Genetic Epidemiology 28: 11–23 (2005)

Power and Robustness of Linkage Tests for Quantitative Traits in General Pedigrees

Wei-Min Chen,[1]* Karl W. Broman,[1] and Kung-Yee Liang[1,2]

[1]*Department of Biostatistics, Johns Hopkins University, Baltimore, Maryland*
[2]*Division of Biostatistics and Bioinformatics, National Health Research Institutes, Taipei, Taiwan*

There are numerous statistical methods for quantitative trait linkage analysis in human studies. An ideal such method would have high power to detect genetic loci contributing to the trait, would be robust to non-normality in the phenotype distribution, would be appropriate for general pedigrees, would allow the incorporation of environmental covariates, and would be appropriate in the presence of selective sampling. We recently described a general framework for quantitative trait linkage analysis, based on generalized estimating equations, for which many current methods are special cases. This procedure is appropriate for general pedigrees and easily accommodates environmental covariates. In this report, we use computer simulations to investigate the power and robustness of a variety of linkage test statistics built upon our general framework. We also propose two novel test statistics that take account of higher moments of the phenotype distribution, in order to accommodate non-normality. These new linkage tests are shown to have high power and to be robust to non-normality. While we have not yet examined the performance of our procedures in the context of selective sampling via computer simulations, the proposed tests satisfy all of the other qualities of an ideal quantitative trait linkage analysis method. *Genet. Epidemiol.* 28:11–23, 2005. © 2004 Wiley-Liss, Inc.

Key words: generalized estimating equations; GEE; variance components

Contract grant sponsor: National Institutes of Health; Contract grant number: GM49909.
*Correspondence to: Wei-Min Chen, Department of Biostatistics, University of Michigan, 1420 Washington Heights, Ann Arbor, MI 48109. E-mail: wechen@umich.edu
Received 26 January 2004; Accepted 9 August 2004
Published online 18 October 2004 in Wiley InterScience (www.interscience.wiley.com)
DOI: 10.1002/gepi.20034

INTRODUCTION

Many human disease phenotypes are inherently quantitative (e.g., hypertension). Others are generally viewed as dichotomous (e.g., diabetes) but are closely associated with intermediate quantitative phenotypes (e.g., glucose tolerance). Numerous statistical methods have been developed for linkage analysis of quantitative traits in human studies [reviewed in Feingold, 2001, 2002]. Haseman-Elston regression [Haseman and Elston, 1972] was one of the first such methods and remains widely used. In this approach, the squared differences between the quantitative phenotypes in sibling pairs are regressed upon the estimated proportion of alleles that they share identical by descent (IBD). A statistically significant negative slope in the regression indicates linkage to a quantitative trait locus (QTL). Based on an observation by Wright [1997], a number of extensions to Haseman-Elston regression, which extract additional information from the sibling pairs' phenotypes, have been proposed [Drigalenko, 1998; Elston et al., 2000; Xu et al., 2000; Forrest, 2001; Sham and Purcell, 2001]. Haseman-Elston regression has also been extended for use with larger sibships [Olson and Wijsman, 1993].

A second approach for quantitative trait linkage analysis in human pedigrees involves the use of variance components models [Amos, 1994; Almasy and Blangero, 1998]. The quantitative phenotypes for the individuals in a pedigree are assumed to follow a multivariate normal distribution, with the correlation between relatives' phenotypes depending on the proportion of alleles IBD at a putative QTL. The variance components approach has been shown to have essentially optimal power in the case that the normal model is correct [Feingold, 2001], but is not robust to departures from normality: when the normal model is not correct, the type I error rate for the test of linkage can be greatly inflated [Allison et al., 1999].

A third approach involves the use of score tests [Tang and Siegmund, 2001; Putter et al., 2002; Wang and Huang, 2002a]. Such score tests have the advantage that, while they are based on a normal model, they can be made robust to departures from normality. Finally, Sham et al. [2002] described a regression-based approach in which the roles of the phenotype and IBD status are interchanged: IBD status is regressed upon the quantitative phenotype. This approach has been shown to be both powerful and robust.

Chen et al. [2004] described a general framework for quantitative trait linkage analysis in human pedigrees, for which many of the above approaches are special cases. The framework makes use of generalized estimating equations (GEE) [Liang and Zeger, 1986], in which one must specify a working covariance matrix. Different choices of this working covariance matrix lead to different methods, and, in particular, one may specify working covariance matrices so that this GEE method is identical to Haseman-Elston regression, certain extensions to Haseman-Elston regression [including those of Sham and Purcell, 2001, and Olson and Wijsman, 1993], and the variance components approach. Under the GEE framework, one obtains estimates of the various genetic parameters, with different choices of the working covariance matrix leading to different estimates. There is additional flexibility in the choice of linkage test statistic.

T. Cuenco et al. [2003] and Szatkiewicz et al. [2003] used computer simulations to investigate the relative performance, in terms of power and robustness, of essentially all available approaches for quantitative trait linkage analysis in sibling pairs, with particular emphasis on the case of selected samples. In this report, we extend their research to investigate a variety of approaches for quantitative trait linkage analysis in sibships and extended pedigrees, though we focus exclusively on the case of random ascertainment. We make use of the general GEE framework of Chen et al. [2004], and investigate the power and robustness of a wide variety of test statistics, including the likelihood ratio test, Wald tests, score tests, and robust versions of these statistics.

In addition, we propose two additional test statistics that take account of the higher moments (skewness and kurtosis) of the phenotype distribution, in order to accommodate non-normality. These new linkage tests are shown to be robust to non-normality but maintain the power of the variance components method.

METHODS

Chen et al. [2004] described a general framework for quantitative trait linkage analysis in general pedigrees that makes use of generalized estimating equations (GEE) and for which many of the current quantitative trait linkage methods are special cases, corresponding to different choices for a working covariance matrix. The approach has considerable flexibility, both in the choice of working covariance matrix and in the ultimate choice of test statistic. In this section, we describe a variety of linkage tests based on this general framework. In the following section, we present the results of computer simulations to investigate the power and robustness of these statistics.

Consider a set of general pedigrees with no inbreeding, and let y_{ki} denote the quantitative phenotype for the ith individual in the kth pedigree. Let Φ_{kij} and Δ_{kij} denote the kinship and fraternity coefficients, respectively, for individuals i and j in pedigree k, and let $\hat{\pi}_{kij}$ and $\hat{\kappa}_{kij}$ denote their expected proportion of alleles shared IBD and the probability that they share 2 alleles IBD, respectively, at a putative QTL, given multipoint marker data. Let σ_a^2 and σ_d^2 denote the additive and dominance variance, respectively, due to a putative QTL, and let σ_{pa}^2, σ_{pd}^2, σ_s^2, and σ_e^2 denote the additive polygenic variance, dominance polygenic variance, shared environmental variance, and non-shared residual environmental variance, respectively. Define $\rho_a = (\sigma_a^2 + \sigma_{pa}^2)/2\sigma^2$, $\rho_d = (\sigma_d^2 + \sigma_{pd}^2)/4\sigma^2$, and $\rho_s = \sigma_s^2/\sigma^2$. Note that $\rho_a + \rho_s$ is the phenotypic correlation for parent-child pairs, and $\rho_a + \rho_d + \rho_s$ is the phenotypic correlation for sibling pairs.

