Behavior Genetics, Vol. 28, No. 5, 1998

A Comparison of Power to Detect a QTL in Sib-Pair Data Using Multivariate Phenotypes, Mean Phenotypes, and Factor Scores

Dorret I. Boomsma[1] and Conor V. Dolan[1]

Received 9 Feb. 1998—Final 28 July 1998

The power to detect a quantitative trait locus (QTL) in sib-pair data is investigated. We assume that we have at our disposal 3 or 4 related phenotypic measures in a sample of sib-pairs. Individual differences in these phenotypes are due to a common QTL and specific (i.e., unique to each phenotype) nonshared environmental and specific polygenic additive effects. In addition, models are considered that include common nonshared environmental effects and/or common polygenic additive effects. We calculate the power to detect the QTL in a genetic covariance structure analysis of the multivariate data, of the mean phenotypic data, and of factor scores. The use of factor scores is shown to be universally more powerful than the use of multivariate or mean phenotypic data. We also investigate the effect of using a single sample of sib-pairs to both calculate the factor score regression matrix and to carry out the QTL analysis. The use of a single sample to both these ends results in a loss of power compared to the theoretical, expected power. The gain in power attributable to the use of factor scores, however, outweighs this observed loss in power. The advantages of using factor scores in selecting sib-pairs are discussed.

KEY WORDS: QTL analysis; sib-pair data; simulation studies; multivariate linkage analysis; genetic factor scores; selection.

INTRODUCTION

In linkage analysis of non-Mendelian, complex, traits the detection of loci that explain a small to medium amount of the genetic variance remains problematic. Nonparametric linkage methods such as identity-by-descent (IBD) mapping in sibling pairs were developed to analyze complex traits without having to specify a mode of inheritance at a single locus. The application of nonparametric methods, however, comes at the cost of a loss of power to localize genes that influence complex qualitative and quantitative traits (e.g., Blackwelder and Elston, 1982).

In a recent overview Allison and Schork (1997) discuss several approaches that have been proposed to increase power in linkage analysis of complex traits. These include multipoint analysis using highly polymorphic markers (Fulker, Cherny, and Cardon, 1995), analysis of the complete bivariate trait distribution in sib-pairs (Fulker and Cherny, 1996) instead of difference scores as originally proposed by Haseman and Elston (1972), selection of extreme discordant and concordant sib-pairs for genotyping (Carey and Williamson, 1991; Eaves and Meyer, 1994; Risch and Zhang, 1995) and multivariate approaches incorporating multiple indicators of the QTL (Amos et al., 1990; Boomsma, 1996).

In this paper we address the question how multivariate genetic analysis can be employed to

[1] Department of Psychology, Vrije Universiteit, De Boelelaan 1111, 1081 HV Amsterdam, The Netherlands. Phone: 31-20-4448789. Fax: 31-20-4448832. Email: dorret@psy.vu.nl

0001-8244/98/0900-0329$15.00/0

increase the power to detect a quantitative trait locus (QTL) that influences a complex trait, using IBD mapping in sibling pairs. Specifically, we explore power to detect a QTL as a function of QTL effect (i.e., amount of variance explained by the locus), allele frequencies at the marker locus and number of measured phenotypes. We also explore how the causes of covariance between multiple phenotypes affect the power to detect the QTL effect. The covariance among the multivariate phenotypes, that we consider, is due to a QTL common to all phenotypes, and to common additive polygenic and/or nonshared environmental influences. The environmental influences are common to the phenotypes, but not shared by the sibs. They are sometimes designated specific environmental influences, or E1, for short.

Analyses were carried out with Mx (Neale, 1997), using a structural equation modeling approach. The application of structural equation modeling (SEM) in the genetic analysis of complex traits involves the specification of the phenotype as a function of unobserved, or latent, variables. The phenotype may be univariate or multivariate. The latent variables are genetic and environmental factors. It is possible to test whether correlated multiple variables, measured in the same group of genetically related subjects, are indicators of the same underlying latent genetic factor, and/or the same environmental factors (Martin and Eaves, 1977). Measured environmental variables or DNA marker data can be incorporated into the structural equation model. The incorporation of marker data allows one to test whether the phenotypic covariance is due to a QTL influencing all measures (Eaves, Neale and Maes, 1996). A common QTL, influencing multiple phenotypic measures, generally is easier to detect in multivariate than in univariate analyses. It is well established that in SEM power can be influenced by manipulating the number of indicator variables of the latent trait of interest. Although discussions of power usually address the influence of sample size, adding another measured indicator of the latent trait to the model can in some circumstances have the same effect as doubling the sample size (Matsueda and Bielby, 1986).

To investigate the power to detect a common QTL in multivariate data, we considered four series of multivariate phenotypes. The covariances between variables were due to different combinations

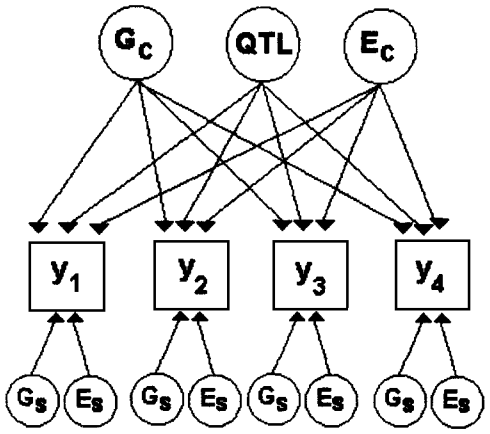

Fig. 1. Path diagram for the multivariate QTL model.

of a common QTL, common polygenic background, and environmental influences. Several different approaches were used to analyze the multivariate data: fitting the complete multivariate model to the data (Eaves, Neale and Maes, 1996), analyzing the mean phenotype in a standard univariate QTL analysis, and analyzing individual genetic factor scores (Boomsma, Molenaar and Orlebeke, 1990).

