Am. J. Hum. Genet. 61:211–222, 1997

A Linkage Strategy for Detection of Human Quantitative-Trait Loci. II. Optimization of Study Designs Based on Extreme Sib Pairs and Generalized Relative Risk Ratios

Chi Gu[1] and D. C. Rao[1,2]

[1]Division of Biostatistics and [2]Departments of Psychiatry and Genetics, Washington University School of Medicine, St. Louis

Summary

We are concerned here with practical issues in the application of extreme sib-pair (ESP) methods to quantitative traits. Two important factors—namely, the way extreme trait values are defined and the proportions in which different types of ESPs are pooled, in the analysis—are shown to determine the power and the cost effectiveness of a study design. We found that, in general, combining reasonable numbers of both extremely discordant and extremely concordant sib pairs that were available in the sample is more powerful and more cost effective than pursuing only a single type of ESP. We also found that dividing trait values with a less extreme threshold at one end or at both ends of the trait distribution leads to more cost-effective designs. The notion of generalized relative risk ratios (the λ method, as described in the first part of this series of two articles) is used to calculate the power and sample size for various choices of polychotomization of trait values and for the combination of different types of ESPs. A balance then can be struck among these choices, to attain an optimum design.

Introduction

In the detection of human quantitative-trait loci (QTLs), the selected sib-pair methods are known to be more powerful than random samples of sib pairs, especially when the heritability is low (Carey and Williamson 1991; Fulker et al. 1991; Eaves and Meyer 1994; Risch and Zhang 1995; Gu et al. 1996). It also has been shown that the power to detect linkage to QTLs is concentrated in three types of extreme sib pairs (ESPs)—those with extremely discordant (ED) trait values, those with extremely high–concordant (HC) trait values, and those with extremely low–concordant (LC) trait values (e.g., see the studies by Risch and Zhang [1995] and Gu et al. [1996]). Among these, ED sib pairs have the greatest power in the detection of linkage to QTLs, under most relevant genetic models, although ED sib pairs are hard to find in human linkage studies.

The availability problem of ED sib pairs and a recent work by Elston (1992) on the cost consideration of a two-stage procedure for genomewide mapping (also see Elston 1994) motivated our current work in design optimization of human QTL studies based on ESP methods. We proposed a test (called "the EDAC test," which is an extension of Blackwelder and Elston's [1985] t_2 test; see Gu et al. 1996) that combines ED pairs with extremely concordant (EC) sib pairs that are available in the same sampling pool from which the ED sib pairs are selected. The EDAC design offers a compromise between the powerfulness and the availability of ESPs, and, as compared with the study designs pursuing solely ED sib pairs, the EDAC test is more likely to be cost effective (Gu et al. 1996). A different type of compromise is to relax the extremeness of ESPs; namely, the method of dividing the trait values can be altered to make more ED sib pairs available in a certain population. But, how far can we relax the extremeness without compromising on power? And how does this strategy compare with the EDAC design? An optimization algorithm is much needed.

To utilize linkage information from different types of ESPs, we believe that an optimum design should answer the following questions: (1) Is it necessary to combine ED and EC sib pairs? (2) Will such a combination enhance the power? (3) What is the most cost-effective combination? At the outset, since the ED pairs are the most powerful, it seems appealing to include all the available ED pairs, for genotyping and linkage analysis, regardless of whether a solely ED-pair design or a combined EDAC design is used. However, Gu et al. (1996) showed that too few of either type of the sib pairs actually could reduce the power, when combined by use of the EDAC test. Depending on the model, the necessary sample size for a particular power could differ by hundreds (see tables 5 and 6 in Gu et al. 1996). This, plus the fact that random fluctuation also may result in too

Received June 5, 1996; accepted for publication April 17, 1997.

Address for correspondence and reprints: Dr. Chi Gu, Division of Biostatistics, Washington University School of Medicine, Box 8067, 660 South Euclid Avenue, St. Louis, MO 63110. E-mail: gc@wubios.wustl.edu

0002-9297/97/6101-0028$02.00

few ED sib pairs, makes the answering of question (2) listed above essential in the search for the optimum design. In the first article of this series (Gu and Rao 1997 [in this issue]), we developed the generalized relative risk ratio (λ) method for quantitative traits and showed that, without direct inference of the underlying model, the power and the necessary sample sizes for ESP tests may be calculated via λ_S (for sib pairs) and λ_O (for parent-offspring pairs), for various types of ESPs. Therefore, we can answer all of the three questions listed above, using estimates of the λ's to determine if a combined EDAC design is warranted. Furthermore, when the EDAC design is indicated, we may apply the method to all possible combinations of ED and EC sib pairs, to select the one that costs the least and/or is the most powerful (i.e., to optimize).

Concerning relaxation of the extremeness of the sib pairs, an optimum design should address the "best" way to define extreme phenotypes; namely, a pair of so-called optimum thresholds should be given to define extreme trait values. This also can be done by application of the λ method. Since a different polychotomy of the trait values should result in a different set of values of the λ's, its effect on power and/or on sample size can be estimated by use of the new values of the λ's (see equations [17] and [18] in Gu and Rao 1997). We then could find the optimum polychotomization by comparing all plausible ways of dividing the trait values.

In this article, we first will give a brief introduction to the concept of generalized λ's and will refer to our previous article (Gu and Rao 1997) for the detailed discussion and equations. We then demonstrate the effect of polychotomization, by calculating the necessary sample sizes for a power of 80% at the significance level (α) of .001, using Risch and Zhang's (1995) ED-only sib-pair (EDSP) test, over a few combinations of upper thresholds and lower thresholds, for extreme trait values. Among these designs, a lower threshold at 30% and an upper threshold at 5%, that is, (30%, 5%), is shown to be consistently more cost effective than the other thresholds, under the genetic models tested. To show the effect of combining different numbers of ED and EC pairs, using the EDAC test, we fix the thresholds at (50%, 5%) and compare the power of various combinations of ED and HC pairs. The optimization for the so-called best combination is depicted by a graph of power contours and lines representing costs (fig. 1). A form of the optimization algorithm is presented in detail, to illustrate application of the λ method. Examples are given at the end, by use of three different sampling methods, to describe the optimization of designs for a hypothesized QTL study, across a variety of underlying genetic models.

