Variance Components Linkage Analysis with Repeated Measurements

Liming Liang[a]   Wei-Min Chen[b, c]   Pak C. Sham[d]   Gonçalo R. Abecasis[a]

[a]Department of Biostatistics, University of Michigan, Ann Arbor, Mich., [b]Department of Public Health Sciences, Division of Biostatistics and Epidemiology, [c]Center for Public Health Genomics, University of Virginia, Charlottesville, Va., USA; [d]Department of Psychiatry and Genome Research Centre, the University of Hong Kong, Pokfulam, Hong Kong

Abstract

Background: When subjects are measured multiple times, linkage analysis needs to appropriately model these repeated measures. A number of methods have been proposed to model repeated measures in linkage analysis. Here, we focus on assessing the impact of repeated measures on the power and cost of a linkage study. **Methods:** We describe three alternative extensions of the variance components approach to accommodate repeated measures in a quantitative trait linkage study. We explicitly relate power and cost through the number of measures for different designs. Based on these models, we derive general formulas for optimal number of repeated measures for a given power or cost and use analytical calculations and simulations to compare power for different numbers of repeated measures across several scenarios. We give rigorous proof for the results under the balanced design. **Results:** Repeated measures substantially improve power and the proportional increase in LOD score depends mostly on measurement error and total heritability but not much on marker map, the number of alleles per marker or family structure. When measurement error takes up 20% of the trait variability and 4 measures/subject are taken, the proportional increase in LOD score ranges from 38% for traits with heritability of ~20% to 63% for traits with heritability of ~80%. An R package is provided to determine optimal number of repeated measures for given measurement error and cost. Variance component and regression based implementations of our methods are included in the MERLIN package to facilitate their use in practical studies.

Copyright © 2009 S. Karger AG, Basel

Introduction

In quantitative trait studies, taking repeated phenotype measures for each subject may increase the power. The approach is especially useful when measurement error is large or the relative cost of recruiting and genotyping additional subjects is high. It is important for a linkage analysis to appropriately take into account these repeated measures. Boomsma and Dolan [1] use structural equation modeling approach to analyze multivariate traits. Levy et al. [2] and de Andrade et al. [3] analyze longitudinal data by extending the standard variance components approach [4, 5]. Although in principle repeated measurements can be treated as multivariate traits or longitudinal data [6–8, 22], here we restrict our attention to modeling for repeated measurements for traits whose

Liming Liang
Department of Biostatistics, School of Public Health
1420 Washington Heights
Ann Arbor, MI 48109-2029 (USA)
Tel. +1 734 763 4901, Fax +1 734 615 8322, E-Mail lianglim@umich.edu

variance components do not change appreciably across time (except due to random measurement error). This allows us to focus the relationship between the power and cost of different study designs for quantitative trait linkage analysis and the number of repeated measures of the phenotype of interest taken for each subject. We also provide general implementations of these approaches, for both variance component [Amos et al.] and regression-based [Sham et al.] linkage analysis, in our MERLIN software package.

To analyze repeated measures, summary statistics such as the average of observed measurements are usually used to take advantage of the models and implementations designed for single measure. In this case, standard packages such as SOLAR [5] and MERLIN [9] can then be used to analyze the averaged measurements. Unfortunately, when different numbers of measures are available for each subject, this approach is invalid and likely to result in a loss of efficiency.

Here repeated measures are modeled explicitly and we use asymptotic theorems to explore the power of QTL linkage tests. Combining these theorems with a cost function that summarizes phenotyping, genotyping and general fixed costs, the optimal number of repeated measures and sample size can be determined for a proposed study.