METHOD

Four sets of analyses were carried out on p related phenotypes, where p = 3, or p = 4, to determine the increase in power that can be achieved by analyzing multivariate, rather than univariate, phenotypes. The model that was used to simulate the multivariate data sets in sibling pairs is depicted in Fig. 1. We considered multivariate phenotypes whose variation may be caused by: a) a QTL that is common to all phenotypes, b) additive polygenic influences (G_c) common to all phenotypes, c) nonshared environmental influences (E_c) common to all phenotypes, d) specific additive genetic influences, that is, genetic influences specific to a given phenotype (G_s), and e) specific, nonshared environmental influences (E_s). The covariance between phenotypes is a function of the loadings of the phenotypes on the common factors Q, G_c and E_c. We considered 4 models that vary with respect to these influences:

1) the QTL is the only source of covariation between measures;

Table I. Percentages of Variance for Phenotypes in 3- and 4-Variate Models

Model 1		3 phenotype model			4 phenotype model		
QTL factor	28.5	14.3	07.1	28.5	14.2	07.1	21.4
genetic specific	28.5	42.8	50.0	28.5	42.8	50.0	35.7
genetic factor	00.0	00.0	00.0	00.0	00.0	00.0	00.0
environmental specific	42.8	42.8	42.8	42.8	42.8	42.8	42.8
environmental factor	00.0	00.0	00.0	00.0	00.0	00.0	00.0
Model 2		3 phenotype model			4 phenotype model		
QTL factor	28.5	14.2	07.1	28.5	14.2	07.1	21.4
genetic specific	14.2	28.5	35.7	14.2	28.5	35.7	21.4
genetic factor	14.2	14.2	14.2	14.2	14.2	14.2	14.2
environmental specific	42.8	42.8	42.8	42.8	42.8	42.8	42.8
environmental factor	00.0	00.0	00.0	00.0	00.0	00.0	00.0
Model 3		3 phenotype model			4 phenotype model		
QTL factor	28.5	14.2	07.1	28.5	14.2	07.1	21.4
genetic specific	28.5	42.8	50.0	28.5	42.8	50.0	35.7
genetic factor	00.0	00.0	00.0	00.0	00.0	00.0	00.0
environmental specific	28.5	28.5	28.5	28.5	28.5	28.5	28.5
environmental factor	14.2	14.2	14.2	14.2	14.2	14.2	14.2
Model 4		3 phenotype model			4 phenotype model		
QTL factor	28.5	14.2	07.1	28.5	14.2	07.1	21.4
genetic specific	14.2	28.5	35.7	14.2	28.5	35.7	21.4
genetic factor	14.2	14.2	14.2	14.2	14.2	14.2	14.2
environmental specific	28.5	28.5	28.5	28.5	28.5	28.5	28.5
environmental factor	14.2	14.2	14.2	14.2	14.2	14.2	14.2

2) the covariation among the measures is due to Q and G_c;
3) the covariation is due to Q and E_c; and
4) the covariation is due to Q, G_c, and E_c.

The parameter values that were used to simulate the data according to these 4 models are summarized in Table I. The total heritability for each variable was 57%, but the heritability due to the QTL varied from phenotype to phenotype. In the 3- and 4-variate data sets the QTL accounted for 28.5, 14.2 and 7.1% of the total phenotypic variance of the first three variables, respectively. In the 4-variate data set, 21.4% of the variance in the fourth variable was accounted for by the QTL. Environmental factors (common plus specific) accounted for 43% of the total variance in all variables. The phenotypic correlations between the measures, implied by the parameter values in Table I, are given in Table II.

Multi-group structural equation models were fitted to covariance matrices of the sib-pairs to estimate the common and specific genetic and environmental effects. As we are working with data of sib-pairs, or DZ twins, the input covariance matrices are of the order 2p. For each model that we consider, we compute the total number of sib-pairs

Table II. Phenotypic Correlation Matrices (Last Row of Each Correlation Matrix Is Absent in the Case of 3 Phenotypes)

Model 1				Model 2			
1				1			
.202	1			.345	1		
.143	.101	1		.286	.244	1	
.247	.175	.124	1	.390	.318	.267	1
Model 3				Model 4			
1				1			
.345	1			.487	1		
.286	.244	1		.428	.386	1	
.390	.318	.266	1	.533	.461	.409	1

required to detect the QTL with a power of 80% and an α of .001. The distribution of the total number of subjects over the IBD groups and the number of IBD groups specified in the analyses both depend on the informativeness of the marker. The expected proportion of alleles shared IBD, π, determines the correlation between the QTL effects in sibling 1 and sibling 2 (Fulker and Cherny, 1996; Eaves, Neale, and Maes, 1996). IBD status at the QTL is established on the basis of parental and offspring genotypic data relating to a single marker at 0 centimorgans (cM) distance from the

Table III. Distribution of π Given Number of Equi-Frequent Marker Alleles, M

π	probability	M= PIC=	8 .861	12 .910	16 .933	∞ 1.00
0, 1	$(M-1)^*(M^2-M-1)/(4M^3)$		.188	.208	.219	.250
.25, .75	$(M-1)/M^2$		.109	.076	.059	.000
.5	$(M^3-2M^2+4M-1)/(2M^3)$		.405	.430	.445	.500

QTL. The number of equi-frequent alleles at the marker locus was set to 8, 12, 16, or ∞. In the first 3 cases, a five group analysis was carried out, with values of π within each group equaling 0, .25, .5, .75, and 1.0 (Haseman and Elston, 1972, Table 2). The proportion of sibling pairs in each of these groups depends on the number of alleles at the marker locus, as is shown in Table III. The derivation of Table 3 is given in Appendix I. As the number of marker alleles increases, the proportion of sibling pairs in the intermediate groups, where $\pi = .25$ and $\pi = .75$, approaches zero. In the infinite allele case, a three-group analysis is carried out. In each group, the correlation between common polygenic influences (G_c) and between specific polygenic influences (G_s) in sib 1 and sib 2 are 0.5.

