Cost Efficiency of Genetic Linkage Studies Using Mixtures of Selected Sib-pairs

Jian Han and Rudy Guerra

Abstract

Sib-pairs are relatively easy to collect and use of extreme quantitative phenotypes provide high statistical power. Thus, selected sib-pair (discordant, concordant) study designs are among the most useful in quantitative genetic linkage analysis. Dudoit and Speed [5, 6] proposed a score test for linkage that allows analysis of any sample, random or selected, by conditioning on phenotype and analyzing genotype. Selected sampling strategies have largely focused on studies collecting data on a single type of sib-pair. Using the score test statistic, we demonstrate that sampling designs based on a mixture of sib-pair types are more cost efficient than the traditional single selection scheme. In particular, there is no need to discard a large fraction of screened individuals. Cost efficient designs are based on a mixture of concordant and discordant sib-pairs, with the selection threshold of concordant sib-pairs more stringent than that of discordant pairs. General guidelines for the thresholds are given as a function of mode of inheritance, allele frequency, and residual correlation, as well as the cost ratio of phenotyping to genotyping. Since in many cases the mode of inheritance is not completely known, robustness with respect to assumed genetic models is also addressed.

Keywords: linkage; sample size; selected sample; sib-pair; study design

1 Introduction

It is well known that quantitative genetic linkage analysis based on random sampling of sib-pairs usually has low statistical power to detect non-Mendelian, quantitative, or complex disease loci. For example, Blackwelder and Elston [3] showed in simulations that even when heritability is moderate (30%) at a single locus, the power of the Haseman-Elston [11] linkage test based on random sampling of sib-pairs is low. Significant improvement in power can be achieved when an unselected sibling is regressed on the proportion (π) of alleles shared identical-by-descent (IBD) with a selected sibling [4]. Eaves and Meyer [7] provide evidence of additional power increases depending on the types of sib-pairs selected: discordant, with the sib-pair representing both tails of the phenotypic distribution; or concordant high (low), where both siblings are selected from the upper (lower) tail of the phenotypic distribution. However, see Allison *et al.* [1] for some limitations on the general utility of selected sib-pairs. Risch and Zhang

[13] also argue for selected sib-pair designs, but unlike previous authors they propose conditioning on the sampled (observed) phenotypes to analyze IBD sharing among the sib-pairs. Their discussion, however, is limited to study designs sampling a fixed type of sib-pair.

Dudoit and Speed [5, 6] generalize the work of Risch of Zhang [13] in three respects. First, although Dudoit and Speed also analyze IBD data, they specifically test for linkage in the traditional sense of evaluating a null hypothesis involving a recombination fraction ($H_0 : \theta = 0.5$), whereas Risch and Zhang evaluate a null hypothesis involving average allele-sharing ($H_0 : \pi = 0.5$). Second, Dudoit and Speed condition on observed phenotypes, while Risch and Zhang condition on phenotypic deciles. Lastly, the mean IBD statistic of Risch and Zhang is interpretable only for a fixed sib-pair type (*e.g.*, all discordant sib-pairs using a fixed threshold); the Dudoit and Speed score test statistic is not restricted to a fixed sampling scheme. By definition, both approaches reflect the actual sampling (conditioning on phenotype) and stochastic nature of the outcome (allele-sharing). In this sense they depart from making assumptions that are clearly violated under methods that model or analyze the phenotype, while viewing allele-sharing as "fixed" design variables. Both approaches are seemingly limited by having to specify knowledge of the gene action underlying the phenotype-genotype association. Robustness studies by Risch and Zhang [13], Zhao, Zhang, and Rotter [15], and Goldstein, Dudoit, and Speed [8], however, show that various characteristics of the approaches are fairly insensitive to misspecifying the mode of inheritance.

One drawback of selected study designs is that a large number of sib-pairs usually need to be screened in order to obtain the minimum sample size (number of sib-pairs) for the desired power. The more stringent the selection thresholds, the more screening that has to be done. Zhao *et al.* [15] evaluate cost efficiency across extremely discordant (ED), concordant high (CH), and concordant low (CL) sib-pair study designs. Considering the three types of designs separately, they conclude that ED sib-pair studies are the most cost efficient and robust against incorrect mode of inheritance and allele frequencies. They note, however, that more cost efficient studies may be possible by using all three types of selected sib-pairs.

Gu *et al.* [10] and Gu and Rao [9] report increased power over ED designs by combining all three types of sib-pairs into a single test statistic. In addition, they show that using all three types of sib-pairs is more cost effective than using just ED pairs. One issue that has yet to be fully addressed, in these and other investigations, is that the three types of sib-pairs may not be equally available in the population. Indeed, their prevalence is highly dependent on the underlying mode of inheritance. Ignoring this fact may lead to inefficient study designs, especially at screening where a lot of time and resources may be required to obtain certain extreme phenotypes that are relatively rare under the true gene action. Related to this idea is that better power and cost efficiency may be achieved by allowing different thresholds for the various sib-pair types.

In summary, there are not yet available general optimal sampling designs, defined by power or cost efficiency, for genetic linkage studies using selected sib-pairs. In this

paper we use the score test of Dudoit and Speed [5, 6] to develop such optimal sampling strategies.

2 Methods

Following standard major locus models (*e.g.*, Haseman and Elston [11]; Amos and Guerra [2]), we assume a locus A with two alleles, A_1 and A_2, with population allele frequencies p, $q\ (=1-p)$, respectively. Let x_{1i} and x_{2i} be the sib-pair phenotypic values of sib-pair i. The sib-pair phenotypes are modeled as:

$$\begin{aligned} x_{1i} &= \mu + g_{1i} + e_{1i}, \\ x_{2i} &= \mu + g_{2i} + e_{2i}, \end{aligned}$$

where μ is an overall mean of x; g_{ji} is the genetic effect due to trait locus A; e_{ji} represents combined residual genetic and environmental contributions with variance σ_e^2. The genetic effect g_{ji} equals a, d, and $-a$ according to genotypes A_1A_1, A_1A_2, and A_2A_2, respectively. To account for residual genetic and environmental correlations, we assume that the sib-pair model error (e_{1i}, e_{2i}) is distributed as a bivariate normal distribution with zero mean vector and correlation coefficient ρ. The additive and dominant components of genetic variation at locus A are defined as $\sigma_a^2 = 2pq[a-d(p-q)]^2$ and $\sigma_d = (2pqd)^2$. The heritability due to locus A is defined as $H = \sigma_g^2/(\sigma_g^2+\sigma_e^2)$, where $\sigma_g^2 = \sigma_a^2 + \sigma_d^2$.

Under the null hypothesis of no linkage between a marker locus and trait locus, the proportion of genes shared IBD at the marker locus is expected to be $1/2$ regardless of the type of sib-pairs collected. When linkage is present mean IBD sharing among ED (CH/CL) pairs is expected to be less (more) than $1/2$. The test statistic used by Risch and Zhang [13] is the sample average for IBD sharing; it has an (asymptotic) null Gaussian distribution with mean zero and variance $1/8$. The Gaussian distribution under the alternative hypothesis of linkage depends on the selection scheme.