We consider three analytical strategies: (a) a full model that explicitly incorporates all measurements for all subjects; (b) a simplified model that uses only the average phenotypic measurement and the number of measurements taken for each subject; and (c) a further simplified model that only considers the average phenotypic measurement for each subject. We find that repeated measures provide substantial power improvements across genetic models. The proportional increase in expected LOD score depends mostly on measurement error and total heritability but not much on marker map or number of alleles per marker. Given a fixed sample size, analysis of repeated measures can have a dramatic impact on power. For example, when measurement error takes up 20% of the trait variability and 4 measures per subject are taken, the proportional increase in expected LOD score ranges from 38% for traits with low heritability (e.g. 20%) to 63% for traits with high heritability (e.g. 80%). When 2 measures per subject are taken, the increase ranges from 23 to 36%. We identify the optimal number of repeated measures for different settings and show that when the number of measures is appropriately taken into account the average measure is a good balance between statistical power and computation efficiency.

Methods

In this section, we briefly review the variance component method for quantitative trait linkage analysis and then extend the model to accommodate repeated measures for arbitrary pedigrees, without inbreeding.

Variance Component Model
Let $Y = (y_1, ..., y_n)'$ be the vector of quantitative trait values for a pedigree with n subjects and no inbreeding. Y is assumed to follow a multivariate normal distribution with mean $\mu = (\mu_1, ..., \mu_n)'$ and variance-covariance matrix Ω. The effect of covariates can be modeled by letting $\mu = X\beta$, where X is the design matrix for covariates and β are the coefficients for each covariate.

In general, Ω will have the form:

$$\Omega = \sum_i \sigma_i^2 \Omega_i,$$

where σ_i^2 is a scalar variance component and Ω_i is the corresponding covariance structure matrix which depends on the effect σ_i^2 is representing. When major gene effect and polygenic effect are of interest, the Ω can be defined as: $\Omega = \Pi \sigma_{mg}^2 + 2\Phi \sigma_{pg}^2 + I_n \sigma_e^2$ where σ_{mg}^2 is the additive genetic variance due to the major gene; the element π_{ij} of Π is the proportion of alleles shared IBD at the test locus between subjects i and j; σ_{pg}^2 denotes the polygenic variance which is the genetic variance due to all residual additive effects not explained by the QTL; Φ is a matrix of genetic kinship coefficients; σ_e^2 is the subject-specific environmental variance and I_n is the $n \times n$ identity matrix [4, 5, 10, 11]. The model can be readily extended to include other effects of interest, such as genetic dominance.

The effects in variance component model can be assessed through likelihood ratio tests. For example, the test comparing $H_0 : \sigma_{mg}^2 = 0$ versus $H_1 : \sigma_{mg}^2 > 0$ is used to assess evidence for a major gene impacting the quantitative trait.

Full Model with Repeated Measures
Let Y_{ij} be the j-th measurement of the i-th subject for the quantitative phenotype of interest. Assume m_i repeated measures are taken for subject i. Then, let:

$$\begin{aligned} Var(Y_{ij}) &= \sigma_{mg}^2 + \sigma_{pg}^2 + \sigma_e^2 + \sigma_m^2 \quad i = 1, ..., n \quad j = 1, ..., m_i \\ Cov(Y_{ij_1}, Y_{ij_2}) &= \sigma_{mg}^2 + \sigma_{pg}^2 + \sigma_e^2 \quad\quad\quad\quad j_1 \neq j_2 \quad\quad (1) \\ Cov(Y_{i_1 j_1}, Y_{i_2 j_2}) &= \pi_{i_1 i_2} \sigma_{mg}^2 + 2\phi_{i_1 i_2} \sigma_{pg}^2 \quad\quad i_1 \neq i_2 \end{aligned}$$

Here, σ_m^2 represents the error specific to each measurement. This model is rather general. The covariance between repeated measurements of the same subject follows the compound symmetry structure [12]. This model is valid when measurement errors within a subject are (a) independent or (b) equally correlated. In the latter setting the correlation between measurements is absorbed by the σ_e^2 component.

Under the assumption of normality and because the variance-covariance structure of residuals does not involve the fixed effect parameters β, the distribution of the likelihood ratio statistics about a variance component does not depend on the fixed effects β [13]. Although our model assumes no time effect in the variance-covariance matrix, if the time effect were included as a fixed effect, the results of this paper remain unchanged. Longitudinal data can therefore be accommodated in this limited manner by specifying time dependent covariates as the fixed effects. For sim-

plicity and without loss of generality we assume the mean of quantitative trait is zero, with no covariate effects. Hence all the phenotypic variation can be explained through the similarity between relatives and the variance components σ_{mg}^2, σ_{pg}^2, σ_e^2 and σ_m^2.