Power calculations were carried out by fitting the known model to the 2p × 2p (exact, or population) covariance matrices. Model fitting was carried out with and without the QTL effect. Power calculations were based on the noncentral chi-square distribution (e.g., Hewitt and Heath, 1988; see Neale and Cardon, 1992, p. 190, for an outline of the procedure). Model fitting and power calculations were carried out using the Mx program (Neale, 1997). Throughout models were fitted by minimizing the multi-group log-likelihood ratio function. We report the number of subjects required to reject the alternative model excluding the QTL with an α of 0.001 and a power $(1 - \beta)$ of .80. We also report the noncentrality parameters so that the reader may calculate their own estimates of power for different α levels and sample sizes (see Hewitt and Heath, 1988). Calculations were carried out for the following cases: The complete multivariate model, a univariate analysis in which the phenotype is the mean (unweighted linear combination) of the p phenotypes, and univariate analyses in which the phenotype is a weighed linear combination of the p phenotypes. The weights used were obtained from the factor score regression matrix (Boomsma, Molenaar and Orlebeke, 1990). Two approaches to obtain the factor scores were explored: one employing the phenotypic data from the subject and his or her sibling and one in which only the data from the single subject were used (see below).

Let Σ_i denote the multivariate (2p × 2p) population covariance matrix in group i (i = 1,. . ., 3, or i = 1,. . ., 5) and let **w** denote a (2 × 2p) matrix of weights. Covariance (2 × 2) matrices of siblings' mean phenotypic scores or of siblings' genetic factor scores are calculated as $\mathbf{w}\Sigma_i\mathbf{w}^t$. The univariate mean phenotypic covariance matrix is calculated by specifying:

$$\mathbf{w} = \begin{bmatrix} 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \end{bmatrix}$$

in the 3-variate case, and in the 4-variate case:

$$\mathbf{w} = \begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/4 & 1/4 & 1/4 & 1/4 \end{bmatrix}$$

The factor score regression matrix (**w**) is obtained after a standard genetic covariance structure analysis of the pooled phenotypic covariance matrices (i.e., pooled over the different IBD groups). The absence of environmental influences shared by sibs must be assumed in the absence of MZ data, or can be tested when MZ data are available. In the genetic covariance structure analysis of the pooled phenotypic covariance matrix, we include a single common genetic factor that accounts for both the effects of the common QTL and the common polygenic effects, if present. The phenotypic covariance matrix (2p × 2p) is modeled as: $\Sigma = \Lambda\Psi\Lambda^t$. The matrix Ψ is the covariance matrix of the latent genetic and environmental factor scores. The matrix Λ contains the factor loadings of the phenotypic variables on the latent variables. The dimensions of Ψ and Λ depend on the number of observed (p) and latent (k) factors in the analysis. If we consider one common genetic and one common nonshared environmental factor for each sibling and genetic and environmental factors specific to each phenotype, then k = 2 + 2p. The Λ matrix thus is of the order 2p × 2k and Ψ is 2k × 2k. As we are considering sib-pairs, Σ can be partitioned as follows:

$$\begin{bmatrix} \Sigma_{11} & \Sigma_{21}{}^t \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} = \begin{bmatrix} \Lambda_{11} & 0 \\ 0 & \Lambda_{22} \end{bmatrix} \begin{bmatrix} \Psi_{11} & \Psi_{21}{}^t \\ \Psi_{21} & \Psi_{22} \end{bmatrix} \begin{bmatrix} \Lambda_{11}{}^t & 0^t \\ 0^t & \Lambda_{22}{}^t \end{bmatrix}$$

Table IV. Number of Sibling Pairs Required to Achieve Power of .80 (α = .001) for the Four Models[a] Considered

M=	8	12	16	∞	8	12	16	∞
PIC=	.861	.910	.933	1.00	.861	.910	.933	1.00
Model 1		3-variate				4-variate		
multivariate	907	857	837	776	604	571	557	516
mean phenotype	859	813	793	736	504	477	465	431
factor scores 1	683	647	631	586	431	408	398	369
factor scores 2	720	681	664	616	447	422	412	382
Model 2		3-variate				4-variate		
multivariate	1191	1127	1100	1021	886	838	817	758
mean phenotype	1207	1142	1114	1035	786	744	726	674
factor scores 1	978	925	903	838	673	636	621	576
factor scores 2	997	943	920	854	680	643	627	582
Model 3		3-variate				4-variate		
multivariate	1369	1296	1264	1174	1081	1023	998	926
mean phenotype	1454	1376	1343	1248	1005	950	928	861
factor scores 1	993	940	917	852	738	698	681	633
factor scores 2	1086	1027	1003	931	800	757	739	686
Model 4		3-variate				4-variate		
multivariate	1618	1531	1494	1388	1365	1292	1260	1170
mean phenotype	1910	1808	1764	1639	1401	1326	1294	1202
factor scores 1	1327	1256	1226	1139	1024	969	946	879
factor scores 2	1414	1338	1306	1213	1086	1027	1003	931

[a] Model 1: covariance due to Q; model 2: covariance due to Q and G; model 3: covariance due to Q and E; model 4: covariance due to Q, G and E. Factor scores 1 refers to genetic factor scores based on familial phenotypic information; factor scores 2 refers to factor scores based on individual phenotypic information.