While our general GEE method allows the easy incorporation of environmental covariates, we will focus here on the simple case of no covariates, and we further assume that the population mean phenotype is known. Without loss of generality, we assume $\mathrm{E}(y_{ki}) = \mathrm{E}(y_{ki}|\boldsymbol{M}_{ki}) = 0$, where $\boldsymbol{M}_{ki}$ denotes the available multipoint marker data for individual i in pedigree k. The covariance of the phenotypes for individuals i and j in pedigree k is

$$\Omega_{kij}^0 = \begin{cases} \sigma^2 & i = j \\ (4\Phi_{kij}\rho_a + 4\Delta_{kij}\rho_d + \rho_s)\sigma^2 & i \neq j \end{cases}$$

The covariance of the phenotypes for individuals i and j in pedigree k, conditioned on the available marker data, is

$$\Omega_{kij} = \begin{cases} \sigma^2 & i = j \\ \sigma_a^2(\hat{\pi}_{kij} - 2\Phi_{kij}) + \sigma_d^2(\hat{\kappa}_{kij} - \Delta_{kij}) + \Omega_{kij}^0 & i \neq j \end{cases}$$

The parameters used are linkage parameters σ_a^2 and σ_d^2, and segregation parameters $\rho_a, \rho_d, \rho_s, \sigma^2$. This parameterization is equivalent to the more commonly used parameters $\{\sigma_a^2, \sigma_d^2, \sigma_{pa}^2, \sigma_{pd}^2, \sigma_s^2, \sigma_e^2\}$, but results in somewhat simplified calculations. In the case of data exclusively on sibships, ρ_a, ρ_d, and ρ_s cannot be separately estimated, and so we consider the reduced parameter set $(\sigma_a^2, \sigma_d^2, \rho, \sigma^2)$. Similar parameterizations have been used before [e.g., Tang and Siegmund, 2001].

In the GEE method of Chen et al. [2004], one considers, for pedigree k, the vector $S_k = (y_k' \ (y_k^2 - \sigma^2)' \ \mathrm{Vec}(y_k y_k' - \Omega_k)')'$, where $\mathrm{Vec}(A)$ is a vector consisting of the upper off-diagonal elements of a matrix A, and a matrix, D_k, whose columns consist of the derivatives of S_k with respect to each of the parameters, as follows:

$$D_k = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \hat{\pi}_k - 2\Phi_k & \hat{\kappa}_k - \Delta_k & 4\sigma^2\Phi_k & 4\sigma^2\Delta_k & \sigma^2 & 4\rho_a\Phi_k + 4\rho_d\Delta_k + \rho_s \end{pmatrix}$$

Here the $\hat{\pi}_k, \hat{\kappa}_k, \Phi_k, \Delta_k$ are vectors of length $n_k(n_k-1)/2$, and the 0s and 1s in the first two rows are vectors of length n_k. One then chooses a working covariance matrix, W_k (that is, an assumed form for the conditional covariance matrix of S_k), and takes as parameter estimates the solutions of the equations

$$\sum_k D_k' W_k^{-1} S_k = 0 \qquad (1)$$

Different choices of the working covariance matrix, W_k, lead to different estimates. In particular, one may choose the following Gaussian working covariance matrix [Prentice and Zhao, 1991]:

$$G_k = \begin{pmatrix} \Omega_k & 0 & 0 \\ 0 & [2\Omega_{kij}^2] & [2\Omega_{kil}\Omega_{kim}] \\ 0 & [2\Omega_{kuj}\Omega_{kvj}] & [\Omega_{kul}\Omega_{kvm} + \Omega_{kum}\Omega_{kvl}] \end{pmatrix}$$

for $1 \le i, j \le n_k$, $1 \le u < v \le n_k$ and $1 \le l < m \le n_k$, where n_k is the number of individuals in pedigree k, and $[2\Omega_{kij}^2]$ denotes a matrix consisting of elements $2\Omega_{kij}^2$. This is the conditional covariance matrix of S_k if y_k given the available marker data is assumed to follow a multivariate normal distribution. When G_k is used as the working covariance matrix, W_k, in the estimating equations (1), then the GEE estimates correspond exactly to the maximum likelihood estimates (MLEs) for the variance components model with the usual normality assumption.

The GEE method, as described so far, provides estimates of the parameters $(\sigma_a^2, \sigma_d^2, \rho_a, \rho_d, \rho_s, \sigma^2)$. In the remainder of this section, we describe a number of possible linkage test statistics, including likelihood ratio tests, Wald tests, and score tests.

LIKELIHOOD RATIO TESTS

In the traditional variance components model [Amos, 1994; Almasy and Blangero, 1998], the trait values of pedigree k, conditional on the marker data, are assumed to follow a multivariate normal distribution with covariance matrix Ω_k (defined above). The test statistic for the likelihood ratio test is

$$T^{\mathrm{LRT}} = \sum_k \ln|\hat{\Omega}_k^0| + \sum_k y_k'(\hat{\Omega}_k^0)^{-1} y_k - \sum_k \ln|\hat{\Omega}_k| - \sum_k y_k' \hat{\Omega}_k^{-1} y_k \qquad (2)$$

where $\hat{\Omega}_k$ and $\hat{\Omega}_k^0$ are the MLEs of the covariance matrix under the full model and under the null model, respectively.

In many previous investigations [e.g., Almasy and Blangero, 1998], the putative QTL was assumed to exhibit no dominance (i.e., $\sigma_d^2 = 0$). The null distribution of the likelihood ratio test statistic is then asymptotically a 50:50 mixture of a $\chi^2(0)$ (that is, a point mass at 0) and a $\chi^2(1)$ distribution [Self and Liang, 1987]. If dominance is considered in forming the test statistic, which we will denote $T^{\mathrm{LR-D}}$, the null distribution is a $1/2 - p : 1/2 : p$ mixture of $\chi^2(0)$, $\chi^2(1)$ and $\chi^2(2)$ [Self and Liang, 1987]. In Appendix A, we derive the mixing proportion, p, which was previously not known [Pratt et al., 2000]. For sibship data, the null distribution is around $0.4 : 0.5 : 0.1$ mixture of $\chi^2(0)$, $\chi^2(1)$ and $\chi^2(2)$, independent of the size of the sibship.

Use of the likelihood ratio test statistic has previously been shown to exhibit an inflated type I error rate in the case that the multivariate normal model is incorrect [Allison et al., 1999]. This problem may be corrected by estimating the true null distribution of the statistic either through

an analytical approach or an empirical approach such as a Monte Carlo or permutation procedure.

Blangero et al. [2000] proposed a robust LOD score approach for general pedigrees. This approach can be derived from the original likelihood ratio test:

$$T^{\text{LRT-R}} = vT^{\text{LRT}}$$

where the coefficient v is the ratio of variance of the MLE estimator versus that of the robust estimator. Although the original procedure proposed by Blangero et al. [2000] is rather complicated, by taking advantage of the data structure of multiple pedigrees, and by virtue of the GEE framework [Chen et al., 2004], the coefficient has a much simpler expression

$$v = \frac{\{(\sum_k D_k'\hat{G}_k^{-1}D_k)^{-1}\}_{11}}{\{(\sum_k D_k'\hat{G}_k^{-1}D_k)^{-1}\sum_k (D_k'\hat{G}_k^{-1}\hat{S}_k)(D_k'\hat{G}_k^{-1}\hat{S}_k)'(\sum_k D_k'\hat{G}_k^{-1}D_k)^{-1}\}_{11}},$$

where the 11 subscript indicates to take the (1,1) element of the matrix.

In the simulation study in the next section, we also consider the following Monte Carlo procedure for the robust likelihood ratio test. We fix the genotypes for all founding individuals in each pedigree and generate random inheritance vectors for the remaining individuals in each pedigree, calculate the likelihood ratio test statistic, and repeat the process multiple times. The null distribution of the test statistic is estimated based on these simulated data; in particular, an appropriate critical value for the statistic is estimated. This procedure is denoted either LR-MC or LR-MC-D, depending on whether dominance is considered.