Model for Average Measures

An alternative to using the model specified in (1) above is to use the average measurement for each subject [e.g. 2] instead of individual measurements. This approach results in smaller variance-covariance matrices and thus requires less computation.

Let

$$\bar{Y}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} Y_{ij}$$

be the average phenotype of subject i, for $i = 1, ..., n$. Using these averages, the model for the variances and covariances becomes:

$$\begin{aligned} Var(\bar{Y}_i) &= \sigma_{mg}^2 + \sigma_{pg}^2 + \sigma_e^2 + \sigma_m^2/m_i & i = 1, ..., n \\ Cov(\bar{Y}_{i_1}, \bar{Y}_{i_2}) &= \pi_{i_1 i_2} \sigma_{mg}^2 + 2\phi_{i_1 i_2} \sigma_{pg}^2 & i_1 \neq i_2 \end{aligned} \quad (2)$$

For balanced designs, where each subject has the same number of repeated measures, it can be shown that, although model (2) requires less computation, model (1) and (2) give identical estimates of genetic variance components (excluding the environmental and measurement error variance component, which are not identifiable) and lead to the same value for linkage test statistics. Details of the equivalence proof are given in the Appendix A.

Furthermore, when the number of repeated measures $m_i = m$ for all i the standard variance component model:

$$\begin{aligned} Var(\bar{Y}_i) &= \sigma_{mg}^2 + \sigma_{pg}^2 + \sigma_e^{2*} & i = 1, ..., n \\ Cov(\bar{Y}_{i_1}, \bar{Y}_{i_2}) &= \pi_{i_1 i_2} \sigma_{mg}^2 + 2\phi_{i_1 i_2} \sigma_{pg}^2 & i_1 \neq i_2 \end{aligned} \quad (3)$$

can be used to construct linkage test without loss of efficiency, where $\sigma_e^{2*} = \sigma_e^2 + \sigma_m^2/m$ and $m_i = m$ for $i = 1, ..., n$. Therefore, standard software packages for QTL linkage analysis can be used.

When m_i's are not all equal, as in unbalanced designs, the standard variance component model (3) is not valid because σ_e^{2*} will be different across subjects, potentially distorting estimates of the genetic variance components and test statistics. Model (2), which takes into account different numbers of measures for each subject, remains a valid model. Through simulation, we show that it is slightly less efficient than the full model (1).

Analytical NCP for Balanced Design

For simplicity we based our analytical calculation on the balanced design. Under general regularity conditions, classical properties of likelihood ratio tests can be used to calculate the power of the test for a given sample size or the sample size required to achieve a desired power [15].

Under the Null hypothesis when there is no major gene effect, the likelihood ratio test statistics is asymptotically distributed as

$$\frac{1}{2}\chi_1^2 : \frac{1}{2}(0),$$

a mixture of a chi-squared distribution with one degree of freedom and a unit point mass at zero. Under the alternative hypothesis, the likelihood ratio test statistics approximately follow a non-central chi-squared distribution with non-centrality parameter

$$\sum_{f=1}^{F} \delta_f,$$

where δ_f is the non-centrality contributed by the f-th of F families and

$$\begin{aligned} \delta_f = &\log \left| (\sigma_{mg}^2 + \sigma_{pg}^2) \, 2\boldsymbol{\Phi}_f + \sigma_e^{2*} \, \mathbf{I}_{n_f} \right| - \\ &E_\pi \log \left| \sigma_{mg}^2 \, \boldsymbol{\Pi}_f + \sigma_{pg}^2 \, 2\boldsymbol{\Phi}_f + \sigma_e^{2*} \, \mathbf{I}_{n_f} \right|. \end{aligned} \quad (4)$$