where Σ_{11} (Σ_{22}) is the (p × p) sib 1 (sib 2) phenotypic covariance matrix, Λ_{11} (p × k) = Λ_{22}, and Ψ_{11} (k × k) = Ψ_{22} and 0 is a (p × k) matrix of zeros. The factors in the model are standardized and uncorrelated within the individual members of the sib-pair, so that $\Psi_{11} = \Psi_{22} = I$. The sib1–sib2 cross-correlation matrix, Ψ_{21}, contains the expected correlations among the factors. These equal zero in the case of environmental factors (which are all nonshared here), or .5 in the case of additive genetic factors. The factor score regression matrix (2k × 2p) is calculated as $\Psi\Lambda^t\Sigma^{-1}$. The regression matrix contains the weights that are used to obtain linear combinations of the observed variables that approximate the latent variables. The rows in $\Psi\Lambda^t\Sigma^{-1}$ that yield weights for the common genetic factor scores for sib 1 and sib 2 are inserted into $\mathbf{w}$ (2 × 2p) and the expected (population) covariance matrix of genetic factor scores is calculated as $\mathbf{w}\Sigma_i\mathbf{w}^t$ for each group defined by its π value. Again using the sib1–sib2 partitioning we may write $\mathbf{w}$ (2 × 2p) as follows:

$$\mathbf{w} = \begin{bmatrix} \mathbf{w}_{11} & \mathbf{w}_{12} \\ \mathbf{w}_{21} & \mathbf{w}_{22} \end{bmatrix}$$

where $\mathbf{w}_{11}$ (1 × p) = $\mathbf{w}_{22}$. When $\mathbf{w}_{21}$ (1 × p) = $\mathbf{w}_{12}$ does not equal zero, the phenotypic scores of sib1 are involved in the calculation of the factor scores of sib2, and vice versa. To obtain an indication of the information that this procedure provides, we also analyze the covariance matrices calculated using factor score weights after setting sub-matrices $\mathbf{w}_{21}$ and $\mathbf{w}_{12}$ to equal zero. We call these factor scores individual factor scores, or factor scores 2 in Table IV.

RESULTS

If a univariate QTL analysis of the phenotypes were carried out, the number of sib-pairs required to detect a QTL accounting for 28.5%, 21.4%, 14.2% and 7.1% of the phenotypic variance would equal 1288, 2307, 5218 and 20938, respectively (α = 0.001, 1 − β = 0.80, and PIC = 1). Table IV contains the number of sibling pairs required to detect the common QTL in multivariate data. The number of sib pairs is given for each of the four models considered (covariance due to Q, to Q and G_c, to Q and E_c, and to Q, G_c and E_c). Either the

full multivariate model or a univariate model was fitted. In the univariate models, either the mean (unweighted) phenotypes, or (weighted) genetic factor scores were analyzed. Results are given for p = 3 and p = 4 phenotypes and for different PIC values at the marker locus (associated with 8, 12, 16, or ∞ marker alleles). It is clear from Table IV that the power to detect the common QTL is greatest when the QTL is the only source of covariation between phenotypes. For example, the full multivariate analysis for the case of p = 4 and PIC = .86 requires 604 sib-pairs when the QTL is the only factor common to all four phenotypes. If, in addition to a common QTL, there is also covariance among the phenotypes that can be ascribed to G_c or E_c, the required number of pairs equals 886 and 1081, respectively. When the additional covariance is due to both G and E, the required number of sib pairs is 1365. The noncentrality parameters that correspond to the entries in Table IV are given in Appendix II.

Table IV shows that the presence of an extra phenotype is clearly beneficial, as it results in a considerable increase in power in all cases. A further systematic result is the gain in power that is achieved by analyzing factor scores instead of the full multivariate data. The explanation for this finding is provided by the difference in the degrees of freedom for the full multivariate tests (df = 3 or df = 4) versus the univariate tests (df = 1) in the case of factor scores. A surprising finding is the difference in power observed when analyzing the mean phenotype and the multivariate phenotypes. In all models, except model 1, the analysis of the mean phenotype is characterized by greater power, although in models 2 and 3 the differences are quite small. The analysis of genetic factor scores, calculated at the familial level (factor scores 1) and at the individual level (factor scores 2), perform better than either the multivariate analysis or the analysis of the mean phenotype. The first method, which uses all available phenotypic information in the construction of the genetic factor scores (both from the subject and the sib) is optimal and performs better in a linkage analysis than the second method which only uses the phenotypic data from the subject himself. However, as we discuss below, in selection of subjects for genotyping the question arises whether we want to construct factor scores that include information from sibs.

Figure 2 depicts the plot of PIC values and the required number of sib-pairs for the 32 cases that were considered. From this figure it is clear that the relationship between PIC values and the required N is perfectly linear in all cases, although the slope depends on the exact model and the number of phenotypes in the analysis.

Results Relating to Factor Scores Based on Simulated Data

The results of the power calculations clearly indicate that it is beneficial to use genetic factor scores when analyzing QTL effects in multivariate sibling data. A complication in calculating factor scores is that the weight matrix **w** is calculated in the same sample as the actual factor scores, i.e., the information that is used to calculate **w** and the information used to test the presence of a QTL are not independent. One way around this problem would be to split the sample into two halves and to calculate **w** in each subsample. One could then use the **w** matrix calculated in sample 1 to calculate factor scores in sample 2 and vice versa. This procedure has the drawback that the weights will be subject to greater sampling fluctuation than the weights calculated in the whole sample.

We investigated the effect of calculating both **w** and the genetic factor scores in one and the same sample. To this end, two models were considered: model 4 (covariance between phenotypes due to Q, G_c and E_c) with 4 phenotypes and a second 4-phenotype model with less variable and smaller QTL effects. For the first model, we simulated phenotypic data for 4000 samples consisting of 1000 sib pairs and genotypic marker data for these sib pairs and their parents. The marker, which is 0 cM from the QTL, has 12 equi-frequent alleles. The bi-allelic QTL is codominant with equi-frequent alleles.