Methods and Results

As we discussed elsewhere (Gu et al. 1996; Gu and Rao 1997), the quantitative phenotype X derives from an additive effect of a biallelic major locus, a residual term e, and the overall phenotypic mean μ: $X = \mu + g + e$. The two alleles, A_1 (corresponding to higher risk) and A_2, at the major locus have frequencies of p and $q = 1 - p$, respectively, and g takes the values $-a$, d, and a for trait genotypes A_2A_2, A_1A_2, and A_1A_1, respectively. The residuals are allowed to be correlated (with correlation ρ) among relatives. The overall heritability is denoted by H. The trait values are divided into a certain number of intervals with specified probabilities. The generalized λ for a relative pair of type R, with trait values in the hth and lth intervals, was defined (for $\rho = 0$) as follows:

$$\lambda_R(h,l) = \frac{K_R(l|h)}{K(l)}, \quad (1)$$

where $K(l)$ is the probability that a randomly selected person has a trait value in the lth interval and where $K_R(l|h)$ is the probability that a person has a trait value in the lth interval given that the trait value of this person's type-R relative is in the hth interval. See our previous article (Gu and Rao 1997) for a definition of $\lambda_R(h,l)$ when $\rho > 0$. Under the assumption that the residual correlation ρ is the same for all relative pairs, we have shown how the generalized $\lambda_R(h,l)$ can be used to calculate the power or the necessary sample size of ESPs, to detect linkage to QTLs.

Two types of ESP methods are discussed in this article. One is Risch and Zhang's ESP test, which uses only one type of ESP (discordant or concordant) and will be referred to as the "ESP test" (e.g., "EDSP" refers to the ED-only sib-pair test). The other is our EDAC test (Gu et al. 1996), which combines both ED and EC (HC or LC or both) sib pairs. Equations for the calculation of power and sample sizes, by use of estimated values of the λ's, were presented in our previous article (Gu and Rao 1997). The notation used here is the same as that used in our previous article (Gu and Rao 1997); in particular, the subscript "R" is used to denote type-R relatives, with an "S" used for sib pairs and an "O" used for parent-offspring pairs.

Effect of Polychotomization on the ESP Test

In the previous article (Gu and Rao 1997), we noted, in an example, how the values of the λ's (hence, the power) depended on a trichotomization of trait values. We now systematically investigate how the division of trait values affects sample sizes of different types of ESPs, as well as the cost effectiveness of a study design. Only three types of ESPs, namely, ED, HC, and LC sib pairs, will be considered. The trait values will be divided into three intervals, by use of two thresholds, $T_l < T_h$, and the trait values of $X > T_h$ are considered to be extremely high and the values of $X \leq T_l$ to be extremely

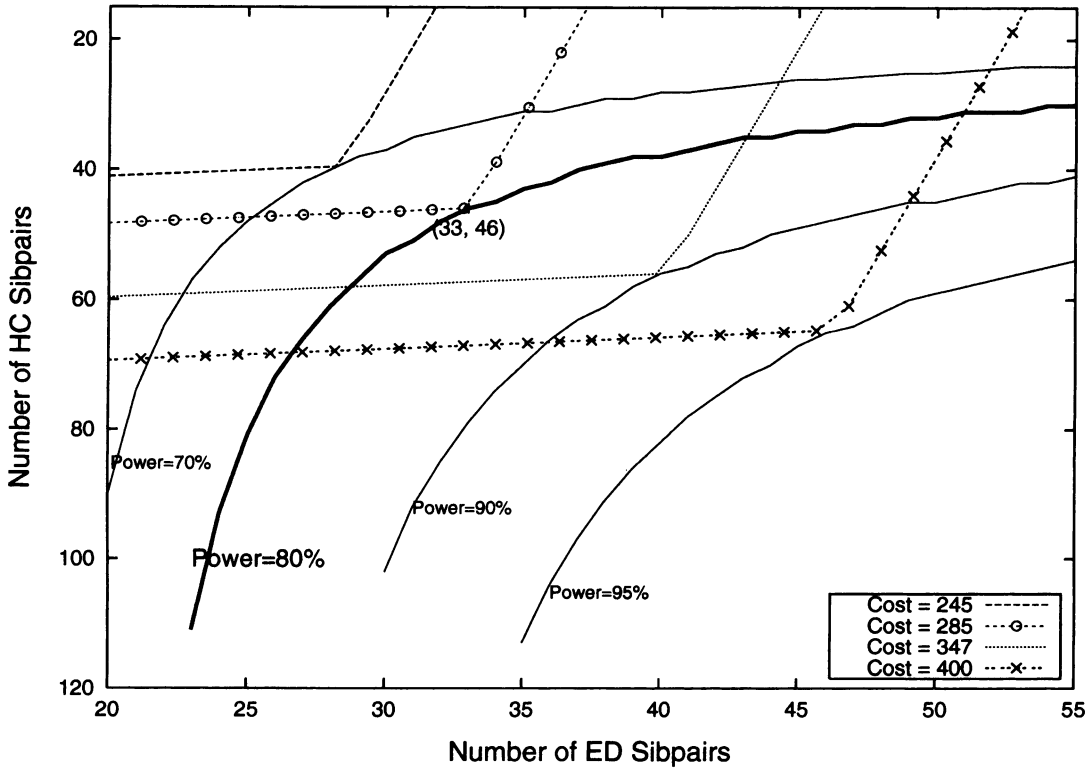

Figure 1 Optimization of the combination of HC and ED sib pairs, for the additive model with $p = .2$, $\rho = .4$, and $H = .3$. The power contours (solid lines) are plotted for 70%, 80%, 90%, and 95% power (at $\alpha = .001$), for various combinations of HC and ED pairs. The dashed lines denote the cost computed as $C = \max(N_{ed}, N_{hc}) + n_{ed} + n_{hc}$.

low. We refer to a trichotomy with $P(X \leq T_l) = x\%$ and $P(X > T_h) = y\%$ as an $(x\%, y\%)$ design. We use "h" and "l" to denote the higher and the lower intervals, respectively.