Here, n_f is the size of the f-th family and E_π denotes an expectation over all possible allele-sharing states that can be calculated by averaging over all possible inheritance vectors [14, 21]. The power of the test is then given by:

$$Power = Pr\left(\chi^2_{1, \sum_{f=1}^{F} \delta_f} > C_\alpha \right)$$

where

$$\chi^2_{1, \sum_{f=1}^{F} \delta_f}$$

follows a one degree of freedom chi-squared distribution with non-centrality parameter

$$\sum_{f=1}^{F} \delta_f$$

and C_α is the $100(1 - \alpha)$ percentile of

$$\frac{1}{2}\chi_1^2 : \frac{1}{2}(0).$$

To simplify the presentation, we consider F families with the same pedigree structure and denote $\delta = \delta_f$ for all f, so that

$$\sum_{f=1}^{F} \delta_f = F\delta.$$

For any desired power the required number of families F or of repeated measures m can then be solved numerically.

Cost-effectiveness

Formula (4) allows us to analytically compare power for different studies; each characterized by a specific family structure, the number of families examined, F, and the number of repeated measures, m, for each subject. To study cost-effectiveness of different designs, we first introduce a cost function for each design. Let:

C_0 = Fixed cost of the study
C_s = Cost per subject recruited and genotyped (total Fn subjects)
C_p = Cost per phenotype measurement (m measures per subject)
Total cost $C = C_0 + F \cdot n \cdot C_s + F \cdot n \cdot m \cdot C_p$

From the last section, we know that the power is determined by $F\delta$ the non-centrality parameter and that δ depends on m through σ_e^{2*}. We denote δ as $\delta(m)$.

For any two combinations of m and F: (m_1, F_1) and (m_2, F_2), maintaining the same power requires $\delta(m_1)F_1 = \delta(m_2)F_2$. Without loss of generality, we assume $m_1 > m_2$ so that $\delta(m_1) > \delta(m_2)$. The total costs for the first design and the second design are $C_0 + F_1 \cdot n \cdot C_s + F_1 \cdot n \cdot m_1 \cdot C_p$ and $C_0 + F_2 \cdot n \cdot C_s + F_2 \cdot n \cdot m_2 \cdot C_p$,

respectively. By simple algebra, taking m_1 (more) measures will provide the same power but a lower cost than taking m_2 (less) measures per subject when the following inequality holds:

$$\frac{C_s}{C_p} > CR_{m_1,m_2} \triangleq \frac{m_1 - m_2 \cdot \delta(m_1)/\delta(m_2)}{\delta(m_1)/\delta(m_2) - 1} \qquad (5)$$

CR_{m_1,m_2} defined above is called the break-event for cost ratio C_s/C_p, where taking m_1 measures is as cost-effective as taking m_2 measures per subject. When this cost ratio is higher (e.g. when phenotyping costs are relatively small compared to subject recruitment and genotyping costs), designs that take more measures per subject are favored.

Note that, for a given total cost (or power), the combination of m and F that maximizes power (or minimizes the total cost) can be identified numerically.

For unbalanced designs, CR can be approximated through simulation by using the ratio of expected LOD (ELOD) scores to replace $\delta(m_1)/\delta(m_2)$ in formula (5).

Simulation

We perform simulations to compare power for different numbers of repeated measures across several scenarios (varying distance between markers from ~0 to ~10 cM, considering SNP and microsatellite markers, and varying major gene heritability, total heritability and measurement error from 2 to 20%, 8 to 80% and 0 to 60% of trait variability, respectively).

For unbalanced designs, we attempted to mimic designs we have encountered in actual studies. For example, we simulated a situation where subjects with an extreme initial measurement were measured a second time. Thus, we first simulated one measurement for every subject. Next, we ordered subjects based on their simulated measurement and generated an additional measurement for $\alpha/2$ subjects at the top and $\alpha/2$ subjects at the bottom of the list. This design reflects the 'intuition' that it may be more fruitful to focus effort on measuring extreme subjects. In this design, the average number of measurements per subject is $1 + \alpha$. We let $\alpha = 20\%$ and $\alpha = 10\%$. In an alternative unbalanced design, referred to as the random design, the number of measures for each subject follows an exponential distribution. This mimics the situation where measurements are missing completely at random. For each subject, we draw independent random number (rounding to the nearest greater integer) from an exponential distribution with mean equal to 0.5, 1 and 2, respectively. The maximum number of measurements per subject was set to 4.