Gu *et al.* [10] proposed the extremely discordant and concordant (EDAC) test statistic, which combines ED, CH, and CL sib-pairs. It is defined as

$$T = \frac{1}{2(n_2+n_0)}\left\{\sum_{i=1}^{n_2}[X_{1i}(h,h)+X_{2i}(h,h)] + \sum_{i=1}^{n_0}[X_{1i}(l,l)+X_{2i}(l,l)]\right\} - \frac{1}{2n_1}\sum_{i=1}^{n_1}[X_{1i}(h,l)+X_{2i}(h,l)],$$

where h and l are indices of high and low tail thresholds [*e.g.* (10%,10%)=(10,10)], respectively; n_0 is the number of CL pairs, n_1 the number of ED pairs, and n_2 the number of CH pairs. X_{1i} is the number of alleles shared IBD from the father and X_{2i} is the number of alleles shared IBD from the mother. Consequently, $X_{1i}+X_{2i}$ is the number of allele shared IBD from the parents. The test statistic is thus a difference

between average proportion IBD-sharing among concordant sib-pairs and that among discordant sib-pairs. Under the null hypothesis, T is asymptotically distributed as a Gaussian random variable with mean 0 and variance $\sigma^2(T) = (n_1 + n_2 + n_0)/(8n_1(n_2 + n_0))$. A one-sided test statistic is given by $T/\sigma(T)$. Formulas for calculating sample size for ED pairs (and therefore for CH and CL pairs) are provided in Gu *et al.* [10].

In general, when a (test) statistic is formed through a linear combination of several available statistics, the number of observations entering each individual statistic and the weights in the combination typically have practical meaning and interpretation. In the present context, both factors may be motivated by mode of inheritance considerations. Risch and Zhang [13], for example, note that ED pairs are universally most useful among all possible types of sib-pairs; whereas CH or CL pairs are useful depending on mode of inheritance and allele frequency, when only a single type of sib-pair is used. Risch and Zhang [13, 14] and Zhao *et al.* [15] also discuss appropriate thresholds for selection sampling. Consider, for example, sampling top 10% and bottom 10% discordant sib-pairs. Under moderate to high positive residual correlation these extreme discordant pairs are relatively more difficult to find than concordant sib-pairs. It is possible to take advantage of the positive correlation by requiring a more stringent selection threshold to recruit more informative concordant pairs. Continuing with our example, we might set a threshold of (10,10) for ED pairs, while selecting top 5% CH pairs and bottom 5% for CL pairs. This flexibility may allow for increased statistical power and better cost efficiency by better selecting more informative sib-pairs. The score test of Dudoit and Speed allows for a broad range of selection strategies.

Dudoit and Speed [5, 6] proposed a score test for evaluating a null hypothesis of no linkage, $H_0 : \theta = 1/2$, against an alternative, $H_1 : 0 \leqslant \theta < 1/2$. The test can be used with the major gene model defined above. The statistic is

$$S(\nu) = 16 \sum_{i=1}^{n} (\pi_{2i} - \pi_{0i})(N_{2i} - N_{0i}),$$

where ν represents genetic parameters, such as values of a, d, p, σ_e^2, ρ, and mode of inheritance (recessive, dominant, additive); π_{2i} is the conditional probability that the *ith* sib-pair shares 2 genes IBD at the trait locus, given sib-pair phenotype (x_{1i}, x_{2i}); π_{0i} is similarly defined as the probability of sharing 0 genes IBD at the trait locus. N_{ji} is an indicator variable, equal to 1 if sib-pair i shares j ($j = 0 or 2$) genes IBD and 0 otherwise.

Under the null hypothesis, S is asymptotically normal with mean 0 and variance $\sigma^2(S) = \frac{1}{2}\sum_{i=1}^{n}(\pi_{2i} - \pi_{0i})^2$. The null hypothesis is rejected at level α when $S/\sigma(S) > z_\alpha$. Under the alternative hypothesis, S is asymptotically distributed as a normal random variable, $N(\mu_A, \sigma_A^2)$, where

$$\begin{aligned} \mu_A &= \sum_{i=1}^{n} (\pi_{2i} - \pi_{0i})(\tau_{2i} - \tau_{0i}), \\ \sigma_A^2 &= \sum_{i=1}^{n} (\pi_{2i} - \pi_{0i})^2 (\tau_{2i}\bar{\tau}_{2i} + \tau_{0i}\bar{\tau}_{0i} + 2\bar{\tau}_{0i}\bar{\tau}_{2i}); \end{aligned}$$

τ_{2i} is the conditional probability that the *ith* sib-pair shares 2 genes IBD at the *marker* locus, given sib-pair phenotype (x_{1i}, x_{2i}); τ_{0i} is similarly defined as the probability of sharing 0 genes IBD at the marker locus; $\bar{\tau} = 1 - \tau$.

The conditional asymptotic power of S, given the phenotypes, is denoted as

$$\Gamma(\theta, \nu; \mathbf{X}) = 1 - \Phi\left(\frac{z_\alpha \sqrt{\frac{1}{2}\sum_{i=1}^{n}(\pi_{2i} - \pi_{0i})^2} - \mu_A}{\sigma_A}\right), \tag{1}$$

where $\mathbf{X}$ is a $n \times 2$ matrix representing the n sib-pair phenotypes, and Φ is standard normal cumulative distribution function such that $\Phi(z_\alpha) = 1 - \alpha$. The unconditional power may be estimated by the average of a set of conditional powers generated under the same model.

The score test is derived through an approximation of the maximum likelihood ratio test and is locally most powerful [5, 6]. One criticism of the test is that it requires specification of a mode of inheritance model for the weights (π) to be determined. (The observed data are the counts, N.) As has been noted, however, the test appears to be sufficiently robust with respect to mode of inheritance assumptions (Goldstein, Dudoit and Speed [8]). The important feature of the test is that it faithfully reflects the non-random sampling that is typical of most genetic epidemiologic studies. We refer readers to the original papers for more technical details.

3 Sample Size Approximation

Although "randomly selected" sib-pairs may be of some utility - and would likely be available through the screening process - in this article we focus on using only ED, CH, and CL sib-pairs. Since the score test is conditional on observed data, there are no closed-form formulas for sample size calculations associated with the unconditional power based on (1); however, it is possible to approximate the power function for a given sampling selection. Under the assumption that the trait and marker loci are in complete linkage ($\theta = 0$), we have $\pi_{ji} = \tau_{ji}$ and expression (1) becomes

$$\Gamma(\nu; \mathbf{X}) = 1 - \Phi\left(\frac{Z_\alpha \sqrt{\frac{1}{2}\mu_A} - \mu_A}{\sigma_A}\right), \tag{2}$$

with

$$\mu_A = \sum_{i=1}^{n}(\pi_{2i} - \pi_{0i})^2, \tag{3}$$

$$\sigma_A^2 = \sum_{i=1}^{n}(\pi_{2i} - \pi_{0i})^2(\pi_{2i}\bar{\pi}_{2i} + \pi_{0i}\bar{\pi}_{0i} + 2\bar{\pi}_{0i}\bar{\pi}_{2i}). \tag{4}$$

To determine the sample sizes, the probability parameters (π) need to be estimated. To this end, define selection schemes (S), $TxBy$, $TxTx$, $BxBx$, where T and B indicate "top" (upper) and "bottom" (lower) tails of the phenotypic distribution, x and y tail areas. For example, the selection scheme $T10B10$ requires sib-pair phenotype (x_1,x_2) to satisfy $max(x_1,x_2) > p_{90}$ and $min(x_1,x_2) < p_{10}$, where p_h is the h^{th} percentile of the (marginal) phenotypic distribution.

By definition, $\pi_2 = P(\text{sib-pair shares 2 trait genes IBD} \mid x_1,x_2)$. Under a given selection scheme (S), π_2 can be estimated by

$$\hat{\pi}_2 \doteq \int\int_S \pi_2 \, dx_1 dx_2.$$

This is equivalent to the estimation of D_2 in equation (1) of Risch and Zhang [13]. Similarly, $\hat{\pi}_0$ can also be used to estimate π_0.