In general, by increasing the distance between the two thresholds, one increases the extremeness of the resulting ED sib pairs and, consequently, increases the power of a given number of ED pairs, but, naturally, the increased extremeness of the ED sib pairs makes them even harder to find. On the other hand, a decreased distance between the two thresholds will generate more ED sib pairs, but the reduced extremeness makes the same number of ED sib pairs less powerful. Will the increased number of available ED sib pairs make up for the power loss? Alternatively, with a fixed distance between the two thresholds, where is the best place to put them? For example, is the (20%, 10%) design better than the (10%, 20%) design? What will be the best design that will give the most throughput at a minimum cost? We will answer these questions later, under the subject of optimum designs. First, we will show here that a choice must be made among various divisions of trait values, because the effect on cost effectiveness could be dramatic.

In tables 1–3, we display, for a few additive, dominant, and recessive models ($p = .2$, $H = .3$, and $\rho = .4$), the sample sizes of ED sib pairs required for a power of 80%, at $\alpha = .001$, using different pairs of thresholds. The thresholds are placed at points so that the first and the last interval each has a probability ($P[l]$ and $P[h]$) of .05, .10, .20, or .30. That is, T_l will be varied among the 5th, 10th, 20th, or 30th percentiles,

Table 1

Sample Sizes for the EDSP Test Having 80% Power at $\alpha = .001$, under an Additive Model

	Sample Size for ED Pairs/Total Sampling Pool ($\lambda_S[h,l]/\lambda_O[h,l]$), for T_h =			
T_l	95%	90%	80%	70%
5%	19/4,628 (.40/.28)	27/4,621 (.45/.37)	46/5,166 (.54/.50)	74/6,178 (.60/.58)
10%	22/1,954 (.42/.32)	33/2,068 (.48/.42)	58/2,502 (.57/.54)	95/3,188 (.63/,62)
20%	28/847 (.46/.38)	43/968 (.52/.48)	80/1,299 (.61/.59)	138/1,797 (.68/.67)
30%	35/539 (.49/.43)	55/655 (.56/.52)	107/953 (.65/.64)	191/1,404 (.71/.71)

NOTE.—$p = .20$; $H = .30$; and $\rho = .40$.

Table 2

Sample Sizes for the EDSP Test Having 80% Power at α = .001, under a Dominant Model

	Sample Size for ED Pairs/Total Sampling Pool ($\lambda_S[h,l]/\lambda_O[h,l]$), for T_h =			
T_l	95%	90%	80%	70%
5%	36/8,706 (.55/.57)	41/6,652 (.56/.59)	53/5,651 (.59/.61)	72/5,840 (.62/.65)
10%	39/3,269 (.56/.58)	45/2,712 (.57/.60)	62/2,565 (.61/.63)	88/2,874 (.65/.67)
20%	45/1,253 (.57/.60)	54/1,154 (.59/.62)	79/1,248 (.64/.66)	121/1,545 (.68/.70)
30%	52/744 (.59/.61)	66/742 (.61/.64)	102/886 (.66/.68)	163/1,180 (.71/.73)

NOTE.—p = .20; H = .30; and ρ = .40.

and T_h will be varied among the 70th, 80th, 90th, or 95th percentiles. For each combination, values of $\lambda_S(h,l)$ and $\lambda_O(h,l)$ follow the necessary ED-pair sample size and the expected total number of sib pairs that need to be screened (i.e., phenotyped) to obtain the number of ED sib pairs, which we will refer to as the "total sampling-pool size," in this article (see equation [2], presented later, and also equations [3]–[6] in Gu et al. 1996). We assumed, for these calculations, that selective sampling from the upper tail of the trait distribution was used; that is, probands were sampled from a subpopulation with extremely high trait values, and their siblings were screened, to form ED sib pairs. Although in some studies large sampling pools may have been already phenotyped prior to the design of the linkage analysis, one cannot always take that for granted.

Under the additive model, for a fixed T_h, when T_l is relaxed from the 5th percentile to the 30th percentile, the ED-pair sample size is roughly doubled, and the total sampling-pool size typically is reduced by 5–7-fold; for a fixed T_l, when T_h is relaxed from the 95th percentile to the 70th percentile, the ED-pair sample size is more than quadrupled, and the total sampling-pool size also is increased by 1.5–2-fold. Under the dominant model, relaxation of T_l, as described above, increases the ED-pair sample size by 1.5–2-fold and reduces the total sampling-pool size by 5–10-fold; relaxation of T_h increases the ED-pair sample size by 2–3-fold and may reduce the total sampling-pool size when $P(l)$ is small (.5 and .10) or may increase it when $P(l)$ is large (.20 and .30). As for the recessive model, the total sampling-pool size is reduced by 3–6-fold and the ED-pair sample size is increased from just slightly to ~3-fold, when T_l is relaxed as described above; relaxation of T_h, on the other hand, increases the total sampling-pool size by 15–35-fold while increasing the ED-pair sample size by 12–40 fold.

Among these cases, the (30%, 5%) design appears to be the most cost-effective way to divide the trait values, for an EDSP design. Therefore, it seems that probands with extremely high trait values and their siblings with not-so-extremely low trait values constitute a better design. We also did the same calculations with ρ = 0 and reached the same conclusions.

Of course, the conclusions depend on the risk-allele frequency (p), as well as the sampling method. A much larger value of p and/or a different way to select the ESPs may yield different results (see the examples in the following sections).

Effect of the Combination of ED and EC Sib Pairs, in the EDAC Test

As an alternative study design, Gu et al. (1996) proposed to utilize the EC sib pairs available in the same sampling pool with the ED sib pairs and provided a combined EDAC test. This type of study design deals with the availability problem of ED sib pairs, from a different perspective; namely, without compromising on the extremeness (i.e., powerfulness) of the ED pairs, one also utilizes the linkage information in the less but still fairly powerful EC sib pairs, which are readily available. We showed that, when a reasonable number of EC sib pairs are combined with ED sib pairs, the EDAC design is more cost effective than the EDSP design. We also noted that there were so-called power dips in the EDAC test, which were caused by too few ED sib pairs or too few EC sib pairs being available (see figs. 1–4 in Gu et al. 1996).

Here, we consider the version of the EDAC test that combines only HC sib pairs with ED sib pairs, so that the power of the EDAC test can be listed in a two-way table. In reality, the combination of LC or of both HC and LC sib pairs could be more cost effective, which is discussed later.