In each simulation, we simulated 1000 families and the results are based on 2000 simulations. The average of LOD scores at the QTL is used to estimate the ELOD. Power is measured by the proportion of likelihood ratio test p values <0.001. The cost-effectiveness break-event for cost ratio, CR_{m_1,m_2}, is also presented to facilitate comparison between different designs.

Results

Analytical Results

Based on the average model (formulas 2 and 4), we can examine the ELOD (hence the power) for different settings under the balanced design and assuming markers
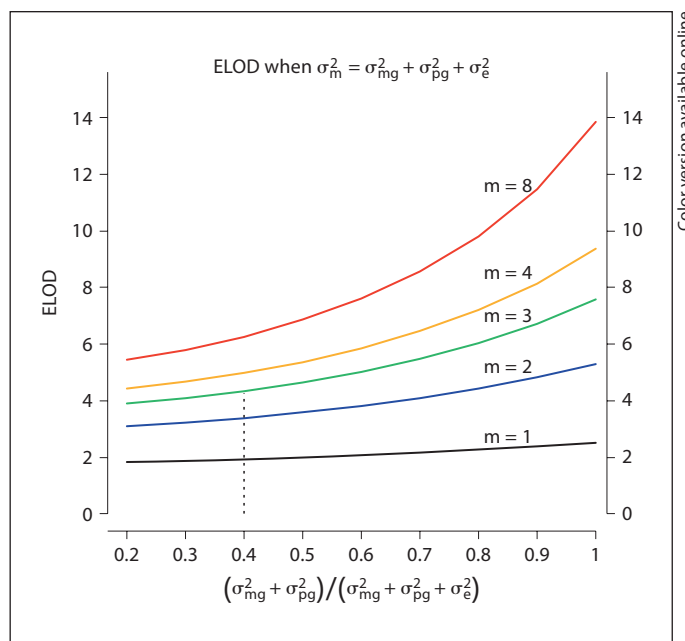

Fig. 1. Expected LOD score for 1000 nuclear families with 4 offspring, where $\sigma_{mg}^2 = 0.2$, $\sigma_{pg}^2 = 0, ..., 0.8$, $\sigma_e^2 = 0.8 - \sigma_{pg}^2$ and $\sigma_m^2 = \sigma_{mg}^2 + \sigma_{pg}^2 + \sigma_e^2 = 1$. m = the number of repeated measures.

are fully informative. Figure 1 shows how the ELOD changes as the heritability, defined as $(\sigma_{mg}^2 + \sigma_{pg}^2)/(\sigma_{mg}^2 + \sigma_{pg}^2 + \sigma_e^2)$, increases for different numbers of repeated measures. For example, when the heritability is 40%, increasing the number of measures from 1 to 3 results in a 2-fold increase in ELOD. We also note that taking more repeated measures results in more rapid increases in ELOD for simulated traits with greater heritability.