Let n be the total selected sample size, $n = n_{ED} + n_{CH} + n_{CL}$, where n_{ED}, n_{CH} and n_{CL} are the sample sizes of selected ED, CH and CL sib-pairs, respectively. For a specified genetic model and selection scheme, let P_{ED}, P_{CH}, and P_{CL} be the probability of randomly selecting an ED, CH, and CL sib-pair from the phenotype distribution; define $r_{ED} = P_{ED}/(P_{ED} + P_{CH} + P_{CL})$, the proportion of ED pairs in the population of ED, CH, and CL sib-pairs. Proportions r_{CH} and r_{CL} are similarly defined. Lastly, let $\hat{\pi}_{ED2}$ and $\hat{\pi}_{ED0}$ be estimates (as defined above) of π_2 and π_0, respectively, for ED sib-pairs. Denote similar estimates for CH and CL sib-pairs.

The mean (μ_A) of the estimated score statistic can thus be estimated as

$$\begin{aligned}
\mu_A &= \sum_{i=1}^{n_{ED}} (\pi_{2i} - \pi_{0i})^2 + \sum_{i=1}^{n_{CH}} (\pi_{2i} - \pi_{0i})^2 + \sum_{i=1}^{n_{CL}} (\pi_{2i} - \pi_{0i})^2 \\
&\approx n_{ED}(\hat{\pi}_{ED2} - \hat{\pi}_{ED0})^2 + n_{CH}(\hat{\pi}_{CH2} - \hat{\pi}_{CH0})^2 + n_{ED}(\hat{\pi}_{CL2} - \hat{\pi}_{CL0})^2 \\
&\approx n[r_{ED}(\hat{\pi}_{ED2} - \hat{\pi}_{ED0})^2 + r_{CH}(\hat{\pi}_{CH2} - \hat{\pi}_{CH0})^2 + r_{ED}(\hat{\pi}_{CL2} - \hat{\pi}_{CL0})^2] \\
&\stackrel{\text{def}}{=} nW.
\end{aligned}$$

In a similar way, the variance (σ_A^2) of the test statistic can be estimated, say nU. Substituting parameter estimates in the conditional power function (2) we obtain

$$Power = 1 - \beta \stackrel{\text{def}}{=} 1 - \Phi\left(\frac{z_\alpha\sqrt{\frac{1}{2}nW} - nW}{nU}\right),$$

and the corresponding sample size n is given by

$$n = \left[\frac{z_\alpha\sqrt{\frac{1}{2}W} - z_\beta\sqrt{U}}{W}\right]^2.$$

The sample sizes for ED, CH, and CL sib-pairs are then calculated as nr_{ED}, nr_{CH}, nr_{CL}, respectively.

We emphasize that when the selection schemes are determined, the number of ED, CH and CL sib-pairs to be selected are calculated according to their selection probability under an assumed genetic model. This method of selection makes use of the extreme sib-pairs relatively readily available in the population and minimizes wasting resources attempting to find sib-pairs that may be difficult to collect under the genetic model. Although the sample size calculation is an approximate one, simulations show that the observed power is always at least as great as the nominal power; see below.

Since the score test weights ED, CH and CL sib-pairs according to a working genetic model, it may be more powerful than the EDAC test whereby the three types of sib-pairs are treated equally. Therefore, for fixed power and type I error probability, the score test may require smaller sample sizes than the EDAC approach.

***Example* 1**

Table 1 shows sample sizes corresponding to $H = 0.3$, $1 - \beta = 0.8$ and $\alpha = 0.001$ under a $T10B10$ selection scheme for ED, $T5T5$ for CH, and $B5B5$ for CL sib-pairs. Under all parameter configurations considered ($\rho = 0.2$, 0.4; $p = 0.1, \ldots, 0.9$; recessive, dominant, and additive models), the two tests indicate the same qualitative pattern of sample size requirements. For example, when $\rho = 0.4$ and $p = 0.3$ under a recessive model, both tests require $n_{ED} \leqslant n_{CL} \leqslant n_{CH}$. However, the score test always requires smaller sample sizes in terms of the total (n) and specific sib-pairs (n_{ED}, n_{CH} or n_{CL}. Table 2 gives the average percent reduction in total sample size of the score test relative to the EDAC test. Higher reductions are obtained in the presence of higher residual correlation. The smallest average reduction (10%) is observed under additive gene action with lower residual correlation. Table 1 shows that ED sib-pairs are less informative when the sib-pairs have a relatively higher degree of (positive) residual correlation; compare n_{ED} sample sizes at $\rho = 0.2$ to $\rho = 0.4$ under each test. This makes sense since a higher degree of (positive) correlation would tend to make the phenotypes more similar. Indeed, both the score test and EDAC test have n_{CH} and n_{CL} each larger than n_{ED} when $\rho = 0.4$. The score test is also less sensitive then EDAC with respect to the given increase in residual correlation. Associated with an increase from $\rho = 0.2$ to $\rho = 0.4$ under the recessive model, the average (across p) percent increase in total sample size (n) for the score test is 28%; under the dominant model the average increase is 24% and under the additive it is 20%. The corresponding results for the EDAC test are 74%, 49%, and 50% under recessive, dominant, and additive models, respectively. Relatively larger sample sizes for extreme discordant sib-pairs are generally observed under a dominant model with lower residual correlation ($\rho = 0.2$), and $n_{ED} \approx n_{CH} > n_{CL}$ under an additive model with $\rho = 0.2$. In most other cases concordant sib-pairs are required more so that discordant pairs. The observed pattern of overall results remained the same when other selection schemes were considered (data not shown). ■

Table 1: Sample sizes requirements for score test and EDAC test. $H = 0.3$, power = .8 and $\alpha = 0.001$, with selection scheme T10B10 for ED, T5T5 for CH and B5B5 for CL sib-pairs. Number of ED, CH and CL sib-pairs denoted by n_{ed}, n_{ch} and n_{cl}, respectively; $n = n_{ed} + n_{ch} + n_{cl}$.

		Score Test	EDAC Test	Score Test	EDAC Test
		$\rho = 0.2$	$\rho = 0.2$	$\rho = 0.4$	$\rho = 0.4$
	p	$n(n_{ed}, n_{ch}, n_{cl})$	$n(n_{ed}, n_{ch}, n_{cl})$	$n(n_{ed}, n_{ch}, n_{cl})$	$n(n_{ed}, n_{ch}, n_{cl})$
Rec	0.1	229(99,75,55)	580(251,190,139)	329(52,145,132)	1036(165,456,415)
	0.3	60(19,27,14)	102(32,46,24)	64(10,31,23)	152(23,73,56)
	0.5	128(38,51,39)	146(44,58,44)	160(18,74,68)	257(28,120,109)
	0.7	165(46,60,59)	165(46,60,59)	241(22,110,109)	311(29,141,141)
	0.9	111(29,32,50)	139(36,40,63)	134(13,54,67)	245(23,99,123)
Dom	0.1	103(36,41,26)	116(41,46,29)	118(18,55,45)	164(25,77,62)
	0.3	138(52,43,43)	143(54,44,45)	194(28,83,83)	211(31,90,90)
	0.5	117(45,31,41)	131(50,35,46)	141(23,56,62)	183(30,73,80)
	0.7	63(24,13,26)	93(35,20,38)	64(14,22,28)	118(25,40,53)
	0.9	231(101,55,75)	578(253,137,188)	330(57,130,143)	970(168,382,420)
Add	0.1	90(30,38,22)	108(36,46,26)	100(14,49,37)	148(21,72,55)
	0.3	105(34,39,32)	109(35,41,33)	128(16,59,53)	155(19,71,65)
	0.5	112(34,39,39)	112(34,39,39)	139(15,62,62)	166(18,74,74)
	0.7	112(31,36,45)	119(33,38,48)	139(14,60,65)	184(18,79,87)
	0.9	94(24,25,45)	126(32,34,60)	110(11,43,56)	209(20,82,107)