WALD TESTS

Due to the complexity of taking appropriate account of the dominance effect in the Wald and score tests, all of the remaining linkage tests assume that the putative QTL acts strictly additively, and the parameter set is reduced to $(\sigma_a^2, \rho_a, \rho_d, \rho_s, \sigma^2)$ for general pedigrees or $(\sigma_a^2, \rho, \sigma^2)$ for sibships. We will discuss the influence of ignoring the dominance effect in the simulation section.

The test statistic for the Wald test [Blangero et al., 2001] is

$$T^{\text{Wald}} = \frac{\hat{\sigma}_a^4}{\left\{(\sum_k D_k'\hat{G}_k^{-1}D_k)^{-1}\right\}_{11}}. \quad (3)$$

A robust Wald test [Liang and Zeger, 1986] has test statistic

$$T^{\text{Wald-R}} = \frac{\hat{\sigma}_a^4}{\left\{(\sum_k D_k'\hat{G}_k^{-1}D_k)^{-1}\sum_k (D_k'\hat{G}_k^{-1}\hat{S}_k)(D_k'\hat{G}_k^{-1}\hat{S}_k)'(\sum_k D_k'\hat{G}_k^{-1}D_k)^{-1}\right\}_{11}} \quad (4)$$

Under the null hypothesis of no linkage, both Wald tests are distributed asymptotically as a 50:50 mixture of $\chi^2(0)$ and $\chi^2(1)$.

SCORE TESTS

Putter et al. [2002] described the theory of score test for quantitative trait linkage analysis. Wang and Huang [2002a] proposed a robust score test specifically for sibships. We first summarize these previously described score tests.

Define $D_k^a = \begin{pmatrix} 0 & 0 & \hat{\pi}_k' - 2\Phi_k' \end{pmatrix}$, $S_k^0 = \begin{pmatrix} y_k' & (y_k^2 - \sigma^2)' & \text{Vec}(y_k y_k' - \Omega_k^0)' \end{pmatrix}$, and

$$G_k^0 = \begin{pmatrix} \Omega_k^0 & 0 & 0 \\ 0 & [2(\Omega_{kij}^0)^2] & [2\Omega_{kil}^0\Omega_{kim}^0] \\ 0 & [2\Omega_{kuj}^0\Omega_{kvj}^0] & [\Omega_{kul}^0\Omega_{kvm}^0 + \Omega_{kum}^0\Omega_{kvl}^0] \end{pmatrix}$$

for $1 \le i, j \le n_k$, $1 \le u < v \le n_k$ and $1 \le l < m \le n_k$. The test statistic for the score test [Blangero et al., 2001] is

$$T^{\text{score}} = \frac{(\sum_k D_k^{a'}(G_k^0)^{-1}S_k^0)^2}{\sum_k D_k^{a'}(G_k^0)^{-1}D_k^a}. \quad (5)$$

A more robust version of the score test [Blangero et al., 2001] is the following:

$$T^{\text{score-R}} = \frac{(\sum_k D_k^{a'}(G_k^0)^{-1}S_k^0)^2}{\sum_k (D_k^{a'}(G_k^0)^{-1}S_k^0)^2}. \quad (6)$$

The test proposed by Sham et al. [2002] and implemented in the software MERLIN [Abecasis

et al., 2002] has been shown to be equivalent to another robust score test [Chen et al., 2004], corresponding to the statistic

$$T^{\text{MERLIN}} = \frac{\left(\sum_k D_k^{a'}(G_k^0)^{-1}S_k^0\right)^2}{\sum_k \left(S_k^{0'}(G_k^0)^{-1}\begin{pmatrix}0 & 0\\ 0 & \hat{\Sigma}_{\hat{\pi}_k}\end{pmatrix}(G_k^0)^{-1}S_k^0\right)} \tag{7}$$

where the elements in the covariance matrix $\hat{\Sigma}_{\hat{\pi}_k}$ have the form $\text{Cov}(\pi_{kij}, \pi_{klm}) - (E[\pi_{kij}\pi_{klm}|M_k] - \hat{\pi}_{kij}\hat{\pi}_{klm})$, where $\text{Cov}(\pi_{kij}, \pi_{klm})$ can be calculated given only the structure of the kth pedigree, and $E[\pi_{kij}\pi_{klm}|M_k]$ can be calculated based on the posterior distribution conditional on marker information M_k.

Wang and Huang [2002a] described a robust score test specific for sibship data; their statistic can be rewritten in matrix form (see Appendix B) as

$$T^{\text{score-S}} = \frac{\left(\sum_k D_k^{a'}(G_k^0)^{-1}S_k^0\right)^2}{\overline{(\pi_{\cdots} - 2\Phi_{\cdots})^2} \times \sum_k \left(S_k^{0'}(G_k^0)^{-1}\begin{pmatrix}0 & 0\\ 0 & I\end{pmatrix}(G_k^0)^{-1}S_k^0\right)} \tag{8}$$

where I is an identity matrix of size $\frac{n_k(n_k-1)}{2}\times\frac{n_k(n_k-1)}{2}$, and $\overline{(\pi_{\cdots} - 2\Phi_{\cdots})^2}$ is the average of squared allele-sharings over all available siblings. The robustness of this test relies on the independence of allele-sharing between different sibling pairs, and so it is generally not applicable for pedigrees of more complex structure [Wang, 2002]. Wang and Huang [2002a] described a further approach, in which the phenotypes are converted to ranks that are then transformed to follow a normal distribution; a robust score test [e.g., score-S used by Wang and Huang, 2002a] can then be applied on the transformed data.

Note that, under the null hypothesis of no linkage, all of the score test statistics are distributed as a 50:50 mixture of $\chi^2(0)$ and $\chi^2(1)$.

HIGHER MOMENT SCORE TESTS

The above score tests are derived from the conditional likelihood under the assumption of normality. The only difference among them is in the method for estimating the variance of the score (the denominator in the statistic). Here we propose an alternative approach: novel score tests based on a quasi-likelihood that incorporates information on the higher moments of the phenotype distribution.

Rather than using the Gaussian working covariance matrix, G_k^0, we use the following:

$$M_k^0 = \begin{pmatrix} \Omega_k^0 & \hat{\gamma}_3\sigma^3 I & 0 \\ \hat{\gamma}_3\sigma^3 I & [2(\Omega_{kij}^0)^2] + \hat{\gamma}_4\sigma^4 I & [2\Omega_{kil}^0\Omega_{kim}^0] \\ 0 & [2\Omega_{kuj}^0\Omega_{kvj}^0] & [\Omega_{kul}^0\Omega_{kvm}^0 + \Omega_{kum}^0\Omega_{kvl}^0] \end{pmatrix} \tag{9}$$

where I is an identity matrix of size $n_k \times n_k$, and $\hat{\gamma}_3$ and $\hat{\gamma}_4$ are empirical moment estimates for skewness and kurtosis parameters of the distribution of the phenotype, respectively, which are both 0 for the case of a normal distribution. To be more specific, define $\hat{\sigma}^2 = \overline{(y-\overline{y})^2}$, where overline represents the sample mean, then

$$\hat{\gamma}_3 = \overline{(y-\overline{y})^3}/\hat{\sigma}^3$$
$$\hat{\gamma}_4 = \overline{(y-\overline{y})^4}/\hat{\sigma}^4 - 3$$

Note that the matrix M_k^0 is the true covariance matrix in the case that $\hat{\gamma}_3$ and $\hat{\gamma}_4$ are the true skewness and kurtosis and that all higher moments are 0.