Data simulations and analyses were carried out using our own programs. The power calculations presented above are based on 5-group or 3-group covariance structure analyses, but in the present simulations, we maximize the raw-data log-likelihood (see Fulker and Cherny, 1996; Eq. 12). These two approaches produce the same results in the present situation. In addition to the genetic factor scores based on the exact weight matrix **w**, henceforth $\mathbf{w}_x$, and the calculated matrix **w**, hence-

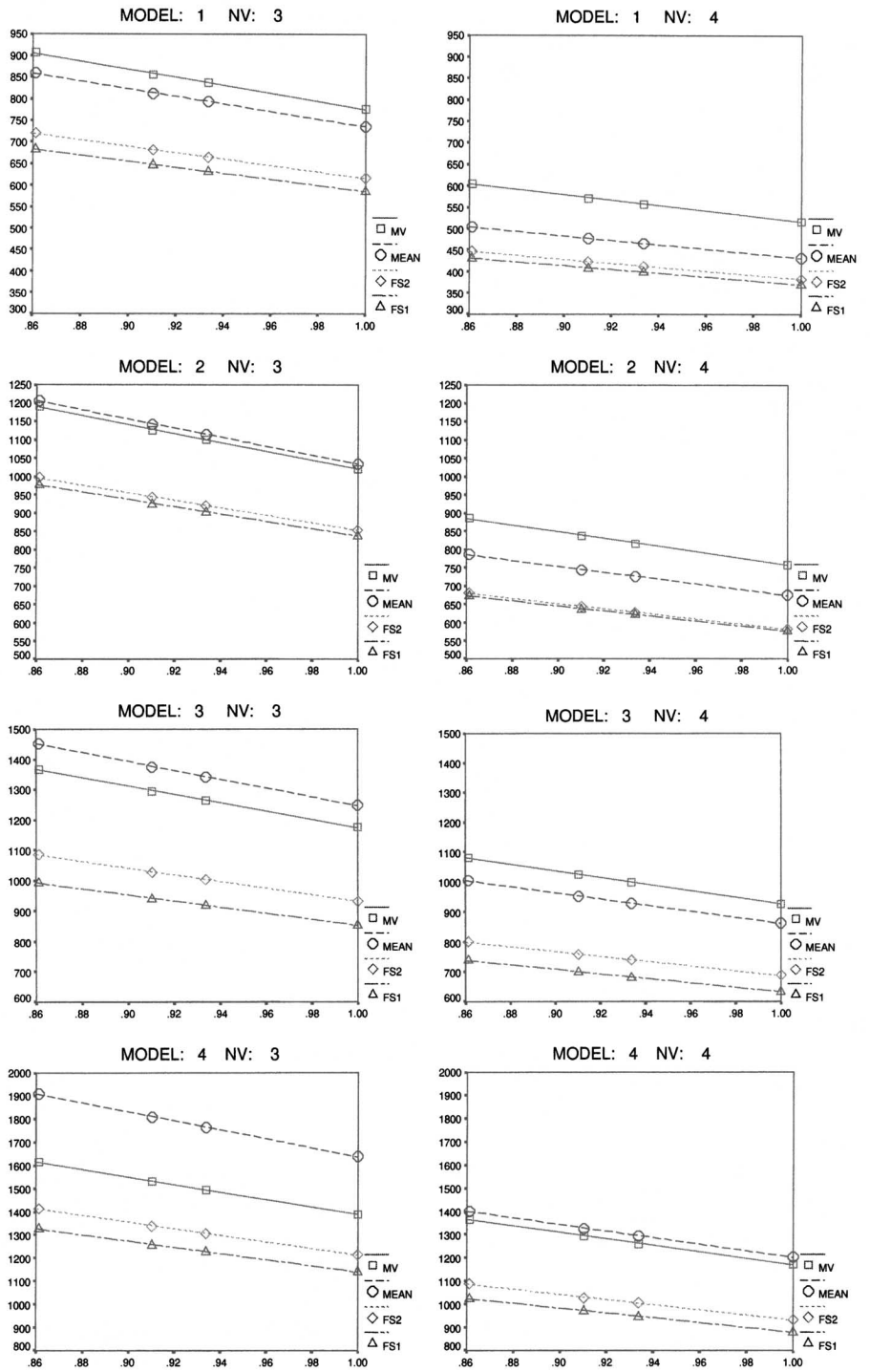

Fig. 2. Number of subjects required, as a function of PIC, to achieve a power of .80 given an α of .001 for the models described in Table IV. Plots are shown for the analysis of multivariate data (mv), mean phenotypic data (mean), factor scores based on familial phenotypic information (fs1) and factor scores based on individual phenotypic information (fs2).

forth $\mathbf{w}_c$, we also analyze the mean phenotypic scores for reasons of comparison.

Table V contains the results of the 4000 replications. Both the analysis of the mean phenotype and the analysis of factor scores using $\mathbf{w}_x$ produced results that are in good agreement with the expected results. The effects of using $\mathbf{w}_c$, instead of $\mathbf{w}_x$, are limited to the power to detect the QTL. We expect a power of .82, but observe a value of .76. That is, given an $\alpha = 0.001$, in 76%, instead of the expected 82%, of the analyses the QTL is detected. With respect to the parameter estimates, we observe a negligible overestimation of the variance due to polygenic effects (.704 vs. .689), and negligible underestimation of the QTL effect (.296 vs. 310). The analysis of the mean phenotype produced results that are in line with expectation. The observed and expected power are the same, but considerably lower than the power observed when analyzing factor scores.

The probability of failure to detect the presence of the QTL is clearly affected by the use of $\mathbf{w}_c$, instead of $\mathbf{w}_x$, in calculating the factor scores. In the present case, the QTL effect on the 4 phenotypes is variable, ranging from 28.5% to 7.1%. In order to further assess the effect of using $\mathbf{w}_c$, we carried out a second simulation. The model used in this second simulation is characterized by a QTL effect accounting for 10% of the variance of each of the four phenotypes and a background genetic variance of 40% (20% due to a common genetic factor and 20% due to genetic factors specific to each phenotype). The environmental effect of 50% was equally divided over common and specific factors. The correlation among the 4 phenotypes thus is 0.55. We carried out 1500 replications, each involving 4629 sib-pairs. Results are given in Table VI. In the present simulation, involving a larger sample size and a smaller and much more uniform QTL effect, we observed no bias in parameter estimates and a very slight difference in power. The expected power given 4629 sib pairs equals .80. The observed power using $\mathbf{w}_c$ equals .78.