Table 3

Sample Sizes for the EDSP Test Having 80% Power at α = .001, under a Recessive Model

	Sample Size for ED Pairs/Total Sampling Pool ($\lambda_S[h,l]/\lambda_O[h,l]$), for T_h =			
T_l	95%	90%	80%	70%
5%	52/2,882 (.67/.84)	79/5,520 (.71/.86)	233/17,950 (.81/.90)	670/46,524 (.88/.94)
10%	54/1,375 (.68/.84)	94/2,957 (.73/.87)	333/10,765 (.83/.92)	1,017/29,169 (.90/.95)
20%	58/668 (.69/.84	123/1,663 (.76/.88)	524/6,820 (.86/.93)	1,677/19,234 (.92/.96)
30%	62/444 (.69/.85)	154/1,230 (.77/.89)	728/5,418 (.88/.94)	2,370/15,606 (.93/.96)

NOTE.—p = .20; H = .30; and ρ = .40.

Table 4

Power of Various Combinations of the Numbers of ED and HC Sib Pairs, by Use of the EDAC Test, under an Additive Model

	POWER, FOR n_{ed} = [b]								
n_{hc} [a]	7 (.03)	14 (.09)	21 (.20)	28 (.34)	35 (.47)	42 (.60)	49 (.71)	56 (.80)	63 (.86)
13 (.03)	.09	.19	.26	.31	.34	.37	.39	.40	.41
26 (.10)	.13	.32	.46	.56	.63	.68	.72	.74	.77
39 (.20)	.15	.38	.57	.69	.77	.82	.86	.89	.91
52 (.32)	.16	.42	.63	.76	.84	.89	.92	.94	.96
65 (.44)	.16	.45	.67	.81	.89	.93	.95	.97	.98
78 (.56)	.17	.47	.70	.84	.91	.95	.97	.98	.99
91 (.66)	.17	.49	.72	.86	.93	.96	.98	.99	.99
104 (.75)	.17	.50	.74	.87	.94	.97	.98	.99	1.00
117 (.82)	.18	.51	.75	.88	.95	.98	.99	.99	1.00

NOTE.—p = .20; H =.30; ρ = .40; and (T_h, T_l) = (50%, 5%).

[a] The nos. in parentheses indicate the power when only HC sib pairs are used.

[b] The nos. in parentheses indicate the power when only ED sib pairs are used.

For the same set of models for which we analyzed, in the previous section, the effects of dividing the trait values, we now fix the (50%, 5%) design and calculate the numbers of ED and HC sib pairs, N_{ed} and N_{hc}, respectively, that are necessary for a nominal power of 80%, when the ESP test is based solely on ED or HC sib pairs. If the number of available ED pairs in a sample is larger than N_{ed} or if that of HC pairs is larger than N_{hc}, then a single-type ESP test will have the necessary power, and the combination of the sib pairs for EDAC is not necessary (although it could yield more power). So, we will focus on the EDAC test that combines the number of ED pairs that is less than N_{ed} and the number of HC pairs that is less than N_{hc}. For the purpose of demonstration, we increase the sample sizes of ED and HC sib pairs in nine equal increments, until they just surpass N_{ed} and N_{hc}, respectively. The power of the EDAC test that combines these numbers of ED and HC pairs is displayed in tables 4–6.

In tables 4–6, we also display, in parentheses after the various numbers of ED or HC sib pairs, the power of the ESP test based on the ED pairs only or on the HC pairs only, so that one can compare the power of these tests with the power of the combined EDAC test, to see which combinations enhance the power. We see that use of a combination of sib pairs will result in a higher power than use of either type of sib pairs alone, as long as enough numbers of both types of pairs are combined. For example, with the additive trait displayed (table 4), if the number of ED sib pairs is <21, pooling of ED and HC sib pairs will never yield a power >75%, regardless of the number of HC pairs (i.e., even when use of the HC pairs alone could have a power >80%). On the other hand, use of a combination of 21 ED pairs and 39 HC pairs has 57% power, which is much better than that for use of the ED pairs only (20%) or the HC pairs only (20%), and use of the combination of 21 ED pairs and 65 HC pairs has 67% power, which is a big improvement over that for use of the ED pairs only (20%) or the HC pairs only (44%). Interested readers may refer to the study by Gu et al. (1996) for the results for a broader range of models. Note that the combinations near the diagonal in the tables always have enhanced power, over the ESP test based on either type of sib pair alone. Also, when approximately one-half of N_{ed} and one-half of N_{hc} are combined, the power of the combined EDAC will surpass the nominal 80% power, and the gain in power, over the ESP test, is nearly maximized (also see fig. 1 and the examples described below, for optimum combinations).

Effect of Polychotomization, on the EDAC Test

We also calculated, for the same set of models with the same set of trait-value divisions, the sample sizes of ED and EC sib pairs, for the EDAC design, when the ED and HC pairs are combined according to their *expected ratio* of availability in a selective sample. The ratio can be estimated via the use of recurrence risks derived from reliable population studies (see equation [3] below).

The patterns of effects are similar to those of the EDSP test (tables 1–3). However, the sampling-pool size for the EDAC design is reduced further for the same trait-value divisions, whereas the combined sample size (ED plus HC pairs) is not always reduced. We present the results only for the additive trait, in table 7. We see that the (30%, 5%) EDAC design will reduce the screening (i.e., phenotyping) burden (and hence the cost), as well as the combined-sample size (for genotyping) by ~4-fold, as compared with the (10%, 10%) EDAC design.

Table 5

Power of Various Combinations of the Numbers of ED and HC Sib Pairs, by Use of the EDAC Test, under a Dominant Model

	POWER, FOR n_{ed} = [b]								
n_{hc} [a]	9 (.02)	18 (.08)	27 (.17)	36 (.28)	45 (.41)	54 (.53)	63 (.65)	72 (.74)	81 (.81)
25 (.03)	.08	.18	.26	.33	.37	.41	.44	.46	.48
50 (.10)	.10	.27	.42	.54	.62	.68	.73	.77	.79
75 (.20)	.10	.31	.50	.64	.74	.81	.85	.89	.91
100 (.32)	.11	.34	.55	.70	.80	.87	.91	.94	.95
125 (.44)	.11	.35	.58	.74	.84	.90	.94	.96	.97
150 (.55)	.11	.37	.60	.76	.86	.92	.95	.97	.98
175 (.66)	.12	.37	.62	.78	.88	.93	.96	.98	.99
200 (.74)	.12	.38	.63	.80	.89	.94	.97	.98	.99
225 (.81)	.12	.39	.64	.81	.90	.95	.98	.99	.99

NOTE.—p = .20; H = .30; ρ = .40; and (T_h, T_l) = (50%, 5%).
[a] The nos. in parentheses indicate the power when only HC sib pairs are used.
[b] The nos. in parentheses indicate the power when only ED sib pairs are used.