We consider two different test statistics based on the working covariance matrix, M_k^0. The first is a score statistic analogous to the statistic T^{score} in equation (5):

$$T^{\text{HM}} = \frac{(\sum_k D_k^{a'}(\hat{M}_k^0)^{-1}\hat{S}_k^0)^2}{\sum_k D_k^{a'}(\hat{M}_k^0)^{-1}D_k^a} \tag{10}$$

We can also apply the MERLIN-type robust estimator [Sham et al., 2002] for the variance of the estimating function, to make the higher moment approach even more robust

$$T^{\text{HM-R}} = \frac{\left(\sum_k D_k^{a'}(\hat{M}_k^0)^{-1}\hat{S}_k^0\right)^2}{\sum_k \left(\hat{S}_k^{0'}(\hat{M}_k^0)^{-1}\begin{pmatrix}0 & 0\\ 0 & \hat{\Sigma}_{\hat{\pi}_k}\end{pmatrix}(\hat{M}_k^0)^{-1}\hat{S}_k^0\right)}. \tag{11}$$

COMPUTER SIMULATIONS

In order to investigate the power and robustness of the linkage methods described in the previous section, we conducted a computer simulation study. While the methods may accommodate pedigrees of varying size and structure, we considered the simple case that all pedigrees in a study were of the same structure: either sibling

pairs, sibships of size four, sibships of size six, or the three-generation cousin pedigree with 10 individuals investigated by Sham et al. [2002] and displayed in Figure 1.

A quantitative phenotype was simulated with a single major, diallelic QTL, with minor allele frequency 0.3 and explaining 10% of the total phenotypic variance, plus 10 additive, unlinked diallelic polygenes. The alleles at the QTL either acted additively, or the more-frequent allele was fully recessive. In the simulations of sibships, the polygenes accounted for 30% of the total phenotypic variance, and there was an additional shared environment effect accounting for 20% of the phenotypic variance and following a normal distribution. In the simulations with the cousin pedigree, the polygenes accounted for 50% of the total phenotypic variance and there was no shared environment effect. The remaining phenotypic variation was due to an unshared environment effect that was either normally distributed or followed a $\chi^2(1)$ distribution.

A single marker was simulated to be either completely linked to the QTL (recombination fraction, $\theta = 0$) or unlinked ($\theta = 0.5$). For most simulations, the marker was fully informative, though in one set of simulations, with sibships of size four, the marker had four equally frequent alleles.

The number of families were chosen so that, analytically, the variance components method would have 80% power to detect the QTL. There were either 2,999 sibling pairs, 440 sibships of size four, 168 sibships of size six, or 387 cousin pedigrees. All simulations were performed with 5,000 replicates, so that the results have standard error <0.007.

The simulation results are presented in Figures 2–6. The methods studied include the likelihood ratio test (LRT, LRT-D, LRT-R), the likelihood ratio test with 100 Monte Carlo simulations used to determine the appropriate critical value (LR-MC, LR-MC-D), the Wald test (Wald), a robustified Wald test (Wald-R), the score test (score), a robust score test (score-R), the robust score test for sibships (score-S) [Wang and Huang, 2002a], the method implemented in MERLIN-REGRESS (MERLIN) [Sham et al., 2002], our higher moment approach (HM), and a robust version of the higher moment approach (HM-R).

Figure 2 corresponds to the case of a normal model with the alleles at the major QTL acting additively and with a fully informative marker. All methods are seen to have appropriate type I error rate, though the robust score test (score-R) is somewhat conservative in the case of a smaller number of larger sibships. All methods have similar power, though the Wald tests and the robust score test have somewhat lower power, especially for sibships of size six. Note that the robust score test of Wang and Huang [2002a] is appropriate only for sibships, and so was not investigated for the case of the cousin pedigree. The LR-MC method also has somewhat reduced

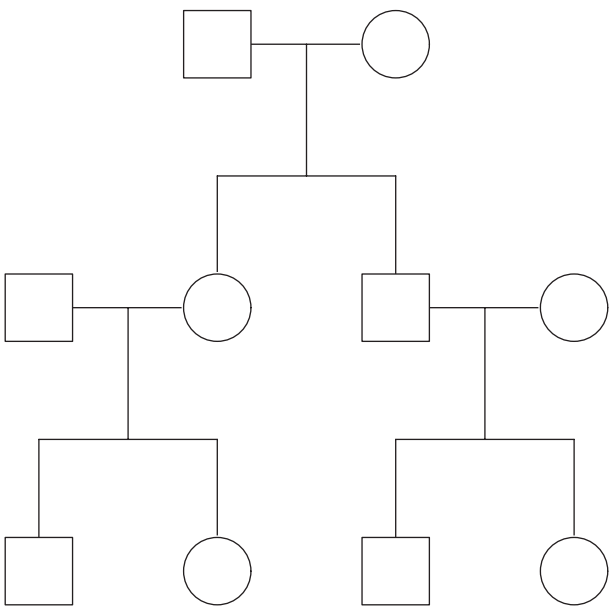

Fig. 1. The first-cousin pedigree considered in the simulation study.

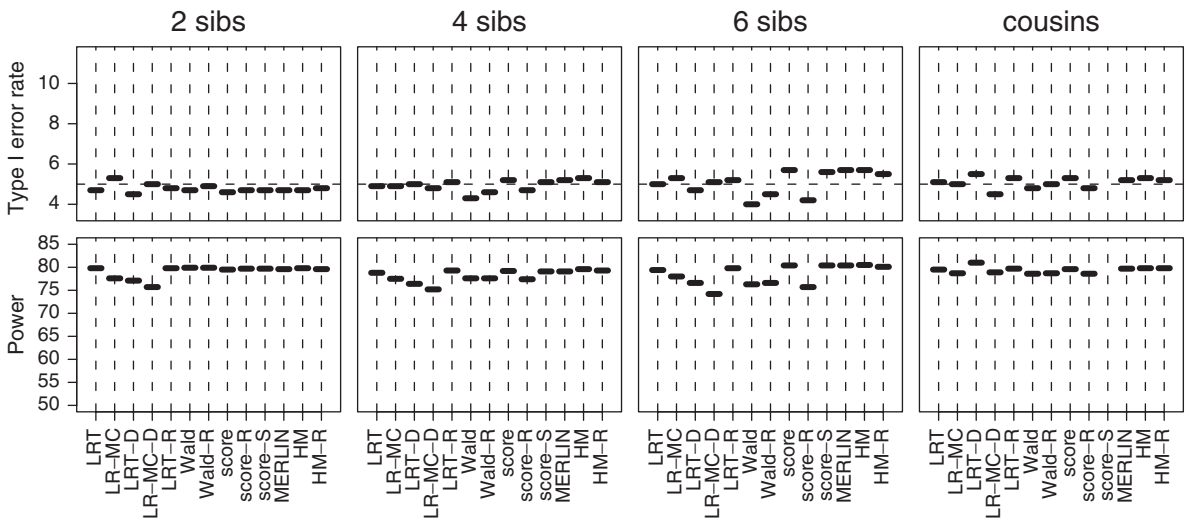

Fig. 2. Type I error rate and power (in percent) for QTL detection in the case that the residual variation follows a normal distribution and the QTL exhibits no dominance.