An additional complication in using the same data to calculate the factor score regression matrix and to fit the QTL effect is that the false positive error rate may be affected. To investigate this possibility, we simulated data without a QTL effect. Parameter values were created by distributing the QTL variance in model 4 (4 phenotypes) evenly among the other sources of individual differences (specific and common genetic and environmental effects). One thousand datasets each consisting of 1000 sibling pairs were simulated and analyzed. Factor scores were calculated using the exact and the estimated factor score regression matrix. Likelihood ratio tests were obtained by fitting the model with and without the QTL effect to the factor scores. As pointed out by Sham (1998), minus twice the difference in loglikelihoods is distributed as a .5:.5 mixture of 0 and a χ^2 (1) variate. We therefore expect the average of the likelihood ratio to equal .5 instead of 1 (expected standard devia-

Table V. Results of Fitting Model 4 (4 Variables) to Mean Scores and to Genetic Factor Scores using Exact and Calculated Weight Matrix (**w**)

	mean phenotype	exact **w**	calculated **w**
ncp[a]	12.880	17.616	17.616
expected χ^2	13.880 (7.31)[b]	18.616 (8.51)[b]	18.616 (8.51)[b]
observed χ^2	13.944 (7.40)	18.811 (8.91)	17.549 (8.72)
expected power	.62	.82	.82
observed power	.62	.81	.76
expected %polyg	.349	.689	.689
observed %polyg	.349	.691	.704
expected %qtl	.287	.310	.310
observed %qtl	.287	.309	.296

[a] ncp = noncentrality parameter, α=.001 (see Appendix II).
[b] Mean of noncentral χ^2 equals ncp + df; standard deviation of noncentral χ^2 equals $[2(ncp+df)(1+(ncp/(ncp+df)))]^{.5}$ (Abramowitz and Stegun, 1970). Number of replications is 4000, 1000 subjects per replication. 12 allele marker. Degrees of freedom for the χ^2 test equal 1.

Table VI. Results of Fitting Model 4 (4 Variables) to Genetic Factor Scores using Exact and Calculated Weight Matrix (**w**)

	exact **w**	calculated **w**
ncp[a]	17.073	17.073
expected χ^2	18.073 (8.38)[b]	18.073 (8.38)[b]
observed χ^2	18.223 (9.10)	17.949 (9.09)
expected power	.80	.80
observed power	.79	.78
expected %polyg	.843	.843
observed %polyg	.844	.846
expected %qtl	.156	.156
observed %qtl	.155	.154

[a] ncp = noncentrality parameter, α = .001
[b] See footnote Table V. N replications 1500. 4629 subjects per replication, 16 allele marker. Degrees of freedom associated with the χ^2 test equal 1.

tion is sqrt (1.25) = 1.118). The mean loglikelihood ratio, based on the 1000 replications, is found to equal .538 (sd = 1.19) using the exact weights, and .539 (sd = 1.229), using the estimated weights. We also counted the number of times the loglikelihood ratio exceeded 3.84. In regular chi-square testing, this is the value associated with an alpha of .05. In view of the unusual null distribution, we expect .025 (25/1000; instead of 50/1000) false positives. We observe .031 using exact factor scores weights and .029 using the estimated factor scores regression weights. Neither of these values differs significantly from the expected value of .025 (p. = .10 and .18, respectively). In conclusion, we find that the false positive error rate does not seem to be affected by the use of the same data to both obtain factor scores regression weights and to model the QTL effects.

DISCUSSION

The analyses presented in this paper illustrate the increase in power that may be accomplished when multivariate indicators of the QTL are used instead of a single phenotypic measure. This general result holds across different PIC values for the marker locus and across different models for the sources of covariance between the multivariate measures.

Recent work (Cardon and Fulker, 1994; Eaves and Meyer, 1994; Carey and Williamson, 1991; Risch and Zhang, 1995, 1996; Gu, Todorov and Rao, 1996; Dolan and Boomsma, 1998) has established the value of selecting extreme scoring sibling pairs, both concordant and discordant extreme pairs, for linkage analysis of complex traits. Selection with multivariate data can be difficult, especially when the phenotypic correlations between variables are not high, as was the case for some of the simulated data sets we presented. If the multivariate data show a genetic factor structure with at least one genetic factor common to all measures (e.g., depression and anxiety [Kendler et al., 1987]) genetic factor scores may be constructed and selection carried out on these scores. In following this procedure one does assume that the QTL effect is embedded in the common genetic factor. The plausibility of this assumption depends in part on the nature of the phenotypes in the analysis. Once the marker data are available, however, this assumption may be tested.

In QTL studies involving sib-pair selection the use of factor scores confers several advantages. First, factor scores are one-dimensional and therefore easy to use in selection procedures. Second, genetic factor scores are less affected by systematic and unsystematic environmental effects than phenotypic scores. Extremely scoring subjects selected on the basis of genetic factor scores are more likely to be extreme because of their genetic constitution, rather than irrelevant (i.e., from the perspective of QTL detection) environmental influences. Third, as linear combinations of several phenotypic scores, factor scores tend to approximate the normal distribution more closely and tend not to be as susceptible to floor and ceiling effects. For instance, floor effects may be a considerable problem in the identification of individuals with low scores on paper and pencil tests of depression. Fourth, the availability of genetic factor scores of the parents of the siblings is useful in selection of sib pairs. As pointed out by Zhang and Risch (1996), it is desirable to select extremely concordant sib pairs whose parents have intermediate scores and are therefore more likely to be heterozygous. It is reasonable to expect that intermediate genetic factor scores will provide a better indication of parental heterozygosity than phenotypic scores.