Also, compared with the (30%, 5%) EDSP design (table 1), this hybrid design reduces the total sampling-pool size by 253 (almost 2-fold), while it increases the sample size for genotyping by 47. Therefore, both the relaxation of extremeness and the utilization of EC sib pairs contribute to a more efficient and more cost-effective design.

Optimization of Study Designs, Using the λ Method

In the preceding discussion, we observed the following: (1) An EDAC design is likely to be more cost effective than a single-type ESP design, when enough numbers of both ED and EC sib pairs are available in the sample, and some combinations are better than others. (2) For either design, some trichotomies of trait values are more cost effective than others. Therefore, a systematic algorithm is needed, to search for the optimum design.

Of course, without any knowledge about the disease etiology, optimization of study designs would not be possible. Fortunately, as we demonstrated in the first article of this series (Gu and Rao 1997), if estimates of various recurrence risks (and, hence, estimates of the λ's) are available, one may calculate statistical power and necessary sample sizes, as well as the cost of different designs, without knowing the parameters of the underlying genetic model. Equations for sample size and power were given in the previous article (Gu and Rao 1997), and the total sampling-pool size may be calculated via the estimated recurrence risk $K_S(h,l)$ as follows (also see equations [3]–[6] in Gu et al. 1996):

Table 6

Power of Various Combinations of the Numbers of ED and HC Sib Pairs, by Use of the EDAC Test, under a Recessive Model

	POWER, FOR n_{ed} = [b]								
n_{hc} [a]	9 (.01)	18 (.06)	27 (.15)	36 (.28)	45 (.43)	54 (.57)	63 (.70)	72 (.80)	81 (.87)
3 (.02)	.07	.09	.10	.11	.11	.11	.11	.11	.12
6 (.08)	.17	.26	.31	.33	.35	.36	.37	.38	.38
9 (.18)	.25	.43	.51	.56	.58	.61	.62	.63	.64
12 (.30)	.32	.56	.66	.72	.75	.78	.79	.80	.81
15 (.43)	.38	.65	.77	.83	.86	.88	.89	.90	.91
18 (.55)	.42	.73	.84	.89	.92	.94	.95	.95	.96
21 (.66)	.46	.78	.89	.93	.95	.97	.97	.98	.98
24 (.76)	.49	.82	.92	.96	.97	.98	.99	.99	.99
27 (.83)	.51	.85	.94	.97	.98	.99	.99	1.00	1.00

NOTE.—p = .20; H = .30; ρ = .40; and (T_h, T_l) = (50%, 5%).
[a] The nos. in parentheses indicate the power when only HC sib pairs are used.
[b] The nos. in parentheses indicate the power when only ED sib pairs are used.

Table 7

Sample Sizes of Ed and HC Sib Pairs for the EDAC Test Having 80% Power at α = .001, under an Additive Model

	Sample Size for ED Pairs/HC Pairs/Total Sampling Pool ($\lambda_S[h,h]/\lambda_O[h,h]$), for T_h =[a]			
T_l	95%	90%	80%	70%
5%	9/520/2,328 (1.89/1.81)	15/746/2,482 (1.59/1.56)	27/1,239/3,023 (1.33/1.33)	46/1,918/3,868 (1.21/1.21)
10%	11/217/971 (1.89/1.81)	17/329/1,094 (1.59/1.56)	33/589/1,437 (1.33/1.33)	58/969/1,955 (1.21/1.21)
20%	14/96/429 (1.89/1.81)	23/155/517 (1.59/1.56)	46/305/745 (1.33/1.33)	84/543/1,094 (1.21/1.21)
30%	18/64/286 (1.89/1.81)	30/108/361 (1.59/1.56)	62/227/555 (1.33/1.33)	117/426/860 (1.21/1.21)

Note.—p = .20; H = .30; and ρ = .40.
[a] The values for the λ's refer to the HC pairs.

$$P[(h,l)_S] = K_S(h,l)P(h) \quad (2)$$

and

$$N = P(h)\frac{N_{ed}}{P[(h,l)_S]}, \quad (3)$$

where $P[(h,l)_S]$ is the probability of a sib pair having trait outcome (h,l), $P(h)$ is the probability of a randomly selected person having the trait value in the hth interval, and N_{ed} is the ED sample size required for the preset power.

Applying this theory, we now can develop an optimum design by answering the following questions: (1) Is the combination of different types of sib pairs going to increase power? (2) Is relaxation of extremeness going to improve cost effectiveness? (3) What is the best combination of ED and EC sib pairs, and what is the best way to divide the trait values? We will describe below an algorithm that answers these questions in steps and that eventually will lead us to an optimum study design.

We will deal with only those instances in which phenotypic data already have been collected for a certain number of families or sib pairs and for which the task is to determine the numbers of the different types of ESPs to be used for genotyping and linkage analysis. We may need to minimize the cost, to achieve a desired power 1 − β, at some α, or to achieve maximum power at a given cost. When a new study is designed in the absence of any phenotypic data, attention must be paid to sampling and to balancing of the costs of phenotyping and genotyping. In the algorithm presented below, (n_0, n_1, n_2) denotes a combination with n_0 LC pairs, n_1 ED pairs, and n_2 HC pairs; for example, a design using n_1 ED pairs only is denoted by (0, n_1, 0).

Algorithm for Optimization

The algorithm to be used with pre-existing phenotypic data is as follows.

1. For each plausible division of trait values (or, at least for the more promising divisions), get estimated values of the generalized λ's for sib pairs and for parent-offspring pairs, from the existing data set or from previous population studies.
2. Then, for each division, calculate the necessary sample sizes, S_0, S_1, and S_2, of LC, ED, and HC sib pairs, respectively, to achieve the desired power 1 − β, if the ESP-test design is to be used for analysis.
3. Let N_0, N_1, and N_2 be the numbers of available LC, ED, and HC sib pairs in the sample. If $N_i \geq S_i$ for any i, calculate the cost for use of only that type of sib pairs in the analysis. Retain the number with the least cost.
4. For each combination (n_0, n_1, n_2), when $n_i \leq N_i$, calculate the power of the combined EDAC test.
5. If the power is less than desired, ignore that combination; if the power is not less than desired, compare its cost, $(n_0 + n_1 + n_2)\,C_G$, where C_G is the genotyping cost of one sib pair, with that of the previously retained combination. If it costs less, retain this combination; otherwise, reject it.
6. Consider the next combination.
7. After all possible combinations have been exhausted, the retained combination attains the desired power with the least cost.