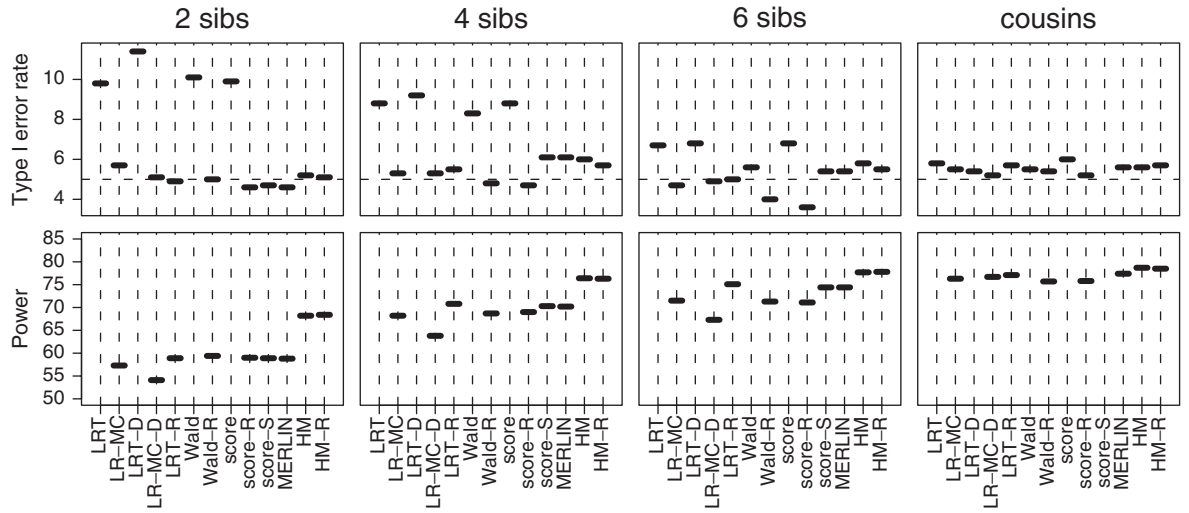

Fig. 3. Type I error rate and power (in percent) for QTL detection in the case that the residual variation follows a χ^2 (1) distribution and the QTL exhibits no dominance.

power, which may be due to the quite limited number of simulations used to estimate the critical value. The allowance for dominance in the likelihood ratio test (LR-D and LR-MC-D) gave slightly reduced power in the case of no dominance, but the type I error rate remained correct.

Figure 3 corresponds to the case that the unshared environment effect followed a $\chi^2(1)$ distribution. Here the likelihood ratio, Wald, and score tests are all seen to have inflated type I error rates (as high as 0.1), and so the power of these methods was not investigated further. The robust tests were generally seen to have type I error under control, though the robust Wald test appears to have an inflated type I error rate in the case of sibling pairs, and the robust score test was again seen to be conservative for the case of a small number of larger sibships. The power of the two higher moment approaches are seen to be higher than the other methods in this non-normal situation, the other approaches all having approximately the same power.

Figures 4 and 5 are analogous to Figures 2 and 3, though with the more-frequent allele at the QTL being fully recessive. The Wald tests had very poor performance and so are not shown in these figures. Of the tests considered here, only the likelihood ratio test (LRT-D, LR-MC-D) takes

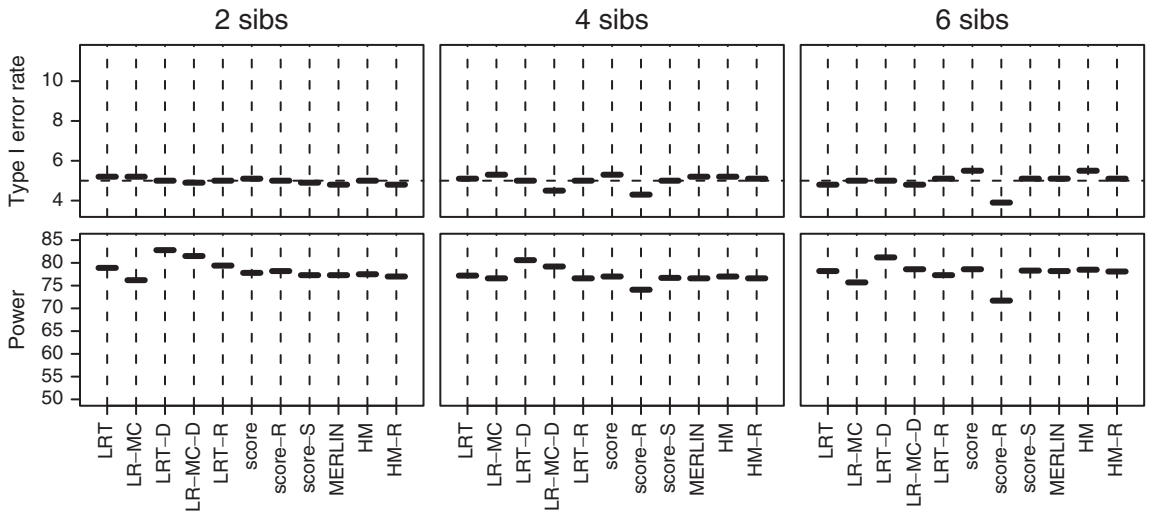

Fig. 4. Type I error rate and power (in percent) for QTL detection in the case that the residual variation follows a normal distribution and the QTL is recessive.

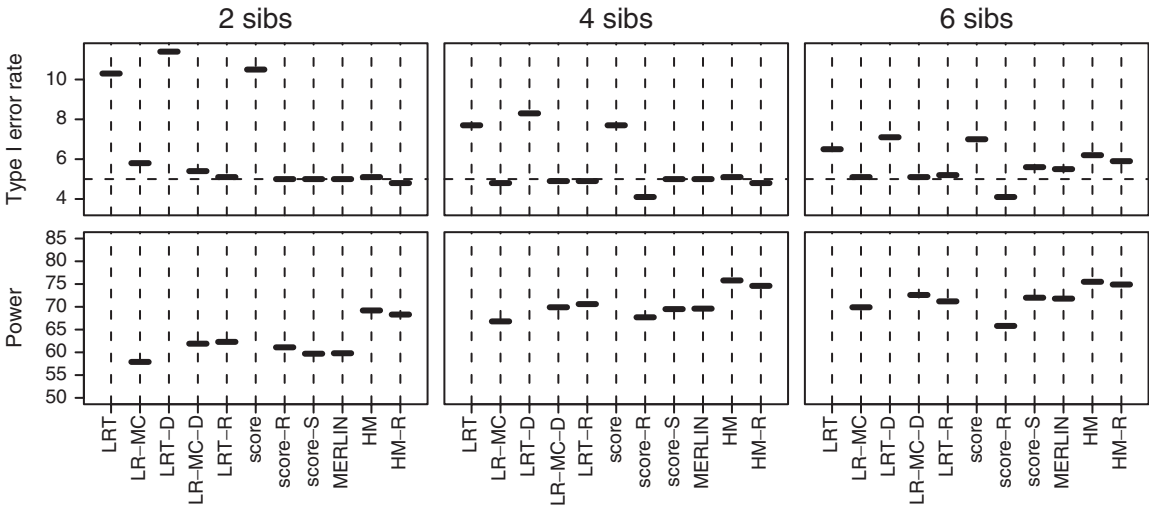

Fig. 5. Type I error rate and power (in percent) for QTL detection in the case that the residual variation follows a χ^2 (1) distribution and the QTL is recessive.

account of the non-additivity at the QTL, and it is seen to have somewhat higher power than other methods under the normal model (Fig. 4). For the likelihood ratio test LRT-D and LR-MC-D, the gain of the power ($\sim 3.5\%$) by taking account of the non-additivity is larger than the loss of power ($\sim 2\%$) for the case of no dominance shown in Figure 2. Tang and Siegmund [2001] and Wang and Huang [2002b] have shown how to modify the score tests to take account of dominance, but since their extension of the score test statistic is asymptotically equivalent to the statistic of the likelihood ratio test, here we only focused on the likelihood ratio test by deriving the critical value of its statistic. In the case of a non-normal model (Fig. 5), the likelihood ratio and score tests again have inflated type I error. The robust versions of the test statistics (including the use of Monte Carlo simulation to identify an appropriate critical value for the likelihood ratio test) have appropriate type I error rates; among these, the higher moment approaches are again seen to have greatest power.