According to (5) we can determine the optimal number of repeated measures for different ratios of genotyping and phenotyping cost and degrees of measurement error. Figure 2 shows the contour plot for the optimal number of repeated measures when the cost ratio C_s/C_p ranges from 0.01 to 50 and measurement error variance ranges from 0.11 to 1.5 (corresponding to 10–60% of the total trait variance). For example, when measurement error variance is 0.4 (corresponding to 28.6% of the total trait variance), taking 2 measurements per subject is cost-effective if the cost ratio is between 1.11 and 4.17. When the ratio of genotyping and recruitment costs to phenotyping costs is <1.11, it is preferable to take a single measurement and collect more subjects. When this ratio is >4.17, it is preferable to take additional measurements and collect fewer subjects. When the cost ratio is between 9.09 and 15.62, taking 4 measurements per subject is the best. The
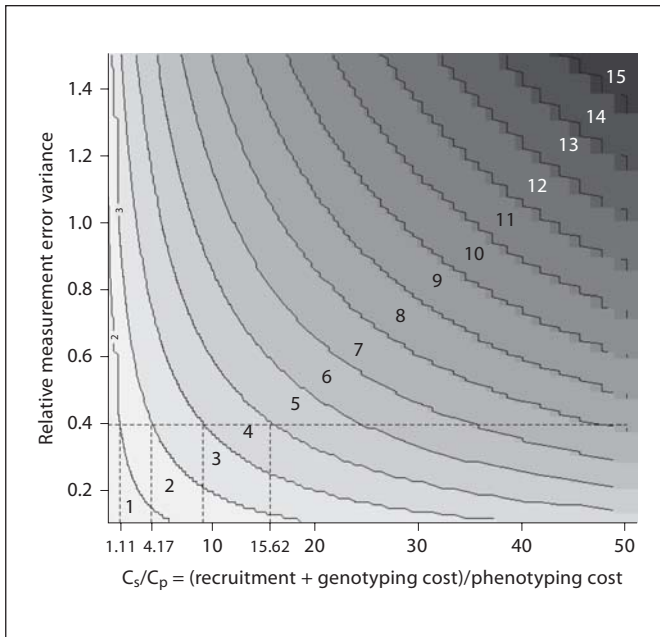
Fig. 2. Contour plot for optimal number of repeated measures when the cost ratio ranges from 0 to 50 and σ_m^2 ranges from 0.11 to 1.5 (10–60% of total trait variance). Trait variance excluding measurement error is fixed to 1 ($\sigma_{mg}^2 = 0.2$, $\sigma_{pg}^2 = 0.4$, $\sigma_e^2 = 0.4$). The numbers on the plot indicate the optimal number of repeated measures.

Fig. 3. Average LOD score profile for balanced design simulations (scenario 2). $\sigma_m^2 = 67$ (40% total variance), $\sigma_{pg}^2 = 40$ (24%). Results based on 500 simulation replications and plotted at every 1 Mb grid point.

ranges of figure 2 should include a variety of realistic scenarios. For example, chip based genotyping for genome-wide linkage studies typically costs a few hundred dollars per subject whereas phenotyping costs are widely variable, ranging from a few dollars per subject (for mail-in questionnaires [24]) to several hundred dollars (for expensive imaging measures or biological assays). The measurement error as well as the intra-individual environmental variance could range from very low (5%), for anthropometric measures such as height, to quite high (40%), for traits such as micro-array summaries of gene expression and questionnaire based assessments of personality.

Simulation Results

We simulated three scenarios: (1) One microsatellite marker with 20 alleles and 0 cM between the marker and the QTL to approximate a fully informative marker. (2) Ten microsatellite markers each with 4 alleles and with 10 cM separating consecutive markers; the QTL placed in the middle of the markers. (3) Fifty SNPs and 2 cM between consecutive markers; the QTL again placed in the middle of the SNPs. For each scenario, the trait variance excluding measurement error was fixed at 100, that is
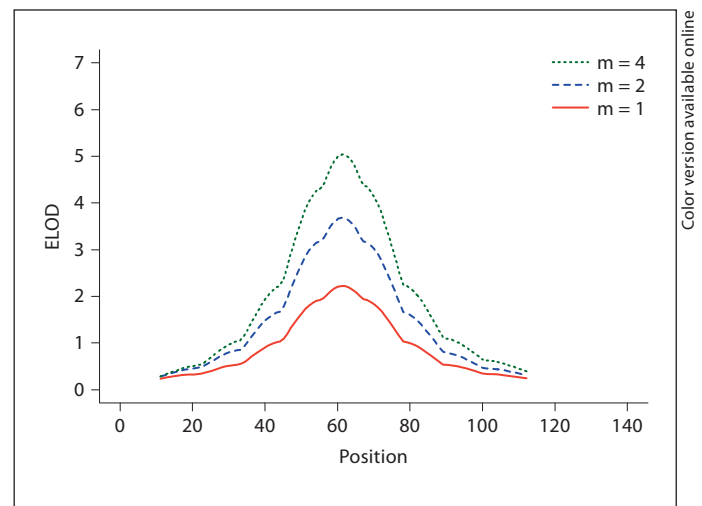
$\sigma_{mg}^2 + \sigma_{pg}^2 + \sigma_e^2 = 100$. The major gene effect σ_{mg}^2 was set at 20. Polygene effects σ_{pg}^2 ranged from 0 to 60. Measurement error variance σ_m^2 ranged from 11 to 150 (corresponding to 10–60% of the total trait variance). In each independent sample, we simulated 1000 nuclear families with 4 offspring each. Relative power for designs with different numbers of measurements varied only slightly for different family structures (online supplementary table 1, www.karger.com/doi/10.1159/000194977), which includes sibships with 2–6 siblings and cousin pedigrees) and so our presentation focuses on nuclear families with 4 off-spring.