When designing a new study in the absence of pheno-

typic data, the cost function discussed in step (5) above takes the form $NC_P + (n_0 + n_1 + n_2)\ C_G$, where C_P is the unit phenotyping cost and where N is the total sampling-pool size. Then, minimization of the cost is equivalent to minimization of $N + (n_0 + n_1 + n_2)\ C_G/C_P$. Thus, the genotyping-to-phenotyping cost ratio is needed in the algorithm, and the sampling method will determine the value of N. In reality, the algorithm should be tailored to suit the actual situation. For example, if partial phenotypic data that yield less power than desired already have been collected, an optimum design then would need to maximize the power of the already-collected data and to minimize the cost of the collection of additional data.

Example

We now illustrate the application of the above algorithm to a hypothetical study of a moderately heritable trait, in which genotyping and phenotypic screening cost the same per subject. Assuming the underlying trait to be additive, dominant, or recessive, with various values of p, we fix $H = .3$. For each of the models, we require that the optimum design retain a power of 80% at $\alpha = .001$. The following three sampling methods were considered: (1) selecting probands from the high-extreme tail of the trait distribution and screening their siblings, to form HC and ED pairs; (2) selecting probands from both extreme tails and the screening their siblings, to form HC, LC, and ED pairs; and (3) randomly selecting sib pairs, to screen for ESPs. Methods (1) and (2) are referred to as "selective sampling," and method (3) is referred to as "random sampling," in this article.

The set of thresholds that trichotomize the trait values for the optimum design is searched for over a grid of possible percentiles spaced 5% apart. In theory, one can search for the thresholds continuously over all possible values, but that would require too many estimates of the corresponding λ's, as well as unrealistic computer time. For each selected set of thresholds, the best combination of the various types of ESPs required for genotyping and for phenotypic screening is derived by comparison of the six possible ESP designs (depending on the sampling scheme), namely, the ED, the HC, the LC, the ED + HC, the ED + LC, and the ED + HC + LC designs. The cost of a design is measured by the sum of the number of unselected sib pairs needed for screening and the number of ESPs to be genotyped. The optimum design will minimize the cost for a given model.

A graphic illustration for the optimization of the combination of HC and ED pairs (by use of selective sampling from the upper-extreme tail), for the thresholds (50%, 5%), under an additive model with $p = .2$, $\rho = .4$, and $H = .3$, is given in figure 1. Contours consisting of ED + HC combinations with 70%, 80%, 90%, and 95% power are plotted as solid curves. Cost lines ($C = \max[N_{ed}, N_{hc}] + n_{ed} + n_{hc}$) are plotted as various types of dashed lines. When the cost is fixed, maximization of power is equivalent to the search for the power contour that intersects the cost line at just one point. Similarly, when the power is preset, minimization of cost is equivalent to the search for the cost line that intersects the power contour at just one point. The value C represented by the cost line will give the minimum cost, and the point of intersection will be the optimum combination (e.g., the combination of 33 ED and 46 HC pairs is optimum for the preset power of 80%).

We summarize the optimization results in tables 8–10, under the categories of additive, dominant, and recessive models and for various values of p. We want to clarify here that, although the results are presented this way, the optimization process does not require prior knowledge of the underlying models per se. All it needs are sets of estimated values of the λ's corresponding to the sets of thresholds over which one wishes to optimize the study design.

As a general pattern, we see that, in far more cases, the EDAC design combining different types of ESPs works more efficiently than those designs utilizing only one type of ESP. This is true more so when the random-sampling scheme is used for selection of the ESPs, for which all the optimum designs correspond to the EDAC design. Only when selective sampling from the high end of the distribution is used does the EDSP design become the choice for the optimum design, for dominant traits and for recessive traits with very high values of p.

A smaller allele frequency, as well as the presence of positive residual correlation among sibs, requires optimum designs to choose more liberal thresholds. For random sampling, it results in less-restricted thresholds on both ends, whereas, for selective sampling (either for one tail or for both tails), it results in a restricted threshold at one end and a somewhat liberal threshold at the other end. Most often, $\rho > 0$ reduces the cost of the optimum design, because EC pairs are more readily available and ED pairs become more powerful.

For additive traits, when $p < .5$, EDAC designs combining ED and HC pairs, together with liberal lower thresholds, are optimum. As the frequency increases and becomes $>.5$, if ESPs are sampled randomly or selected from both ends of the distribution, the results for p correspond to that of $1 - p$, with the positions of the HC and LC pairs switched; so, these results are not displayed in tables 8 and 9. For those cases, the optimum design combines LC pairs with ED pairs by use of less-restrictive *upper* thresholds. If we select ESPs from only the upper tail of the trait distribution, the optimum design turns out to be the EDSP design. The lower threshold depends on the magnitude of ρ.

For dominant traits, the patterns of extreme thresh-

Table 8

Optimization of Study Designs under Various Additive and Dominant Models (H = .30), with the Assumptions of Random Sampling and Equal Unit Cost for Genotyping and Phenotyping