4 Relative Importance of ED, CH, CL Sib-pairs

It is well known [13, 6] that extreme discordant sib-pairs are generally most powerful when a single selection scheme is used. Gu *et al.* [10] argue that concordant sib-pairs available in the screening pool provide an important additional source of linkage information and should be included in the selected sample. However, several practical questions remain unanswered, including the following. What are the relative merits of the various sib-pair types in a given study design? More specifically, how are the individual sample sizes (n_{ED}, n_{CH}, n_{CL}) related to linkage information. For a given level of power, are studies carried out with ED pairs alone more or less cost efficient than those that include mixtures of concordant and discordant sib-pairs? How should the thresholds for the different sib-pair types be chosen?

***Example* 2**

As a motivating example, consider an additive model with heritability $H = 0.3$, allele (A_1) frequency $p = 0.2$, and sib-pair residual correlation $\rho = 0.4$. Under this model a selection scheme of T15B15 for ED, T10T10 for CH, and B10B10 for CL pairs

Table 2: Percent decrease in score test total sample size (n) relative to EDAC test. Model parameters as in Table 1.

	Recessive		Dominant		Additive	
p	$\rho = 0.2$	$\rho = 0.4$	$\rho = 0.2$	$\rho = 0.4$	$\rho = 0.2$	$\rho = 0.4$
.1	61	78	11	28	17	32
.3	41	58	3	8	4	17
.5	12	38	11	23	0	16
.7	0	23	32	46	6	25
.9	20	45	60	66	25	47
Ave	27	48	23	34	10	27

corresponds to selection probabilities $P(ED) = 0.0103, P(CH) = 0.030$, and $P(CL) = 0.027$. At 80% power and type I error probability $\alpha = 0.001$, the required sample sizes are $n_{ED} = 28$, $n_{CH} = 81$, and $n_{CL} = 73$. Figure 1 shows the relative importance of each kind of pair. In plot (a) the number of ED pairs is fixed at 28, and numbers of CH and CL pairs vary from zero to the require sample size. When both n_{CH} and n_{CL} are zero, the power is 38%. The power gradually increases as more CH pairs are added, but the CL pairs do not affect power very much. Plot (b) clearly shows the importance of ED pairs with n_{CH} fixed at 81. Using the required 81 CH pairs alone yields a power of 26%. The relative importance of both ED and CH pairs is jointly exhibited in plot (c) where there are 73 concordant low pairs. Under this particular model, ED pairs affect power the most; CH pairs contribute as well, but the usefulness of CL pairs is very limited (73 CL pairs alone has power of nearly zero). As shown in plot (d), the sample size needed to achieve power of 0.8 using only ED pairs is 52. The expected number of randomly screened sib-pairs to obtain 52 ED pairs is $52/(0.0103) = 5048$; the expected number of randomly screened sib-pairs to obtain the mixed sample is $182/(0.0103 + 0.03 + 0.027) = 2704$. Note that the last calculation is not based on an optimal selection scheme, which may further reduce the screening size. ■

Example 2 makes evident that adding more sib-pairs (concordant or discordant) in the sample provides an increase in the power of the test, albeit possibly small. This is generally true for the score test regardless of the mode of inheritance, allele frequency and residual correlation. The EDAC test, however, occasionally loses power when CH pairs are combined with ED pairs.

***Example* 3**

In Example 2, the CL sib-pairs were least important in their contribution to the power of test, but this is not always the case. Consider a dominant model with $H = 0.3$, allele (A_1) frequency $p = 0.8$, and $\rho = 0.4$. The selection scheme is as in Example 2,

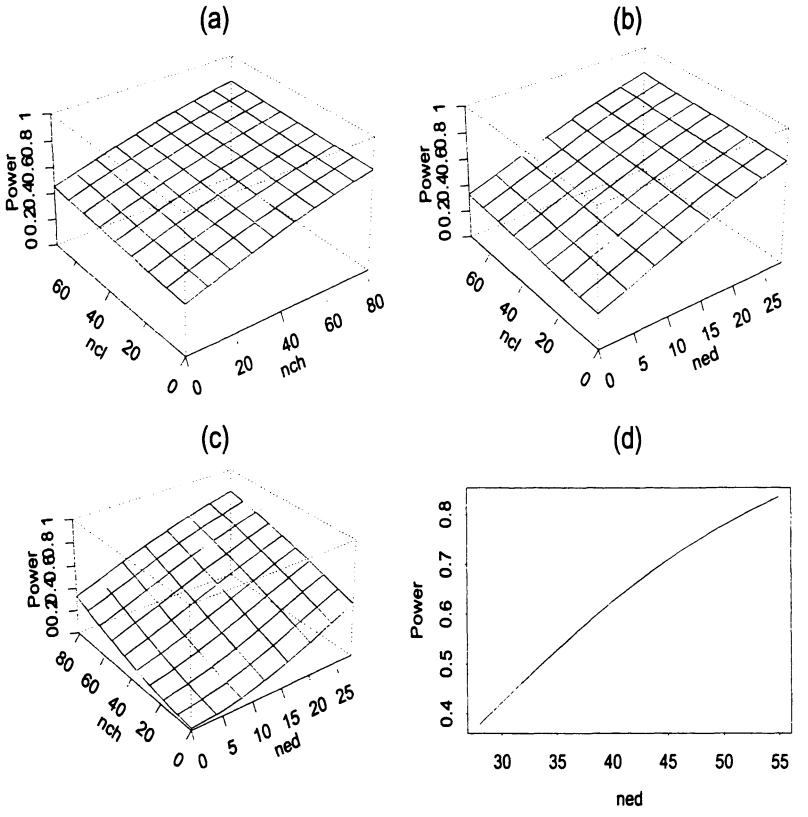

Figure 1: Power of score test for various combinations of sample sizes under selection scheme T15B15 for ED, T10T10 for CH and B10B10 for CL pairs. Nominal power = 0.8 and α = 0.001. True genetic parameters include H=0.3, p=0.2, ρ = 0.4, additive gene action. (a) Power of test for fixed number of ED pairs, $n_{ed} = 28$. (b) Power of test for fixed number of CH pairs, $n_{ch} = 81$. (c) Power of test for fixed number of CL pairs, $n_{cl} = 73$. (d) Power of test using only ED pairs.

$T15B15$ for ED, $T10T10$ for CH and $B10B10$ for CL pairs. Under this dominant model the selection probabilities for ED, CH and CL pairs are 0.0146, 0.0265, and 0.0295, respectively. At 80% power with $\alpha = 0.001$, the sample sizes for ED, CH and CL pairs are 43, 78, and 87, respectively. Contrary to the results of the previous example, the CL pairs are the most important in terms of power contribution, whereas both ED and CH pairs have a limited role; see plots (a), (b) and (c) of Figure 2. Using CL sib-pairs alone the test has moderate power at 68%. As shown in plot (d), a study with only $T15B15$ ED sib-pairs requires approximately 225 pairs to achieve the desired power of 80%. The expected number of sib-pairs screened for this ED-only study is $223/0.0146 = 15273$, compared with $208/(0.0146 + 0.0265 + 0.0295) = 2946$ for a mixture study. ■

The examples illustrate the potential savings in screening by using mixed sib-pair types in the score test. We also see that some sib-pair types are less useful than others in determining the power of the test.