Figure 6 displays the results for the case that the marker is not fully informative (having four equally frequent alleles) and for 440 sibships of size four. The results are similar to those seen in Figures 2–5. In particular, our higher moment approach is seen to be both robust and powerful.

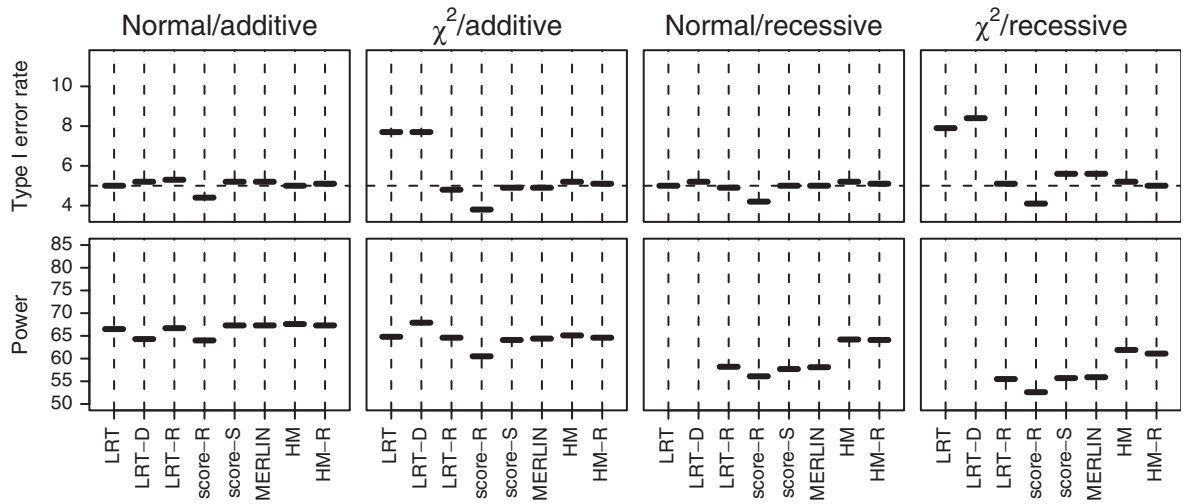

Fig. 6. Type I error rate and power (in percent) using 440 sibships of size 4 and a marker having 4 equally-frequent alleles.

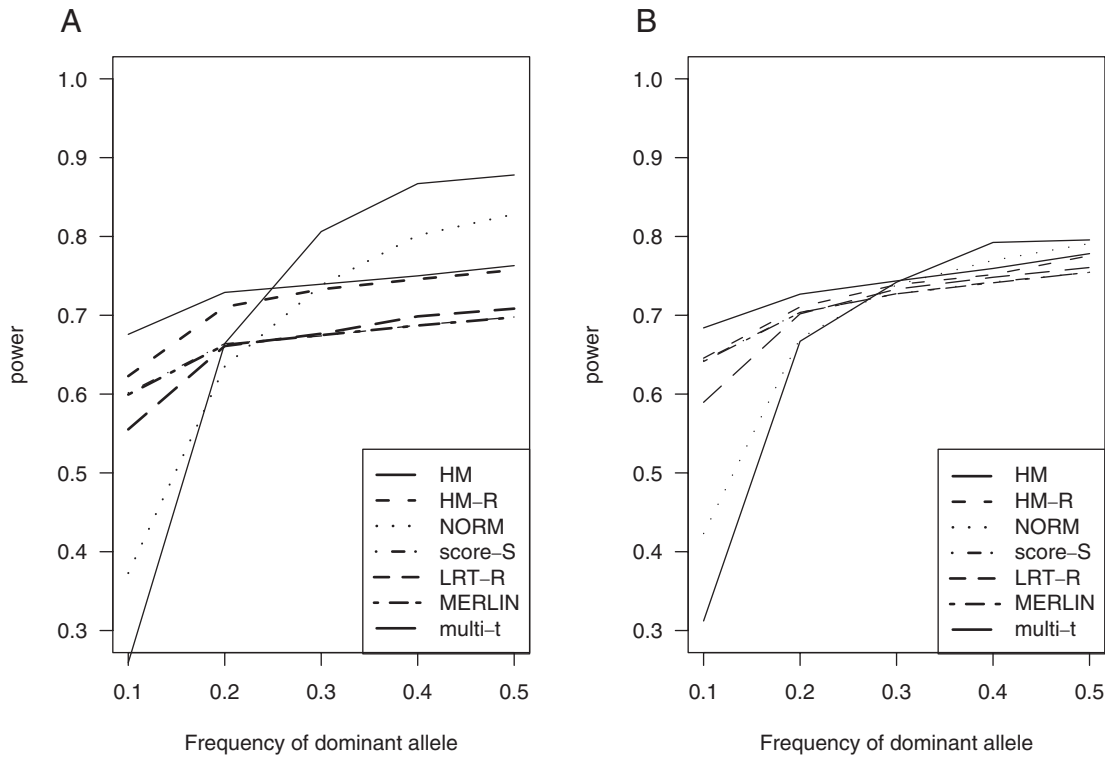

Fig. 7. Power for 440 sibships of size 4, for five different linkage analysis methods. A: Power as a function of the frequency of the dominant allele when the unshared environmental effect follows a χ^2 (1) distribution. B: Power as a function of the frequency of the dominant allele when the unshared environmental effect follows a t(5) distribution.

We also considered the use of a nominal significance level of 0.001 (data not shown); the results were qualitatively similar and would lead to the identical conclusions.

Figure 7 contains the results of further simulations to investigate the effect of the QTL allele frequency on power in the case of non-normality with the more-frequent QTL allele being fully recessive, and for 440 sibships of size four. Here we include results for the transformation procedure proposed by Wang and Huang [2002a] and denoted NORM, and the multivariate t method

[Lange et al., 1989] implemented in SOLAR [Blangero et al., 2001]. Figure 7A corresponds to the case that the environment effect follows a $\chi^2(1)$ distribution, analogous to Figure 5, while Figure 7B corresponds to the case that the environment effect follows a $t(5)$ distribution. Use of the transformation or the multivariate t performs extremely well in the case that the QTL alleles are approximately equally frequent, but performs poorly in the case that the dominant allele has frequency $<20\%$. Note Wang and Huang [2002a] showed this transformation approach reduces the power of the score test when the trait values are approximately normally distributed and the alleles at the major QTL act non-additively. Special attentions should be paid when this empirical approach is applied.

DISCUSSION

Chen et al. [2004] described a general framework for quantitative trait linkage analysis, based on generalized estimating equations (GEE), for which many current methods are special cases. The method has considerable flexibility, both in the choice of working covariance matrix and in the choice of test statistic. In this report, we have expanded upon that work: we proposed two novel higher moment statistics and investigated, through computer simulations, the power and robustness of these new methods relative to previously described approaches, including the variance components method [Amos, 1994; Almasy and Blangero, 1998], the robust LOD score approach [Blangero et al., 2000], the score test proposed by Wang and Huang [2002a], and the method implemented in MERLIN-REGRESS [Sham et al., 2002]. Xu et al. [2000] and Shete et al. [2003] also used GEE in the context of QTL mapping, though they focused exclusively on sibships and particular working covariance matrices. Our approach is more general.

The computer simulations described here were conducted using computer software that we developed, LinkageExplorer (LE). This program is able to simulate general pedigrees and multipoint marker data, perform all of the linkage tests described in this report, and provide analytical sample size calculations (unpublished data). As part of our testing of this software, we compared the results, for simulated data, from our software with those from GeneHunter [Pratt et al., 2000; Kruglyak et al., 1996], SOLAR [Almasy and Blangero, 1998], and MERLIN-REGRESS [Sham et al., 2002]. The likelihood ratio test implemented in LinkageExplorer has similar results to GeneHunter and SOLAR in the case that the QTL alleles acted additively; our implementation of the method of Sham et al. [2002] gave results identical to those of MERLIN-REGRESS.