	Study Design, for ρ = 0			Study Design, for ρ = .40		
p	Thresholds[a] (%, %)	(n_0, n_1, n_2)[b]	Cost[c]	Thresholds[a] (%, %)	(n_0, n_1, n_2)[b]	Cost[c]
			Additive Model			
.05	(50, 10)	(0, 76, 41)	1,468	(60, 10)	(0, 67, 41)	1,292
.10	(45, 15)	(0, 128, 77)	1,831	(45, 15)	(0, 80, 79)	1,567
.20	(30, 20)	(0, 129, 133)	2,198	(35, 20)	(0, 92, 131)	1,837
.30	(30, 20)	(0, 136, 141)	2,364	(35, 20)	(0, 96, 139)	1,952
.40	(30, 25)	(126, 169, 151)	2,369	(35, 20)	(180, 92, 127)	2,020
.50	(25, 25)	(169, 121, 169)	2,248	(25, 25)	(193, 78, 97)	2,098
			Dominant Model			
.05	(55, 10)	(0,79,38)	1,356	(60, 10)	(0, 61, 39)	1,205
.10	(45, 15)	(0, 118, 70)	1,680	(50, 15)	(0, 88, 71)	1,430
.20	(30, 25)	(0, 157, 169)	2,095	(35, 25)	(0, 117, 162)	1,749
.30	(30, 30)	(199, 176, 105)	2,049	(30, 25)	(221, 92, 94)	1,923
.40	(25, 30)	(169, 163, 0)	2,121	(25, 40)	(163, 147, 0)	1,781
.50	(20, 40)	(109, 165, 0)	1,900	(20, 45)	(109, 127, 0)	1,635
.60	(15, 50)	(64, 139, 0)	1,637	(15, 50)	(67, 96, 0)	1,435
.70	(10, 60)	(32, 96, 0)	1,339	(10, 55)	(35, 63, 0)	1,243
.80	(5, 60)	(14, 43, 0)	1,158	(5, 55)	(15, 32, 0)	1,136
.90	(5, 55)	(61, 321, 0)	8,888	(5, 50)	(101, 205, 0)	10,082
.950	(5, 55)	(649, 4387, 0)	119,566	(5, 40)	(1409, 1909, 0)	150,589

NOTE.—The results for additive models with $p > .5$ and those for recessive models are not displayed, since they can be derived from the corresponding results for additive or dominant models (see the Example section in Methods and Results).

[a] (x%, y%) defines the trait values below the xth percentile as extremely low and those above the (100 − y)th percentile as extremely high.

[b] Also indicates which type of ESP test (EDSP or EDAC) was used for the optimum design, since n_0 denotes the no. of LC pairs, n_1 the no. of ED pairs, and n_2 the no. of HC pairs used.

[c] Measured by the sum of the no. of sib pairs to be screened and the no. to be genotyped.

olds in optimum designs are similar to those of the additive case, except that the optimum designs switch from HC + ED to LC + ED earlier than those of the additive case, as p increases. Again, when selecting from the upper tail only, EDSP designs become the choice for the optimum design, as p increases.

For recessive traits, if ESPs are sampled randomly or selected from both tails of the distribution, the results for p are symmetric to those of $1 - p$, for the dominant trait with HC and LC switched. So, we skip displaying the results for these cases. When ESPs are selected only from the upper tail of the distribution, HC designs become favorable when p is small and when ρ = 0.

Finally, it should be noted that this example was limited to the case when the unit costs are the same for genotyping and for phenotyping. In reality, the phenotyping costs can be much larger. However, phenotypic studies often collect information on a number of phenotypes, and it may be desirable to perform linkage analysis for several (if not all) of the phenotypes. Looked at in this light, the unit cost per phenotype may not be grossly different from the unit cost of genotyping. Therefore, the results presented in this example may have a greater degree of validity than may appear at first.

Discussion

When the method of ESPs is applied to quantitative traits, the setting of the thresholds to define extreme trait values and the choice of the right combination of ED and EC sib pairs to pool poses an important practical challenge. We have demonstrated here how these two factors affect the cost effectiveness of an ESP-study design and how an optimum choice can be achieved via use of the λ method. The overwhelming power of and the relative lack of a sufficient number of ESPs generate a dilemma for any investigator searching for the best study design. The following three types of compromises were suggested here: (1) relax the criterion of extremeness so that more ED sib pairs will be available in

Table 9

Optimization of Study Designs under Various Additive and Dominant Models (H = .30), with the Assumptions of Selective Sampling, from both the Upper and Lower Tails, and Equal Unit Cost for Genotyping and Phenotyping

	Study Design, for ρ = 0.0			Study Design, for ρ = 0.40		
p	Thresholds (%, %)	(n_0, n_1, n_2)	Cost	Thresholds (%, %)	(n_0, n_1, n_2)	Cost
	Additive Model					
.05	(55, 5)	(0, 38, 25)	183	(60, 5)	(0, 23, 28)	148
.10	(40, 5)	(0, 47, 28)	266	(60, 5)	(0, 32, 35)	210
.20	(25, 5)	(0, 44, 34)	371	(50, 5)	(0, 33, 46)	285
.30	(20, 5)	(0, 48, 40)	489	(50, 5)	(0, 42, 55)	355
.40	(20, 5)	(0, 62, 46)	627	(45, 5)	(0, 44, 65)	434
.50	(5, 15)	(54, 61, 0)	795	(5, 45)	(79, 54, 0)	525
	Dominant Model					
.05	(55, 5)	(0, 37, 24)	178	(70, 5)	(0, 28, 26)	144
.10	(35, 5)	(0, 55, 31)	338	(55, 5)	(0, 34, 39)	242
.20	(35, 5)	(0, 130, 0)	669	(50, 5)	(0, 58, 67)	458
.30	(5, 25)	(0, 132, 0)	887	(5, 45)	(97, 74, 0)	671
.40	(10, 25)	(77, 92, 0)	688	(5, 50)	(67, 61, 0)	466
.50	(10, 35)	(57, 86, 0)	479	(5, 55)	(46, 49, 0)	320
.60	(5, 40)	(25, 57, 0)	287	(5, 60)	(31, 38, 0)	212
.70	(5, 55)	(18, 39, 0)	160	(5, 65)	(21, 28, 0)	136
.80	(5, 5)	(21, 0, 0)	105	(5, 60)	(19, 23, 0)	111
.90	(5, 5)	(40, 0, 0)	425	(5, 55)	(130, 155, 0)	915
.950	(5, 5)	(241, 0, 0)	4090	(5, 55)	(1,805, 2,185, 0)	13,403

NOTE.—See footnotes to table 8.

the sampling pool; (2) combine EC sib pairs with ED sib pairs, to enhance the power by use of the EDAC design; and (3) both of the above choices. By applying the λ method, we have shown that it is possible to strike a balance between these choices and to arrive at an optimum design. It should be noted that "optimization" of study designs is always relative in practice; namely, the result may be only *relatively* optimum, given all the available data and prior information. For instance, in the example discussed, we searched on a 19 × 19 grid of percentiles to identify the optimum trait thresholds. This search requires 361 sets of estimates of the corresponding λ's, which are hardly available in reality. What is more practical is to search for the optimum design over a much smaller number of thresholds (e.g., 5%, 10%, 20%, and 30%).