5 Optimal Mixture of Sib-pairs

In this section we address the issue of optimal selection thresholds for ED, CH and CL sib-pairs in the selected sample in order to minimize the total cost of phenotyping and genotyping. Given a desired power and type I error probability, our goal is to find the optimal selection scheme for the score test such that the cost of the test is minimized.

We assume that sib-pairs are randomly chosen from the population and that the ratio (R) of phenotyping-to-genotyping cost ranges as 0.02, 0.1, 1, 10, 50. Eight selection thresholds for ED sib-pairs are considered: T10B10, T10B20, T10B25, T15B15, T15B25, T20B20, T25B25 and T30B70. Since moderate to high (positive) residual correlation makes it difficult to find an ED sib-pair, more stringent thresholds for this type of sib-pair are not considered here. Among these selection schemes, some are symmetric (*e.g.*, T10B10) and some are asymmetric (*e.g.*, T10B25). We consider seven symmetric selection schemes for CH (CL) pairs: T1T1, T3T3, T5T5, T10T10, T15T15, T20T20, T25T25 (B1B1, B3B3, B5B5, B10B10, B15B15, B20B20, B25B25). For now we assume equivalent tail areas for CH and CL sib-pairs (*e.g.*, T5T5 and B5B5). The more stringent thresholds for concordant pairs are chosen since they are relatively easier to recruit than ED pairs under positive residual correlation. Thus, the total number of selection schemes considered is 56 (8×7). This seems to be wide enough coverage to be practically useful. Heritability H is fixed at 0.1 or 0.3, allele frequency p ranges from 0.1 to 0.9 (by 0.2), residual correlation (ρ) takes values 0.1 or 0.4. The total number of genetic models considered is $2 \times 2 \times 5 \times 3 = 60$ (heritability $\times$ correlation $\times$ p $\times$ gene action).

The total cost of interest is the sum of the cost for phenotyping all screened sib-pairs required to obtain the total sample size and the cost of genotyping the selected sib-pairs. The total cost (TC) is calculated as $TC = 2RN + 2n$ (Zhao *et al.* [15]), where $n = n_{ED} + n_{CH} + n_{CL}$. N is the expected total number of screened sib-pairs calculated as $n/[P(ED) + P(CH) + P(CL)]$; R the cost ratio of phenotyping to genotyping. Without loss of generality, the cost of genotyping one individual is assumed to be 1 unit in the calculation of total cost. For each of the 60 genetic models, the optimal (minimum cost) sampling is obtained by searching all 56 stated selection schemes for a fixed cost ratio (R) of phenotyping to genotyping. For purposes of comparison with what might be considered accepted convention, we also report results for ED-only study designs; minimum cost is found among the eight ED selection schemes.

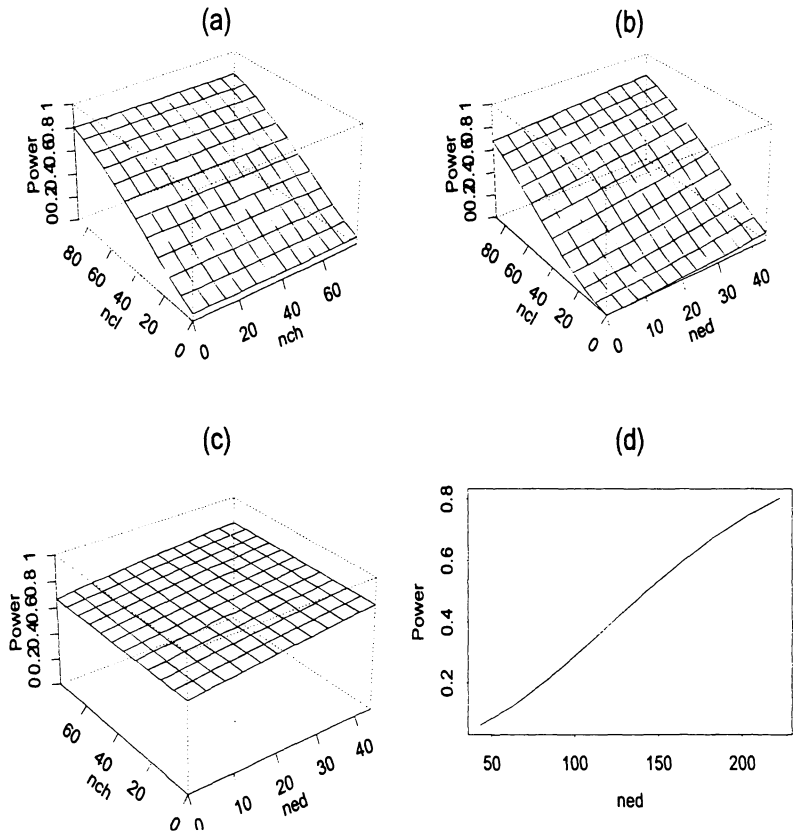

Figure 2: Power of score test for various combinations of sample sizes under selection scheme T15B15 for ED, T10T10 for CH and B10B10 for CL pairs. Nominal power = 0.8 and α = 0.001. True genetic parameters include H=0.3, p=0.8, ρ = 0.4, dominant gene action. (a) Power of test for fixed number of ED pairs, n_{ed} = 43. (b) Power of test for fixed number of CH pairs, n_{ch} = 78. (c) Power of test for fixed number of CL pairs, n_{cl} = 87. (d) Power of test using only ED pairs.

***Example* 4**

We first consider the case where phenotyping cost is very low compared to the cost of genotyping, $R = 0.02$, and $H = 0.3$. Optimal selection schemes with corresponding sample sizes and total costs (in thousand units) are listed in Table 3. Analogous results for an ED-only study are given in the right part of the table. Column 3 shows the optimal selection scheme for the given parameters; since both types of concordant pairs have equal tail areas their selection schemes are summarized as Tx/Bx. Column 4 shows the required sample sizes for the optimal sampling; column 5, the total cost (thousand units) when mixed (m) sib-pairs are used. Columns 6 – 8 show results for ED-only optimal studies. The dominant case is not reported since it is equivalent to a recessive case with upper and lower thresholds switched and p replaced by $1 - p$.

We first consider the recessive model. Here mixed samples require smaller total sample sizes and are generally more cost efficient than ED-only samples. At lower allele frequencies (approximately 0.3 or less), mixed samples are much more cost efficient than ED-only samples, independent of residual correlation. At the higher allele frequencies, discordant sib-pairs are generally more informative than are concordant pairs in that the ED pairs constitute the majority of the total sample size. Concordant high pairs are slightly more (less) informative than concordant low pairs at the lower (higher) allele frequencies; they are equally informative at $p \approx 0.5$. At higher residual correlation ($\rho = 0.4$), the optimal thresholds for discordant pairs become less stringent with increasing allele frequency. This relationship holds less so under weaker residual correlation ($\rho = 0.1$). This is what we might expect, since at higher degrees of (positive) residual correlation, ED pairs are observed as such because of linkage effects overriding the residual correlation. On the other hand, with higher residual correlation it is less clear whether concordant pairs are phenotypically similar because of genes or residual factors, which may or may not reflect genetic factors. Thus, ED pairs are relatively more important and the relaxing thresholds under $\rho = 0.4$ reflect the need to collect them. Conversely, the concordant thresholds are relatively more extreme in order to distinguish the genetic signal from the "noise" in the residual correlation. Under the additive case, the mixed and ED-only samples are about equally cost efficient, with the exception at very high allele frequencies (0.9 or higher). And, as in the recessive case, a higher degree of residual correlation is associated with less (more) stringent optimal thresholds for discordant (concordant) sib-pairs. The recessive and additive cases also share in common the fact that in most cases the concordant pairs represent a small fraction of the total sample size. ■

***Example* 5**

When phenotyping and genotyping costs are about the same ($R \approx 1$, Table 4), the total costs increase compared to the case $R < 1$ not only because of higher costs per individual, but also because the total sample sizes increase as well. Characteristics of study

Table 3: Optimal selection schemes found from all 56 possible selection combinations for recessive (top) and additive (bottom) model with $H = 0.3$. The cost ratio of phenotype-to-genotype is $R = 0.02$. Power = 0.8, $\alpha = 0.001$.

