeufer, A., Illig, T., Wichmann, H. E., Meitinger, T., Hunter, D., Hu, F. B., Colditz, G., Hinney, A., *et al.* (2006). A common genetic variant is associated with adult and childhood obesity. *Science* **312**(5771), 279–283.

Hirschhorn, J. N., and Daly, M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nat. Rev. Genet.* **6**(2), 95–108.

Horvath, S., and Laird, N. M. (1998). A discordant-sibship test for disequilibrium and linkage: No need for parental data. *Am. J. Hum. Genet.* **63**(6), 1886–1897.

Horvath, S., Xu, X., and Laird, N. M. (2001). The family based association test method: Strategies for studying general genotype-phenotype associations. *Eur. J. Hum. Genet.* **9**(4), 301–306.

Horvath, S., Xu, X., Lake, S. L., Silverman, E. K., Weiss, S. T., and Laird, N. M. (2004). Family-based tests for associating haplotypes with general phenotype data: Application to asthma genetics. *Genet. Epidemiol.* **26**(1), 61–69.

Jiang, H., Harrington, D., Raby, B. A., Bertram, L., Blacker, D., Weiss, S. T., and Lange, C. (2006). Family-based association test for time-to-onset data with time-dependent differences between the hazard functions. *Genet. Epidemiol.* **30**(2), 124–132.

Kistner, E. O., and Weinberg, C. R. (2004). Method for using complete and incomplete trios to identify genes related to a quantitative trait. *Genet. Epidemiol.* **27**(1), 33–42.

Kistner, E. O., Infante-Rivard, C., and Weinberg, C. R. (2006). A method for using incomplete triads to test maternally mediated genetic effects and parent-of-origin effects in relation to a quantitative trait. *Am. J. Epidemiol.* **163**(3), 255–261.

Laird, N. M., and Lange, C. (2006). Family-based designs in the age of large-scale gene-association studies. *Nat. Rev. Genet.* **7**(5), 385–394.

Laird, N. M., Horvath, S., and Xu, X. (2000). Implementing a unified approach to family-based tests of association. *Genet. Epidemiol.* **19,** S36–S42.

Lake, S. L., and Laird, N. M. (2004). Tests of gene-environment interaction for case-parent triads with general environmental exposures. *Ann. Hum. Genet.* **68**(Pt 1), 55–64.

Lake, S. L., Blacker, D., and Laird, N. M. (2000). Family-based tests of association in the presence of linkage. *Am. J. Hum. Genet.* **67**(6), 1515–1525.

Lange, C., and Laird, N. M. (2002a). Power calculations for a general class of family-based association tests: Dichotomous traits. *Am. J. Hum. Genet.* **71**(3), 575–584.

Lange, C., and Laird, N. M. (2002b). On a general class of conditional tests for family-based association studies in genetics: The asymptotic distribution, the conditional power, and optimality considerations. *Genet. Epidemiol.* **23**(2), 165–180.

Lange, C., DeMeo, D. L., and Laird, N. M. (2002). Power and design considerations for a general class of family-based association tests: Quantitative traits. *Am. J. Hum. Genet.* **71**(6), 1330–1341.

Lange, C., DeMeo, D., Silverman, E. K., Weiss, S. T., and Laird, N. M. (2003a). Using the noninformative families in family-based association tests: A powerful new testing strategy. *Am. J. Hum. Genet.* **73**(4), 801–811.

Lange, C., Silverman, E. K., Xu, X., Weiss, S. T., and Laird, N. M. (2003b). A multivariate family-based association test using generalized estimating equations: FBAT-GEE. *Biostatistics* **4**(2), 195–206.

Lange, C., Blacker, D., and Laird, N. M. (2004a). Family-based association tests for survival and times-to-onset analysis. *Stat. Med.* **23**(2), 179–189.

Lange, C., Van Steen, K., Andrew, T., Lyon, H., DeMeo, D., Raby, B., Murphy, A., Silverman, E., MacGregor, A., Weiss, S., and Laird, N. M. (2004b). A family-based association test for repeatedly measured quantitative traits adjusting for unknown environmental and/or polygenic effects. *Stat. Appl. Genet. Mol. Biol* **3**(1), http://www.bepress.com/sagmb/vol3/iss1/art17

Lasky-Su, J., Biederman, J., Laird, N., Tsunag, M., Doyle, A. E., Smoller, J. W., Lange, C., and Faraone, S. V. (2007). Evidence for an association of the dopamine D5 receptor gene on age at onset of attention deficit hyperactivity disorder. *Am. J. Hum. Genet.* **71**(5), 648–659.

Lunetta, K. L., Faraone, S. V., Biederman, J., and Laird, N. M. (2000). Family-based tests of association and linkage that use unaffected sibs, covariates, and interactions. *Am. J. Hum. Genet.* **66**(2), 605–614.

Martin, E. R., Kaplan, N. L., and Weir, B. S. (1997). Tests for linkage and association in nuclear families. *Am. J. Hum. Genet.* **61**(2), 439–448.

Martin, E. R., Monks, S. A., Warren, L. L., and Kaplan, N. L. (2000). A test for linkage and association in general pedigrees: The pedigree disequilibrium test. *Am. J. Hum. Genet.* **67**(1), 146–154.

Martin, E. R., Bass, M. P., Hauser, E. R., and Kaplan, N. L. (2003). Accounting for linkage in family-based tests of association with missing parental genotypes. *Am. J. Hum. Genet.* **73**(5), 1016–1026.

Naylor, M., Raby, B., Weiss, S. T., and Lic, C. (2005). Generalizability of regression models for the fev1-measurements in asthmatic children. *Far East J. Theor. Stat.* **16,** 351–369.