An interesting problem is whether factor scores for selection and linkage analysis should be computed using only the data of the individual subject, or whether phenotypic data of family members (siblings or parents, for example) should be used in addition. The use of phenotypic data of other family members (weighted according to their genetic and environmental relatedness) provides more accurate estimates of genetic factor scores. The greater accuracy results in more power to detect a QTL, as is clear in Table IV (factor scores 1 vs. factor scores 2). However, inclusion of information from other family members in the construction of genetic factor scores may lead to practical problems. When working with unbalanced pedigrees, the reliability of factor scores based on all available phenotypic data may be expected to vary with the number of subjects in the pedigree. In addition, a separate weight matrix has to be calculated for each unique pedigree. The results shown in Table 4 indicate that the loss of power incurred when working with individual genetic factor scores is not very large. For practical purposes, therefore, we believe that using factor scores, based on the phenotypic

information of only the individual, is a feasible option.

It has sometimes been argued that in the construction of factor scores, unit weighting should be used instead of sample-based weights, although this seems mainly an issue in small samples (e.g., Cohen, 1990). The results presented in Tables IV and V include results of the analysis of the mean phenotypic scores (i.e., unit weighting of observations). It is clear that in this particular situation the genetic factor scores perform substantially better than the unweighted mean scores, especially when covariances between measures are due to more than only a pleiotropic QTL effect. This difference between mean scores and factor scores is amplified when one of the variables does not load on the QTL factor (Martin, Boomsma, and Machin, 1997).

Finally, our simulations have indicated that the use of the same sample to calculate the factor score regression matrix and to test the presence of a QTL does have a slightly adverse effect on the power. In both simulations the observed probability to detect a QTL was somewhat lower that the theoretical probability. Still the gain in power attributable to the use of factor scores outweighs this observed loss in power. The linkage analysis of the data that were simulated without a QTL effect indicated that the false positive error rate does not seem to be affected by the use of the same data to obtain factor score regression weights and to model the QTL effect.

APPENDIX I: DERIVATION OF EXPECTED π FREQUENCIES

Let M denote the number of alleles of the marker. We assume equi-frequent alleles so that the allele frequencies each equal q = 1/M. Given this assumption, we reproduce Table 2 of Haseman and Elston (1972).

Haseman and Elston (1972) derive their Table 2 on the basis of possible mating types and sib-pair types which are defined independently of the actual genotype. For example, mating type ii-ii and jj-jj are identical (type I), as are ii-ij and kk-kl (type III), etc. The fifth column of Table AI contains the number of possible mating types when genotype is taken into consideration. For instance, if M = 4, there are 4 possible type I matings (ii-ii, jj-jj, kk-kk, ll-ll). The sum of the products of the elements in the third column and the corresponding element in the fifth column equals 1.0.

Table AI. Adapted from Haseman and Elston (1972, Table 2). Number of Mating sib-pair Types Given M Equifrequent Marker Alleles

Mating type	sib-pair type	prob.	π	number of mating sib-pair types
I:ii-ii	I:ii-ii	q^4	0.5	M
II:ii-jj	V:ij-ij	$2q^4$	0.5	M*(M-1)/2
III:ii-ij	I:ii-ii	q^4	0.75	M*(M-1)
	III:ii-ij	$2q^4$	0.25	M*(M-1)
	V:ij-ij	q^4	0.75	M*(M-1)
IV:ii-jk	V:(2)[1]	q^4	0.75	M*(M-1)*(M-2)/2
	VI:ij-ik	$2q^4$	0.25	M*(M-1)*(M-2)/2
V:ij-ij	I:(2)	$q^4/4$	1	M*(M-1)/2
	II:ii-jj	$q^4/2$	0	M*(M-1)/2
	III:(2)	q^4	0.5	M*(M-1)/2
	V:ij-ij	q^4	0.5	M*(M-1)/2
VI:ij-ik	I:ii-ii	$q^4/2$	1	M*(M-1)*(M-2)/2
	III:(2)	q^4	0.5	M*(M-1)*(M-2)/2
	IV:ii-jk	q^4	0	M*(M-1)*(M-2)/2
	V: (3)	$q^4/2$	1	M*(M-1)*(M-2)/2
	VI:ij-ik	q^4	0	M*(M-1)*(M-2)/2
	VI:ij-jk	q^4	0.5	M*(M-1)*(M-2)/2
	VI:ik-jk	q^4	0.5	M*(M-1)*(M-2)/2
VII:ij-kl	V: (4)	$q^4/2$	1	M*(M-1)/2*(M-2)*(M-3)/4
	VI: (4)	q^4	0.5	M*(M-1)/2*(M-2)*(M-3)/4
	VII: (2)	q^4	0	M*(M-1)/2*(M-2)*(M-3)/4

[a] A given mating type can give rise to several sib-pairs of the same type. For instance, mating type IV gives rise to two sib-pair types V, i.e., ij-ij and ik-ik.

To obtain the probability of a sib pair having $\pi = 0$, we calculate the sum of the products of the elements in column 3 and those in column 5, where $\pi = 0$:

$$\text{prob}(\pi = 0) = q^4/2 * [M^* (M-1)/2] + 2^* [q^{4*} M^* (M-1) * (M-2)/2] + 2^* [q^{4*} M^* (M-1)/2^* (M-2) * (M-3)/4]$$

This can be simplified to $M^* q^{4*} (M - 1) * (M2 - M - 1)/2$, which in turn is written as $(M - 1)(M^2 - M - 1)/(4M^3)$, because q = 1/M. Calculating the probabilities associated with the other values of π in the same manner, we arrive at Table III.