Recessive Model

ρ	p	ED, CH/CL	$n(n_{ed}, n_{ch}, n_{cl})$	COST	ED	n	COST
0.1	0.1	T10B10, T1/B1	53(43,9,1)	0.234	T10B25	8604	26.4
	0.3	T10B10, T5/B5	63(26,26,1)	0.237	T10B25	163	0.58
	0.5	T10B25, T5/B5	125(85,23,17)	0.412	T10B25	113	0.441
	0.7	T15B15, T1/B1	113(109,2,2)	0.477	T15B15	111	0.476
	0.9	T15B15, T5/B5	134(84,17,33)	0.444	T15B15	138	0.565
0.4	0.1	T10B10, T1/B1	24(12,9,3)	0.172	T10B20	1004	5.92
	0.3	T10B25, T5/B5	68(25,24,19)	0.216	T10B25	76	0.387
	0.5	T15B25, T1/B1	86(78,4,4)	0.353	T15B25	82	0.355
	0.7	T25B25, T1/B1	114(108,3,3)	0.357	T25B25	110	0.352
	0.9	T20B80, T3/B3	105(68,16,21)	0.359	T20B20	91	0.38

Additive Model

ρ	p	ED,CHCL	$n(n_{ed}, n_{ch}, n_{cl})$	COST	ED	n	COST
0.1	0.1	T10B25,T3B3	95(79,11,5)	0.307	T10B25	106	0.366
	0.3	T10B20,T3B3	100(81,11,8)	0.367	T10B20	96	0.389
	0.5	T15B15,T3B3	114(95,9,10)	0.413	T15B15	108	0.425
	0.7	T15B15,T5B5	126(80,19,27)	0.423	T15B15	112	0.464
	0.9	T10B10,T5B5	92(33,19,40)	0.374	T15B15	153	0.617
0.4	0.1	T10B25,T1B1	50(43,4,3)	0.229	T10B25	50	0.251
	0.3	T15B25,T1B1	69(63,3,3)	0.263	T15B25	67	0.264
	0.5	T20B20,T1B1	75(68,3,4)	0.284	T20B20	71	0.283
	0.7	T20B20,T1B1	77(69,4,4)	0.307	T20B20	74	0.312
	0.9	T15B15,T3B3	77(33,18,26)	0.236	T20B20	97	0.404

design and costs when $R = 1$, compared to $R = 0.02$, include a more prominent role of concordant sib-pairs, less stringent optimal thresholds, and higher gains in sample sizes and costs by the mixed sampling scheme. By relaxing the thresholds, we are able to recruit the desired number of sib-pair types without the need to screen prohibitively large numbers of sib-pairs. ■

The impact of residual correlation on the optimal mixture selection scheme is summarized in Table 5. The selection schemes for ED sib-pairs are listed in column 1 from most stringent (top) to least stringent (bottom); selection schemes for CH and CL are given in the first row from most stringent (left) least stringent(right). The top (bottom) panel gives results for $\rho = 0.1$ ($\rho = 0.4$). The entry is the frequency of the intersecting combination of ED and CHCL pairs defining an optimal design among 56 choices. For

Table 4: Optimal selection schemes found from all 56 possible selection combinations for recessive (top) and additive (bottom) models with $H = 0.3$. The cost ratio of phenotype-to-genotype is $R = 1$. Power = 0.8, $\alpha = 0.001$.

Recessive Model

ρ	p	ED,CHCL	$n(n_{ed}, n_{ch}, n_{cl})$	COST	ED	n	COST
0.1	0.1	T10B10,T1B1	53(43,9,1)	6.44	T10B25	8604	475
	0.3	T10B25,T10B10	105(42,40,23)	3.5	T10B25	163	13
	0.5	T15B25,T20B20	311(73,129,109)	4.78	T30B30	445	9.03
	0.7	T30B30,T25B25	492(178,157,157)	4.67	T30B30	306	6.92
	0.9	T25B25,T15B15	308(150,64,94)	4.9	T25B25	352	10.7
0.4	0.1	T10B20,T1B1	43(31,9,3)	6.27	T10B25	1255	177
	0.3	T15B25,T10B10	134(38,52,44)	3.64	T15B25	125	11.3
	0.5	T30B30,T15B15	300(120,93,87)	4.18	T30B30	208	6.63
	0.7	T30B30,T25B25	450(95,178,177)	4.11	T30B30	157	5.58
	0.9	T30B30,T15B15	308(118,89,101)	4.22	T30B30	229	7.44

Additive Model

ρ	p	ED,CHCL	$n(n_{ed}, n_{ch}, n_{cl})$	COST	ED	n	COST
0.1	0.1	T10B25,T15B15	197(53,85,59)	4.28	T10B25	106	7.87
	0.3	T25B25,T20B20	380(153,118,109)	4.64	T30B30	381	6.99
	0.5	T25B25,T20B20	379(144,117,117)	4.77	T30B30	468	7.32
	0.7	T25B25,T20B20	394(141,121,132)	5.1	T30B30	381	8.16
	0.9	T25B25,T15B15	337(166,71,100)	5.31	T25B25	405	12.1
0.4	0.1	T15B25,T10B10	151(43,58,50)	3.97	T15B25	75	6.51
	0.3	T30B30,T15B15	285(122,84,79)	3.72	T30B30	176	4.91
	0.5	T30B30,T15B15	287(116,85,86)	3.88	T30B30	170	5.18
	0.7	T30B30,T15B15	301(114,92,95)	4.2	T30B30	176	5.18
	0.9	T30B30,T15B15	338(131,97,110)	4.6	T30B30	259	8.28

example, among the 150 parameter configurations defining a genetic model, there were 16 that had as an optimal design T10B10 ED, T1T1 CH and B1B1 CL sib-pair types. When residual correlation increases (decreases), the marginal counts of discordant pairs shift toward less (more) stringent thresholds. This general pattern corroborates the specific results seen in Tables 3 and 4. Considering discordant and concordant selection jointly, we observe that the majority of counts occur along the diagonal at $\rho = 0.1$, while most of the counts are located below the diagonal at $\rho = 0.4$. Consequently, in the presence of positive residual correlation we should not plan studies that combine extreme discordant pairs with less extreme concordant sib-pairs.

Similar summary counts in Table 6 are stratified by low ($R = 0.02, 0.1$) and high ($R = 1, 10, 50$) phenotype-to-genotype costs . When phenotyping cost is relatively low,

Table 5: Counts of optimal mixture selection schemes among all 150 possible selection combinations when residual correlation $\rho = 0.1$ (top) and $\rho = 0.4$ (bottom). Other parameters are $H = 0.1, 0.3$, $p = 0.1, 0.3, 0.5, 0.7, 0.9$, the cost ratio of phenotype-to-genotype $R = 0.02, 0.1, 1, 10, 50$, under recessive, dominant, and additive gene action. Power = 0.8, α is 0.001. Under CHCL is shown the selection scheme for CH and CL pairs; for example, T10B10 means T10T10 for CH pairs, and B10B10 for CL pairs.