Simulation results again show repeated measures can provide substantial power improvements (table 1, fig. 3). Table 1 shows the ELOD and power of balanced designs for a simulated microsatellite panel (scenario 2). Taking 2 measures per subject increases ELOD by 52% to 75% and power at $\alpha = 0.001$ by 63% to 78%. Figure 3 shows the average LOD score profile for the microsatellite panel (scenario 2, major gene effect 20 (or 12% total variance), polygene effect 40 (or 24% total variance), and measurement error 67 (or 40% total variance). In this case, taking 1 measure per subject results in an average peak LOD score of

Table 1. Power increment by taking repeated measures (scenario 2)

Polygene effect (% total var.)	No measurement error or M = ●●			M = 4			M = 2			M = 1	
	ELOD	power	ratio	ELOD	power	ratio	ELOD	power	ratio	ELOD	power
0.0 (0%)	4.88	0.94	2.71	3.58	0.80	1.99	2.74	0.64	1.52	1.80	0.36
0.2 (12%)	5.91	0.97	2.93	4.16	0.87	2.06	3.11	0.70	1.54	2.02	0.41
0.4 (24%)	7.48	1.00	3.35	5.01	0.94	2.25	3.63	0.80	1.63	2.23	0.48
0.6 (36%)	10.30	1.00	4.17	6.32	0.98	2.56	4.32	0.88	1.75	2.47	0.54

Measurement error variance = 67 (40% of the total trait variance). M = the number of repeated measures. The ratio is the ELOD ratio between M measures and 1 measures per subject. Scenario (2): Ten microsatellite markers each with 4 alleles and spaced 10 cM apart; the QTL placed in the middle of the markers.

Table 2. Cost-effectiveness analysis for 4 repeated measures vs. 1 measure

Measurement error var. (% total var.)	Heritability (% total var.)	Ave ELOD ratio	Sample size savings	$CR_{4,1}$	$CR_{4,2}$
11 (10%)	0.20 (18%)	1.17	0.15	16.31	31.20
	0.60 (54%)	1.23	0.19	12.04	26.75
25 (20%)	0.20 (16%)	1.38	0.28	6.83	14.44
	0.60 (48%)	1.51	0.34	4.84	10.41
67 (40%)	0.20 (12%)	2.01	0.50	1.97	4.55
	0.60 (36%)	2.28	0.56	1.34	3.33
150 (60%)	0.20 (8%)	3.07	0.67	0.45	1.33
	0.60 (24%)	3.39	0.71	0.25	0.78

CR_{m_1,m_2} is defined in (5). $CR_{4,2}$ is also listed here for comparison purpose. When $C_s/C_p > CR_{m_1,m_2}$, taking m_1 measures is better than taking m_2 measures per subject. Heritability is defined as $(\sigma_{mg}^2 + \sigma_{pg}^2)/(\sigma_{mg}^2 + \sigma_{pg}^2 + \sigma_e^2)$ where $\sigma_{mg}^2 + \sigma_{pg}^2 + \sigma_e^2 = 100$ and the major gene effect s_{mg}^2 is fixed to 20. Average of ELOD ratio is the average across three scenarios that give similar results: (1) a highly informative microsatellite marker with 20 alleles and 0 cM between the marker and the QTL. (2) Ten microsatellite markers each with 4 alleles and spaced 10 cM apart; the QTL placed in the middle of the markers. (3) Fifty SNPs spaced 2 cM apart; the QTL again placed in the middle of the SNPs.

only 2.22. Taking repeated measures increases the average peak LOD to 3.69 (2 measures) and 5.04 (4 measures).