As has been shown previously [Feingold, 2001], the variance components approach has high power in the case that the normal model is correct, but can have greatly inflated type I error rates in the case of a non-normal phenotype. This non-robustness applies to likelihood ratio test, Wald test, and score test. The robust approaches LRT-R, MERLIN, and score-S are found to have similar power and robustness, while the robust approaches Wald-R and score-R have somewhat lower power unless the number of pedigrees is large. Further simulations (data not shown) showed that while the Haseman-Elston regression [Haseman and Elston, 1972] and its derivatives have proper type I error rate, they have much lower power than the robust approaches investigated in this review. Our higher moment approaches have power similar to the variance components method in the case that the normal model is correct and have a properly controlled type I error rate in the case that the normal model is not correct. Further, in the case that the normal model is not correct, the higher moment approaches are the most powerful methods investigated here.

By using samples selected from normally distributed population, Sham et al. [2002] showed their approach is robust to selective sampling, as long as one can correctly specify the segregation parameters in the random population. This robustness also applies to the higher moment approach HM-R. To see this property, note that with higher moments γ_3 and γ_4 being estimated as 0 in the random population, HM-R is equivalent to Sham et al.'s approach. The practical performance of our higher moment approaches in the context of selective sampling deserves further investigation.

It should be noted that Amos et al. [1996] also proposed a quantitative trait linkage analysis that made use of higher moments of the phenotype distribution, but their approach was based on a Wald test, and they found it did not perform well. In addition, Blangero et al. [2000] made use of higher moments of the phenotype distribution in order to correct the type I error rate of the variance components method, but did

not consider a modification of the test statistic itself.

The use of a transformation to normality [e.g., Wang and Huang, 2002a] and the multivariate-t approach [Blangero et al., 2001] were investigated only in the case of a fully recessive QTL (Fig. 7), and were seen to have especially high power in the case that the two alleles were equally frequent, but had low power in the case that the dominant allele had frequency 10%. In contrast, the higher moment approaches that we have proposed were less affected by the frequency of the QTL alleles.

In the simulations described herein, the population mean phenotype, μ, was assumed to be known. In practice, μ will not be known; further, one will often need to accommodate environmental covariates. Estimation of μ, and of the effects of covariates, may be easily accommodated within our general GEE framework [see Chen et al., 2004]; in fact, this is an important advantage of the approach. If our simulations were repeated, with μ estimated based on the data, the results would likely be the same as those reported, as this nuisance parameter can be quite precisely estimated on the basis of randomly ascertained pedigrees. We examined the scenarios in Figure 6 with μ being estimated (data not shown), and the power differences were found to be less than 1%.

An ideal method for quantitative trait linkage analysis in human studies would have high power to detect a QTL, would be robust to departures from normality (i.e., it would maintain the appropriate type I error rate yet maintain reasonable power to detect a QTL), would be applicable for general pedigrees rather than simply sibling pairs, could incorporate information from environmental and other covariates, and would be appropriate in the presence of selective sampling (e.g., the selection of discordant sibling pairs). While we have not yet examined the performance of our proposed procedures in the context of selective sampling via computer simulations, the higher moment score tests, implemented within the GEE framework of Chen et al. [2004], satisfy all of the other qualities of an ideal quantitative trait linkage analysis method.

ACKNOWLEDGMENTS

The authors thank two anonymous reviewers for valuable suggestions for improving the manuscript.

ELECTRONIC DATABASE INFORMATION

The URL for computer program Linkage Explorer is
http://www.biostat.jhsph.edu/~wmchen/le.html

REFERENCES

Abecasis GR, Cherny SS, Cookson WO, Cardon LR. 2002. Merlin: rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 30:97–101.

Allison DB, Neale MC, Zannolli R, Schork NJ, Amos CI, Blangero J. 1999. Testing the robustness of the likelihood-ratio test in a variance-component quantitative-trait loci-mapping procedure. Am J Hum Genet 65:531–544.

Almasy L, Blangero J. 1998. Multipoint quantitative-trait linkage analysis in general pedigree. Am J Hum Genet 62:1198–1211.

Amos CI. 1994. Robust variance-components approach for assessing genetic linkage in pedigree. Am J Hum Genet 54:535–543.

Amos CI, Zhu DK, Boerwinkle E. 1996. Assessing genetic linkage and association with robust components of variance approaches. Ann Hum Genet 60:143–160.

Blangero J, Williams JT, Almasy L. 2000. Robust LOD score for variance component-based linkage analysis. Genet Epidemiol 19:s8–s14.

Blangero J, Williams JT, Almasy L. 2001. Variance component methods for detecting complex trait loci. Adv Genet 42:151–181.

Chen WM, Broman KW, Liang KY. 2004. Quantitative trait linkage analysis by generalized estimating equations: unification of variance components and Haseman-Elston regression. Genet Epidemiol 26:265–272.

Drigalenko E. 1998. How sib pairs reveal linkage. Am J Hum Genet 63:1242–1245.

Elston RC, Buxbaum S, Jacobs KB, Olson JM. 2000. Haseman and Elston revisited. Genet Epidemiol 19:1–17.

Feingold E. 2001. Methods for linkage analysis of quantitative trait loci in humans. Theor Popul Biol 60:167–180.

Feingold E. 2002. Regression-based quantitative-trait-locus mapping in the 21st century. Am J Hum Genet 71:217–222.

Forrest W. 2001. Weighting improves the "new Haseman-Elston" method. Hum Hered 52:47–54.

Haseman JX, Elston RC. 1972. The investigation of linkage between a quantitative trait and a marker locus. Behavior Genetics 2:3–19.

Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. 1996. Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 58:1323–1337.

Lange KL, Little RJA, Taylor JMG. 1989. Robust statistical modeling using the t distribution. J Am Stat Assoc 84:881–896.

Liang KY, Zeger SL. 1986. Longitudinal data analysis using generalized linear models. Biometrika 73:13–22.

Olson JM, Wijsman EM. 1993. Linkage between quantitative trait and marker loci: methods using all relative pairs. Genet Epidemiol 10:87–102.

Pratt SC, Daly MJ, Kruglyak L. 2000. Exact multipoint quantitative-trait linkage analysis in pedigrees by variance components. Am J Hum Genet 66:1153–1157.

Prentice RL, Zhao LP. 1991. Estimating equations for parameters in means and covariances of multivariate and continuous responses. Biometrics 47:825–839.

Putter H, Sandkuijl LA, van Houwelingen JC. 2002. Score test for detecting linkage to quantitative traits. Genet Epidemiol 22:345–355.

Self SG, Liang KY. 1987. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. J Am Statist Asso 82:605–610.

Sham PC, Purcell S. 2001. Equivalence between Haseman-Elston and variance-components linkage analyses for sib pairs. Am J Hum Genet 68:1527–1532.

Sham PC, Purcell S, Chemy SS, Abecasis GR. 2002. Powerful regression-based quantitative-trait linkage analysis of general pedigrees. Am J Hum Genet 71:238–253.

Shete S, Jacobs KB, Elston RC. 2003. Adding further power to the Haseman and Elston method for detecting linkage in larger sibships: Weighting sums and differences. Hum Hered 55:79–85.

Szatkiewicz JP, T.Cuenco K, Feingold E. 2003. Recent advances in human quantitative-trait-locus mapping: comparison of methods for discordant sibling pairs. Am J Hum Genet 73:874–885.