Ott, J. (1989). Statistical properties of the haplotype relative risk. *Genet. Epidemiol.* **6**(1), 127–130.

Rabinowitz, D. (1997). A transmission disequilibrium test for quantitative trait loci. *Hum. Heredity* **47**(6), 342–350.

Rabinowitz, D., and Laird, N. (2000). A unified approach to adjusting association tests for population admixture with arbitrary pedigree structure and arbitrary missing marker information. *Hum. Heredity* **50**(4), 211–223.

Rakovski, C. S., Xu, X., Lazarus, R., Blacker, D., and Laird, N. M. (2007). A new multimarker test for family-based association studies. *Genet. Epidemiol.* **31**(1), 9–17.

Risch, N., and Zhang, H. (1995). Extreme discordant sib pairs for mapping quantitative trait loci in humans. *Science* **268**(5217), 1584–1589.

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

Schaid, D. J. (1996). General score tests for associations of genetic markers with disease using cases and their parents. *Genet. Epidemiol.* **13**(5), 423–449.

Schaid, D. J., and Li, H. Z. (1997). Genotype relative-risks and association tests for nuclear families with missing parental data. *Genet. Epidemiol.* **14**(6), 1113–1118.

Schneiter, K., Laird, N., and Corcoran, C. (2005). Exact family-based association tests for biallelic data. *Genet. Epidemiol.* **29**(3), 185–194.

Self, S. G. (1991). An adaptive weighted log-rank test with application to cancer prevention and screening trials. *Biometrics* **47**(3), 975–986.

Sham, P. C., Cherny, S. S., Purcell, S., and Hewitt, J. K. (2000). Power of linkage versus association analysis of quantitative traits, by use of variance-components models, for sibship data. *Am. J. Hum. Genet.* **66**(5), 1616–1630.

Shih, M. C., and Whittemore, A. S. (2002). Tests for genetic association using family data. *Genet. Epidemiol.* **22**(2), 128–145.

Spielman, R. S., and Ewens, W. J. (1996). The TDT and other family-based tests for linkage disequilibrium and association. *Am. J. Hum. Genet.* **59**(5), 983–989.

Spielman, R. S., and Ewens, W. J. (1998). A sibship test for linkage in the presence of association: The sib transmission/disequilibrium test. *Am. J. Hum. Genet.* **62,** 450–458.

Spielman, R. S., McGinnis, R. E., and Ewens, W. J. (1993). Transmission test for linkage disequilibrium: The insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am. J. Hum. Genet.* **52**(3), 506–516.

Terwilliger, J. D., and Ott, J. (1992). A haplotype-based haplotype relative risk approach to detecting allelic associations. *Hum. Heredity* **42**(6), 337–346.

Thomas, D., Xie, R., and Gebregziabher, M. (2004). Two-Stage sampling designs for gene association studies. *Genet. Epidemiol.* **27**(4), 401–414.

Umbach, D. H., and Weinberg, C. R. (2000). The use of case-parent triads to study joint meets of genotype and exposure. *Am. J. Hum. Genet.* **66**(1), 251–261.

Van Steen, K., McQueen, M. B., Herbert, A., Raby, B., Lyon, H., DeMeo, D. L., Murphy, A., Su, J., Datta, S., Rosenow, C., Christman, M., Silverman, E. K., *et al.* (2005). Genomic screening and replication using the same data set in family-based association testing. *Nat. Genet.* **37**(7), 683–691.

Vansteelandt, S., DeMeo, D. L., Su, J., Smoller, J., Murphy, A. J., McQueen, M., Schneiter, K., Celedon, J. C., Weiss, S. T., Silverman, E. K., and Lange, C. (2007). Testing and estimating gene-environment interactions with complex exposure variables. *Biometrics* (in press).

Ware, J. H., and Weiss, S. (1996). Statistical issues in longitudinal research on respiratory health. *Am. J. Respir. Crit. Care Med.* **154**(6 Pt 2), S212–S216.

Weinberg, C. R. (1999). Methods for detection of parent-of-origin effects in genetic studies of case-parents triads. *Am. J. Hum. Genet.* **65**(1), 229–235.

Weinberg, C. R., Wilcox, A. J., and Lie, R. T. (1998). A log-linear approach to case-parent-triad data: Assessing effects of disease genes that act either directly or through maternal effects and that may be subject to parental imprinting. *Am. J. Hum. Genet.* **62**(4), 969–978.

Whittaker, J. C., and Lewis, C. M. (1998). The effect of family structure on linkage tests using allelic association. *Am. J. Hum. Genet.* **63,** 889–897.

Whittemore, A. S., and Tu, I. P. (2000). Detection of disease genes by use of family data. I. Likelihood-based theory. *Am. J. Hum. Genet.* **66**(4), 1328–1340.

Witte, J. S., Gauderman, W. J., and Thomas, D. C. (1999). Asymptotic bias and efficiency in case-control studies of candidate genes and gene-environment interactions: Basic family designs. *Am. J. Epidemiol.* **149**(8), 693–705.

Xu, X., Rakovski, C., Xu, X. P., and Laird, N. (2006). An efficient family-based association test using multiple markers. *Genet. Epidemiol.* **30**(7), 620–626.

Yu, J. M., Pressoir, G., Briggs, W. H., Bi, I. V., Yamasaki, M., Doebley, J. F., McMullen, M. D., Gaut, B. S., Nielsen, D. M., Holland, J. B., Kresovich, S., and Buckler, E. S. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* **38**(2), 203–208.

Zhao, H. (2000). Family-based association studies. *Stat. Methods Med. Res.* **9**(6), 563–587.