	$\rho = 0.1$							
				CHCL				
ED	T1B1	T3B3	T5B5	T10B10	T15B15	T20B20	T25B25	Sum
T10B10	16	0	14	2	0	0	0	32
T10B20	0	1	2	0	0	0	0	3
T10B25	4	1	2	14	1	0	0	22
T15B15	2	1	3	11	0	0	0	17
T15B25	0	0	0	2	5	8	0	15
T20B20	0	0	1	5	0	0	0	6
T25B25	0	0	0	3	5	11	0	19
T30B30	0	0	0	0	1	5	30	36
Sum	22	3	22	37	12	24	30	150
	$\rho = 0.4$							
				CHCL				
ED	T1B1	T3B3	T5B5	T10B10	T15B15	T20B20	T25B25	Sum
T10B10	14	0	1	0	0	0	0	15
T10B20	1	0	0	0	0	0	0	1
T10B25	11	1	3	0	0	0	0	15
T15B15	1	4	1	0	0	0	0	6
T15B25	2	0	6	5	0	0	0	13
T20B20	8	1	3	3	0	0	0	15
T25B25	2	0	10	3	1	0	0	16
T30B30	0	1	0	6	27	12	23	69
Sum	39	7	24	17	28	12	23	150

fairly stringent discordant and concordant sib-pairs (upper left region) should be collected. Conversely, when phenotyping cost is relatively high, less stringent conditions are indicated. Of course, these observations are general guidelines; more specific designs are possible with more information other than just the phenotype-to-genotype cost ratio. However, in cases when very little is known about the underlying genetic factors, one may not know more than the costs involved.

Lastly, we summarize the comparison of costs between the optimal mixed sample and optimal ED-only sample; Figure 3 gives an overview. The y-axis represents ED:mixed cost ratio, and the x-axis indexes an ordered set of parameters as given below:

for R=(0.02, 0.1, 1, 10, 50)
  for H=(0.1, 0.3)

Table 6: Counts of optimal mixture selection schemes among all 150 possible selection combinations when phenotype-to-genotype cost ratio $R = 0.02, 0.1$ (top) and $R = 1, 10, 50$ (bottom). Other parameters are $H = 0.1, 0.3$; $p = 0.1, 0.3, 0.5, 0.7, 0.9$; $\rho = 0.2, 0.4$; recessive, dominant, and additive gene action. Nominal power = 0.8 and $\alpha = 0.001$. Under CHCL is shown the selection scheme for CH and CL pairs; for example, T10B10 means T10T10 for CH pairs, and B10B10 for CL pairs.

	$R = 0.02, 0.1$							
				CHCL				
ED	T1B1	T3B3	T5B5	T10B10	T15B15	T20B20	T25B25	Sum
T10B10	16	0	15	2	0	0	0	33
T10B20	0	1	2	0	0	0	0	3
T10B25	6	2	5	8	0	0	0	21
T15B15	3	5	4	8	0	0	0	20
T15B25	2	0	6	2	0	0	0	10
T20B20	8	1	4	4	0	0	0	17
T25B25	2	0	10	3	0	0	0	15
T30B30	0	1	0	0	0	0	0	1
Sum	37	10	46	27	0	0	0	120

	$R = 1, 10, 50$							
				CHCL				
ED	T1B1	T3B3	T5B5	T10B10	T15B15	T20B20	T25B25	Sum
T10B10	14	0	0	0	0	0	0	14
T10B20	1	0	0	0	0	0	0	1
T10B25	9	0	0	6	1	0	0	16
T15B15	0	0	0	3	0	0	0	3
T15B25	0	0	0	5	5	8	0	18
T20B20	0	0	0	4	0	0	0	4
T25B25	0	0	0	3	6	11	0	23
T30B30	0	0	0	6	28	17	53	104
Sum	24	0	0	27	40	36	53	180

for Mode=(recessive, additive)
  for ρ=(0.1, 0.3)
    for p=(0.1, 0.3, 0.5, 0.7, 0.9)

For example, the first 5 points correspond to R=0.02, H=0.1, recessive gene action, ρ=0.1 and p=0.1, 0.3, 0.5, 0.7 or 0.9; the next 5 points correspond to R=0.02, H=0.1, recessive gene action, ρ=0.3 and p=0.1, 0.3, 0.5, 0.7 or 0.9, and so on. Although the costs for both designs increase as R increases, the ratio ED:mixed is generally between 1 and 2. In plot (a), the 10 pairs of high-low peaks correspond to the 20 combinations of R, mode-of-inheritance, and ρ. Within each pair the decrease reflects increases in ρ; across pairs the peak magnitudes reflect changes in mode-of-inheritance. When the model is recessive with infrequent allele $(1 - p)$, the optimal cost from the test with

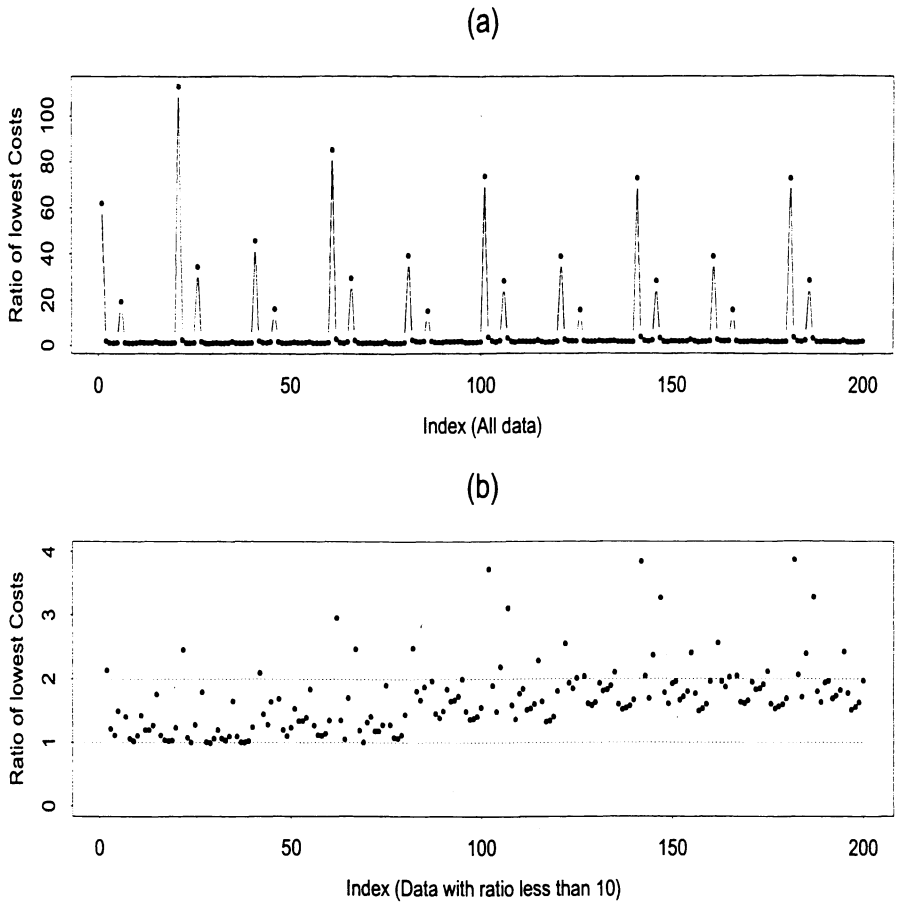

Figure 3: The ratio of lowest cost from the test using ED sib-pairs alone to the lowest cost of the mixture test. The lowest cost is chosen from all possible selection schemes under each genetic models: $H = 0.1, 0.3$, recessive, additive, $\rho = 0.1, 0.4$, and $p = 0.1, 0.3, 0.5, 0.7, 0.9$. Power = 0.8 and $\alpha = 0.001$. (a) Plot of ratios of all available data. (b) Plot of ratios below 10.