T.Cuenco K, Szatkiewicz JP, Feingold E. 2003. Recent advances in human quantitative-trait-locus mapping: comparison of methods for selected sibling pairs. Am J Hum Genet 73:863–873.

Tang HK, Siegmund D. 2001. Mapping quantitative trait loci in oligogenic models. Biostatistics 2:147–162.

Wang K, Huang J. 2002a. A score-statistic approach for the mapping of quantitative-trait loci with sibships of arbitrary size. Am J Hum Genet 70:412–424.

Wang K, Huang J. 2002b. Score test for mapping quantitative-trait loci with sibships of arbitrary size when the dominance effect is not negligible. Genet Epidemiol 23:398–412.

Wang K. 2002. Efficient score statistics for mapping quantitative trait loci with extended pedigrees. Hum Hered 54:57–68.

Wright FA. 1997. The phenotypic difference discards sib-pair QTL linkage information. Am J Hum Genet 60:740–742.

Xu X, Weiss S, Xu X, Wei LJ. 2000. A unified Haseman-Elston method for testing linkage with quantitative traits. Am J Hum Genet 67:1025–1028.

APPENDIX A

Self and Liang [1987] showed for the situation when 2 parameters of interest are on the boundary of the parameter space, the asymptotic distribution for the likelihood ratio test statistic is a $\frac{1}{2}-p:\frac{1}{2}:p$ mixture of $\chi^2(0)$, $\chi^2(1)$, and $\chi^2(2)$, where

$$p=\frac{1}{2\pi}\cos^{-1}\frac{I_{12}}{\sqrt{I_{11}I_{22}}},$$

and I_{11}, I_{12}, I_{22} are elements of the information matrix. In this Appendix, we apply this theory in order to obtain the null distribution of the likelihood ratio test of the variance components analysis when the dominance effect is considered.

Let the index $u:v$ denote the row corresponding to the pair (u,v) with $u<v$. In a variance components model, suppose a matrix B has $(u:v,l:m)$ element $(\Omega_0^{-1})_{ul}(\Omega_0^{-1})_{vm}+(\Omega_0^{-1})_{um}(\Omega_0^{-1})_{vl}$, where Ω_0 is the covariance matrix under the null hypothesis of no linkage. Then the information is

$$I_{11}=(\pi-2\Phi)'B(\pi-2\Phi)$$
$$I_{12}=(\kappa-\Delta)'B(\pi-2\Phi)$$
$$I_{22}=(\kappa-\Delta)'B(\kappa-\Delta)$$

Therefore, for a general pedigree, we have the following formula to calculate the mixing probabilities

$$p=\frac{1}{2\pi}\cos^{-1}\frac{(\kappa-\Delta)'B(\pi-2\Phi)}{\sqrt{(\pi-2\Phi)'B(\pi-2\Phi)\times(\kappa-\Delta)'B(\kappa-\Delta)}}. \quad (12)$$

For sibship data, since $\mathrm{Cov}[\pi_{ij},\pi_{lm}]=\mathrm{Cov}[\kappa_{ij},\kappa_{lm}]=0$ when $i\neq l$ or $j\neq m$, and diagonal elements of matrix B are identical, the B matrix can be canceled out, and thus (12) can be further simplified as

$$p=\frac{1}{2\pi}\cos^{-1}\frac{\sum_{i<j}E[(\pi_{ij}-1/2)(\kappa_{ij}-1/4)]}{\sqrt{\sum_{i<j}E[(\pi_{ij}-1/2)^2]\sum_{i<j}E[(\kappa_{ij}-1/4)^2]}}$$
$$=\frac{1}{2\pi}\cos^{-1}\sqrt{\frac{2}{3}}\approx 0.1.$$

The critical value corresponding to the 0.05 nominal level is 3.417. When the marker has two alleles with equal frequency, following a procedure similar to Wang and Huang [2002b], we have $p=0.083$ and critical value becomes 3.32. In a multipoint linkage analysis, markers tend to be much more informative. Therefore, a $0.4:0.5:0.1$ mixture of $\chi^2(0)$, $\chi^2(1)$, and $\chi^2(2)$ is a reasonable approximation for the null distribution of likelihood ratio test for sibship data. For fully informative markers and a mixture of sibships of various sizes, this mixing proportions would be the same as for a large sample of sibships of the same size. These mixing proportions have been obtained for an extension of score test [Wang and Huang, 2002b].

APPENDIX B

Here we show that the robust test statistic of Wang and Huang [2002a] may be expressed in matrix form and fit within the framework of our general GEE framework. Suppose $\omega_k=\Omega_k y_k$ and $b_k=\sum_{i<j}[\hat{\pi}_{kij}-\mathrm{E}(\hat{\pi}_{kij})][\omega_{ki}\omega_{kj}-\mathrm{E}(\omega_{ki}\omega_{kj})]$. Under the null hypothesis, the variance estimate for

$\sum_k b_k$ is $\overline{(\hat{\pi}\ldots - \mathrm{E}(\hat{\pi}\ldots))^2} \times \sum_k \sum_{i<j} (\omega_{ki}\omega_{kj} - \mathrm{E}(\omega_{ki}\omega_{kj}))^2$. Then the test statistic proposed by Wang and Huang [2002a] is $(\sum b_k)^2 / Var(\sum b_k)$ is identical to statistic (8) following the next two equalities:

$$
\begin{aligned}
&\sum_k D_k^{a'} (G_k^0)^{-1} S_k^0 \\
&= \sum_k \sum_{i<j} \sum_l [\hat{\pi}_{kij} - \mathrm{E}(\hat{\pi}_{kij})](\Omega_k^{-1})_{il}(\Omega_k^{-1})_{jl} \\
&\quad [y_{kl}^2 - \mathrm{E}(y_{kl}^2)] \\
&\quad + \sum_k \sum_{i<j} \sum_{l<m} [\hat{\pi}_{kij} - \mathrm{E}(\hat{\pi}_{kij})](\Omega_k^{-1})_{il}(\Omega_k^{-1})_{jm} \\
&\quad + (\Omega_k^{-1})_{im}(\Omega_k^{-1})_{jl})[y_{kl}y_{km} - \mathrm{E}(y_{kl}y_{km})] \\
&= \sum_k \sum_{i<j} [\hat{\pi}_{kij} - \mathrm{E}(\hat{\pi}_{kij})][(\Omega_k^{-1} y_k y_k' \Omega_k^{-1})_{ij} \\
&\quad - \mathrm{E}((\Omega_k^{-1} y_k y_k' \Omega_k^{-1})_{ij})] \\
&= \sum_k \sum_{i<j} [\hat{\pi}_{kij} - \mathrm{E}(\hat{\pi}_{kij})][(\omega_k \omega_k')_{ij} - \mathrm{E}((\omega_k \omega_k')_{ij})] \\
&= \sum_k \sum_{i<j} [\hat{\pi}_{kij} - \mathrm{E}(\hat{\pi}_{kij})][\omega_{ki}\omega_{kj} - \mathrm{E}(\omega_{ki}\omega_{kj})],
\end{aligned}
$$

and similarly

$$
\begin{aligned}
&\overline{(\hat{\pi}\ldots - \mathrm{E}(\hat{\pi}\ldots))^2} \times \sum_k \sum_{i<j} (\omega_{ki}\omega_{kj} - \mathrm{E}(\omega_{ki}\omega_{kj}))^2 \\
&= \overline{(\hat{\pi}\ldots - \mathrm{E}(\hat{\pi}\ldots))^2} \\
&\quad \times \sum_k \sum_{i<j} \left(S_k^{0'} (G_k^0)^{-1} \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} (G_k^0)^{-1} S_k^0 \right).
\end{aligned}
$$