From our investigation, it appears that the use of an EDAC design and the selection of siblings from a broader interval at the lower end and from a restricted range at the higher end provide a more cost-effective design. Of course, a much larger p will reverse the situation, as we have seen in the example. When ESPs are selected from randomly sampled sib pairs, cost is noticeably larger, and EC pairs from both ends of the distribution should be collected for analysis when p is close to .5. When ESPs are sampled from the tails of the distribution, cost is reduced substantially. Which type of EC pairs should be combined depends on p. However, as we have shown, as long as one has estimates of the λ's, one can search for the optimum design without knowing the actual frequency. Finally, as discussed by Zhang and Risch (1996), the parental phenotypic status also is relevant when dealing with EC sib pairs, which is an issue not discussed here.

The sampling scheme is another important factor affecting the choice of optimum study designs and certainly deserves a separate investigation. We simply point out here that selection of ESPs from extremely high-risk or from extremely low-risk populations likely would result in better designs and that application of the λ method enables us to decide which selective-sampling scheme may be more suitable for a study.

We should clarify two points here. First, although the proposed optimization may involve all three types of extreme sibpairs (LC, ED, and HC), it does not require the genotyping of all the sib pairs available in the sample. On the contrary, what the optimization achieves is either

Table 10

Optimization of Study Designs under Various Additive, Dominant, and Recessive Models (H = .30), with the Assumptions of Selective Sampling, from the Upper Tail, and Equal Unit Cost for Genotyping and Phenotyping

	Study Design, for ρ = 0.0			Study Design, for ρ = 0.40		
p	Thresholds (%, %)	(n_0, n_1, n_2)	Cost	Thresholds (%, %)	(n_0, n_1, n_2)	Cost
			Additive Model			
.05	(55, 5)	(0, 38, 25)	183	(60, 5)	(0, 23, 28)	148
.20	(25, 5)	(0, 44, 34)	371	(50, 5)	(0, 33, 46)	285
.40	(20, 5)	(0, 62, 46)	627	(45, 5)	(0, 44, 65)	434
.60	(25, 5)	(0, 140, 0)	971	(40, 5)	(0, 55, 97)	643
.80	(15, 5)	(0, 122, 0)	1,350	(35, 5)	(0, 93, 0)	1,078
			Dominant Model			
.05	(55, 5)	(0, 37, 24)	178	(70, 5)	(0, 28, 26)	144
.20	(35, 5)	(0, 130, 0)	669	(50, 5)	(0, 58, 67)	458
.40	(20, 5)	(0, 152, 0)	1,203	(40, 5)	(0, 112, 0)	983
.60	(10, 5)	(0, 132, 0)	1,939	(25, 5)	(0, 104, 0)	1,617
.80	(5, 5)	(0, 144, 0)	3,885	(5, 25)	(0, 61, 0)	2,724
			Recessive Model			
.05	(5, 5)	(0, 0, 241)	4,090	(55, 5)	(0, 2,185, 1,805)	13,403
.20	(5, 5)	(0, 0, 21)	105	(60, 5)	(0, 23, 19)	111
.40	(40, 5)	(0, 57, 25)	287	(60, 5)	(0, 38, 31)	212
.60	(25, 10)	(0, 92, 77)	688	(50, 5)	(0, 61, 67)	466
.80	(20, 5)	(0, 143, 0)	1,167	(40, 5)	(0, 105, 0)	950

NOTE.—See footnotes a, b, and c in table 8.

to minimize the genotyping of sib pairs as much as possible, for a preset power, or to determine the numbers of various types of sib pairs to be genotyped, so as to attain as high a power as possible. Second, the use of neither HC pairs nor ED pairs should be regarded as the traditional affected-affected or affected-unaffected design for dichotomous traits. The extremeness resulted from a trichotomy of the trait values that enhances power in ways that a simple affected/unaffected dichotomy of the trait cannot.

When a genomewide search using a dense map is performed, estimation of the expected identity-by-descent (IBD) proportion is possible from the λ's (Gu and Rao 1997). Therefore, the placement of markers also may be incorporated into the optimization process, to accommodate a genomewide mapping project.

We note that both the ESP test and the EDAC test deal with only ESPs. Some linkage information is discarded by the trichotomization and by the throwing away of sib pairs falling in the middle interval. It will be interesting to see if a test utilizing this information would lead to an even better study design. Such an endeavor will involve weighting of the IBDs of sib pairs with not-so-extreme trait values and certainly is out of the scope of this article.

Acknowledgments

This work was supported, in part, by NIH and NIMH grants GM28719 and MH31302, respectively. We thank the two anonymous reviewers for their many helpful comments, which improved this article.

References

Blackwelder WC, Elston RC (1985) A comparison of sib-pair linkage tests for disease susceptibility loci. Genet Epidemiol 2:85–97

Carey G, Williamson J (1991) Linkage analysis of quantitative traits: increased power by using selected samples. Am J Hum Genet 49:786–796

Eaves L, Meyer J (1994) Locating human quantitative trait loci: guidelines for the selection of sibling pairs for genotyping. Behav Genet 24:443–455

Elston RC (1992) Design for the global search of the human genome by linkage analysis. In: Proceedings of the XVIth

International Biometric Conference. Hamilton, New Zealand, pp 39–51

Elston RC (1994) P values, power, and pitfalls in the linkage analysis of psychiatric disorders. In: Gershon ES, Cloninger CR (eds) Genetic approaches to mental disorders. American Psychiatric Press, Washington, D.C., pp 3–21

Fulker DW, Cardon LR, DeFries JC, Kimberling WJ, Pennington BF, Smith SD (1991) Multiple regression analysis of sib-pair data on reading to detect quantitative trait loci. Reading Writing 3:299–313

Gu C, Rao DC (1997) A linkage strategy for detection of human quantitative-trait loci. I. Generalized relative risk ratios and power of sib pairs with extreme trait values. Am J Hum Genet 61:200–210 (in this issue)

Gu C, Todorov A, Rao DC (1996) Combining extremely concordant sibpairs with extremely discordant sibpairs provides a cost effective way to linkage analysis of quantitative trait loci. Genet Epidemiol 13:513–533

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

Zhang H, Risch N (1996) Mapping quantitative-trait loci in humans by use of extreme concordant sib pairs: selected sampling by parental phenotypes. Am J Hum Genet 59: 951–957