APPENDIX II: NONCENTRAL CHI-SQUARES FOR N = 1000 SIB PAIRS FOR THE FOUR MODELS CONSIDERED IN THE SIMULATIONS

M=	8	12	16	∞	8	12	16	∞
PIC=	.861	.910	.933	1.00	.861	.910	.933	1.00
Model 1	3-variate				4-variate			
multivariate	23.75	25.14	25.74	27.76	38.24	40.45	41.46	44.76
mean phenotype	19.87	21.00	21.53	23.19	33.87	35.79	36.70	39.61
factor scores 1	24.99	26.38	27.05	29.13	39.61	41.79	42.90	46.27
factor scores 2	23.71	25.07	25.71	27.71	38.19	40.45	41.44	44.68
Model 2	3-variate				4-variate			
Multivariate	18.09	19.11	19.58	21.10	26.07	27.56	28.27	30.47
mean phenotype	14.14	14.95	15.32	16.49	21.72	22.94	23.51	25.33
factor scores 1	17.45	18.45	18.90	20.39	25.36	26.83	27.49	29.64
facotr scores 2	17.12	18.10	18.55	19.99	25.11	26.54	27.23	29.33
Model 3	3-variate				4-variate			
Multivariate	15.73	16.62	17.04	18.35	21.37	22.57	23.14	24.94
mean phenotype	11.74	12.40	12.71	13.68	16.98	17.97	18.39	19.83
factor scores 1	17.19	18.16	18.61	20.03	23.13	24.45	25.07	26.97
factor scores 2	15.72	16.62	17.02	18.33	21.34	22.55	23.10	24.88
Model 4	3-variate				4-variate			
Multivariate	13.31	14.07	14.42	15.52	16.92	17.87	18.33	19.74
mean phenotype	8.93	9.44	9.67	10.41	12.18	12.87	13.19	14.20
factor scores 1	12.86	13.59	13.92	14.98	16.67	17.62	18.04	19.42
factor scores 2	12.07	12.76	13.07	14.07	15.72	16.62	17.02	18.33

Model 1: covariance due to Q; Model 2: covariance due to Q and G; Model 3: covariance due to Q and E; Model 4: covariance due to Q, E and G

Factor scores 1 refer to genetic factor scores based on familial phenotypic information and factor scores 2 on individual phenotypic information.

ACKNOWLEDGEMENT

This study was made possible by NWO grant 904-61-090 to DI Boomsma and by a fellowship of the Royal Netherlands Academy of Arts and Sciences to CV Dolan.

REFERENCES

Abramowitz, M. and Stegun, I.A. (1970). *Handbook of Mathematical Functions*. New York: Dover.

Amos, C.I., Elston, R.C., Bonney, G.E., Keats, B.J.B., and Berenson, G.S. (1990). A multivariate method for detecting genetic linkage, with application to a pedigree with an adverse lipoprotein phenotype, *American Journal of Human Genetics*, **47**:247–254.

Allison, D.B. and Schork, N.J. (1997). Selected methodological issues in meiotic mapping of obesity genes in humans: issues of power and efficiency. *Behavior Genetics*, **27**: 401–422.

Blackwelder, W.C. and Elston, R.C. (1982). Power and robustness of sibpair linkage tests and extension to larger sibships. *Comm. Stat. Theory Meth.* **11**:449–484.

Boomsma, D.I. (1996). Using multivariate genetic modeling to detect pleiotropic quantitative trait loci. *Behavior Genetics*, **26**:161–166.

Boomsma, D.I., Molenaar, P.C.M. and Orlebeke, J.F. (1990). Estimation of individual and environmental factor scores. *Genetic Epidemiology*, **7**:83–91.

Cardon, L.R. and Fulker, D.W. (1994). The power of interval mapping of quantitative trait loci, using selected sib pairs. *American Journal of Human Genetics*, **55**:825–833.

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

Cohen, J. (1990) Things I have learned (so far), *American Psychologist*, **45**:1304–1312.

Dolan, C.V. and Boomsma, D.I. (1998). Optimal selection of sib-pairs from random samples for linkage analysis of a QTL using the EDAC test. *Behavior Genetics*, **28**:197–206.

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

Eaves, L.J., Neale, M.C., and Maes, H. (1996). Multivariate multipoint linkage analysis of quantitative trait loci. *Behavior Genetics*, **26**:519–525.

Fulker, D.W. and Cherny, S.S. (1996). An improved multipoint sib-pair analysis of quantitative traits. *Behavior Genetics*, **26**:527–532.

Fulker, D.W., Cherny, S.S. and Cardon, L.C. (1995). Multipoint interval mapping of quantitative trait loci, using sib pairs. *American Journal of Human Genetics*, **56**:1224–1233.

Gu, C., Todorov, A. and Rao, D.C. (1996). 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.

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

Hewitt, J.K. and Heath, A.C. (1988). A note on computing the chi-square noncentrality parameter for power analyses. *Behavior Genetics*, **18**:105–108.

Kendler, K.S., Heath, A., Martin, N.G., and Eaves, L.J. (1987). Symptoms of anxiety and symptoms of depression. Same genes, different environments? *Archives of General Psychiatry*, **44**:451–457.

Martin, N., Boomsma, D., and Machin, G. (1997). A twin-pronged attack on complex traits, *Nature Genetics*, **17**: 387–391.

Martin, N.G., and Eaves, L.J. (1977). The genetical analysis of covariance structure. *Heredity*, **38**:79–95.

Matsueda, R.L. and Bielby, W.T. (1986). Statistical power in covariance structure models. In: N. Brandon-Tuma, Ed. *Sociological Methodology*. pp. 120–158, American Sociological Association, Washington D.C., 1986.

Neale, M.C. (1997). *Mx: Statistical Modeling*, 3rd Edition, Box 710, Department of Psychiatry, MCV, Richmond, VA, 23298.

Neale, M.C. and Cardon, L.C. (1992). *Methodology for Genetic Studies of Twins and Families*. Dordrecht: Kluwer Academic Publishers.

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

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

Sham, P. (1998) *Statistics in Human Genetics*, John Wiley & Sons, New York

Zhang H. and Risch N. (1996). Mapping quantitative trait loci in humans by use of extreme concordant sib pairs: selected sampling by parental phenotypes. *American Journal of Human Genetics*, **59**:951–957.

Edited by Stacey Cherny