ED pairs alone can attain 20 – 100-fold increases over the mixed sample. A similar conclusion holds for a dominant model with frequent allele (p). More information about the low ratios is shown in plot (b), where the high ratios (greater than 10) are not shown. In some cases the ratios are close to 1, especially when $R = 0.02$ or $R = 0.1$.

6 Discussion

Extremely discordant and extremely concordant (high or low) sib-pairs are among the most useful sib-pairs in genetic linkage analysis of quantitative trait loci (Risch and

Zhang [13]). Essentially, selecting sib-pairs mimics a designed experiment whereby known genotypes are typically compared by phenotypic averages (*e.g.*, analysis of variance). Comparing averages (first moments) is much more powerful than analyzing variances (second moments), as is the basis of the Haseman and Elston [11] robust sib-pair method for linkage. By selecting extreme discordant or concordant sib-pairs we are enriching the sample with individuals that are more likely to be in the tails of the phenotypic distribution because of genotype rather than chance.

Gu *et al.* [10] extended the mean IBD test of Risch and Zhang to incorporate the three types of pairs into a single test (EDAC). In this paper, we show the advantages of the score test for linkage, developed by Dudoit and Speed [5, 6], when multiple selection schemes are possible. Under the mixed selection strategy, the score test provides more power than the EDAC test by weighting each kind of sib-pair according to its linkage information under an assumed genetic model. As the basis for inclusion is the underlying biological mechanisms, it is not surprising that the score test performs better than an alternative that combines test statistics largely on the basis of statistical principles, although the latter has been shown to significantly increase power over unselected sib-pairs.

Compared with the ED-only selection scheme, the mixture selection scheme not only makes better use of the screening process it is also more cost efficient. Considerable savings in cost are seen under recessive and dominant modes of inheritance. Residual correlation between sib-pairs plays a key role in the optimal design of selected samples. At higher degrees of correlation (perhaps larger than 0.3-0.4) discordant pairs become increasingly difficult to obtain. Therefore, the threshold for ED pairs should be accordingly relaxed. Conversely, the threshold for concordant pairs may be more stringent. The results shown in Tables 3, 4 and 5 provide some useful guidelines when the cost ratio (R) of phenotype-to-genotype is known. More specific guidelines are possible when there is knowledge of residual correlation.

We have assumed that the trait locus and marker locus are in complete linkage, but this is not an unrealistic assumption as more and more genetic markers are available for many organisms. Also, we have set a conservative type I error probability of $\alpha = 0.001$, as discussed Lander and Kruglyak [12], to more closely resemble a "search" for trait loci, whether by a scan or a relatively large panel of candidate genes. An error rate of $\alpha = 0.01$ or $\alpha = 0.0001$ gives the same basic patterns in optimal designs as discussed in the text. The IBD mean test of Risch and Zhang and more general methods as developed by Dudoit and Speed are needed to more faithfully reflect the reality of genetic epidemiology studies. Analyzing genotypes conditional on phenotypes provides a realistic framework under which to study genetic traits. The specific assumptions underlying the Dudoit-Speed score test allow one to evaluate the appropriate use of the method in any given situation. This is an important step when assessing the validity of study results, especially in observational studies.

Dedication

This paper is dedicated to Terry Speed. As thesis advisor he provided unending support and motivation; learning from him was truly inspirational. At this particular time in the development of statistical methods for statistical genetics and bioinformatics we are indeed fortunate to have Terry play a major role in shaping the field. As an occasional lone voice in the desert, he reminds us that we are trying to solve real problems that more often than not require thinking outside the box. This is perhaps the most important thing I learned from him – solve the problem. I am privileged to have his attention as friend, colleague, and advisor. – *Rudy Guerra*

Rudy Guerra, Program in Biostatistics, Department of Statistics, Rice University, Houston rguerra@rice.edu

Jian Han, Biostatistics, PPD, Inc., Austin jian.han@austin.ppdi.com

References

[1] D. B. Allison, M. Heo, N. J. Schork, S-L Wong, and R. C. Elston. Extreme selection strategies in gene mapping studies of oligogenic traits do not always increase power. *Human Heredity*, 48:97–107, 1998.

[2] C. I. Amos and R. Guerra. Statistics in human genetics. In S. Kotz, D. Banks, and C. Read, editors, *Encyclopedia of Statistical Science, Update*, volume 3, pages 334–346. Wiley, New York, 1998.

[3] W. C. Blackwelder and R. C. Elston. A comparison of sib-pair linkage tests for disease suscpetibility loci. *Genetic Epidemiology*, 2:85–97, 1985.

[4] G. Carey and J. A. Williamson. Linkage analysis of quantitative trait: increased power by using selected samples. *American Journal of Human Genetics*, 49:786–796, 1991.

[5] S. Dudoit and T. P. Speed. A score test for linkage using identity by descent data from sibships. *Annals of Statistics*, 27:943–986, 1999.

[6] S. Dudoit and T. P. Speed. A score test for the linkage analysis of quantitative and qualitative traits based on identity by descent data from sib-pairs. *Biostatistics*, 1:1–26, 2000.

[7] L. Eaves and J. Meyer. Locating human quantitative trait loci: Guidelines for the selection of sibling pairs for genotyping. *Behavior Genetics*, 24:443–455, 1994.

[8] D. R. Goldstein, S. Dudoit, and T. P. Speed. Power and robustness of a score test for linkage analysis of quantitative traits using identity by descent data on sib pairs. *Genetic Epidemiology*, 20:415–431, 2001.

[9] C. Gu and D. C. Rao. Linkage strategy for detection of human quantitative-trait loci. ii optimization of study design based on extreme sib pairs and generalized relative risk ratios. *American Journal of Human Genetics*, 61:211–222, 1997.

[10] C. Gu, A. Todorov, and D. C. Rao. Combining extremely concordant sibpairs with extremely discordant sibpairs provides a cost effective way to linkage analysis of quantitative trait loci. *Genetic Epidemiology*, 13:513–533, 1996.

[11] J. K. Haseman and R. C. Elston. The investigation of linkage between a quantitative trait and a marker locus. *Behavior Genetics*, 2:3–19, 1972.

[12] E. Lander and L. Kruglyak. Genetic dissection of complex traits: Guidlines for interpreting and reporting linkage results. *Nature Genetics*, 11:241–247, 1995.

[13] N. Risch and H. Zhang. Extreme discordant sib pairs for mapping quantitative trait loci in humans. *Science*, 268:1584–1589, 1995.

[14] N. Risch and H. Zhang. Mapping quantitative trait loci with extreme discordant sib pairs: sampling considerations. *American Journal of Human Genetics*, 58:836–843, 1996.

[15] H. Zhao, H. Zhang, and J. Rotter. Cost-effective sib-pair designs in the mapping of quantitative-trait loci. *American Journal of Human Genetics*, 60:1211–1221, 1997.