Since IBD estimation does not affect the accuracy of estimates of measurement error variance, the proportional increase in expected LOD score (ELOD ratio) depends mostly on measurement error and total heritability but not much on marker map or number of alleles per marker (table 2), which mostly impact the precision of QTL effect size estimates. This suggests that the optimal design (in terms of optimal number of repeated measures) is relatively insensitive to the genotyping platform selected. Table 2 shows the average ELOD ratios for 4 repeated measures under three scenarios. Based on the ELOD ratio and the condition to maintain the same power, $\delta(m_1)F_1 = \delta(m_2)F_2$, we can calculate the savings in sample size when using 4 repeated measures. For example, for the first setting when measurement error variance is 11 (10% total variance) and total heritability is 20%, the sample size (number of subjects) required when taking 4 measures per subject is 85% (1/1.17) of the sample size required when taking 1 measure per subject.

When the ELOD ratios are available, it is possible to calculate the break-event CR_{m_1,m_2} for cost ratio C_s/C_p using (5). For example, when measurement error variance is 25 (20% of the total variance) and heritability is 20%, if genotyping and recruitment costs per subject are more than 6.83 higher than phenotyping costs, taking 4 measures per subject is more cost-effective than taking 1 measure per subject. The cost ratio needs to exceed 14.44 so that taking 4 measures is better than taking 2 measures per subject (table 2).

For the unbalanced design where 20% (or 10%) of subjects with an extreme first measurement are measured one more time, the cost ratios can be calculated in a similar way because the total number of measures is fixed. We denote these two designs as 'm = 1.2' and 'm = 1.1' respectively. Now we can compare different designs using the cost ratio CR_{m_1,m_2}. The results are summarized in table 3. The cost ratio $CR_{1,1,1}$ is relative large, $CR_{1,1,1} = ●●$ in the first row means designs with m = 1.1 are never more cost-effective than taking one measure per subject, when the measurement error is small. Note that since $CR_{2,1.2} <$
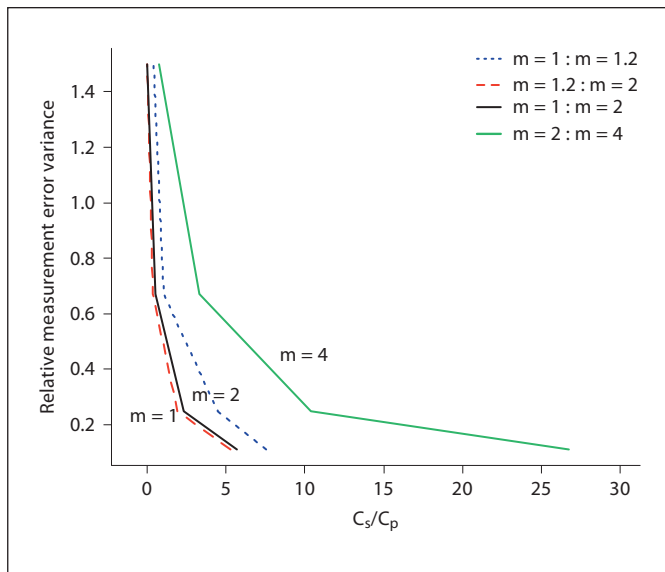
Fig. 4. Contour plot of optimal number of repeated measures when the cost ratio ranges from 0 to 30 and σ_m^2 ranges from 11 to 150 (10–60% total variance). Trait variance excluding measurement error is fixed to 100 ($\sigma_{mg}^2 = 20$, $\sigma_{pg}^2 = 40$, $\sigma_e^2 = 40$). This setting is equivalent to the setting in figure 2. Each line separates two regions in which one design is better than the other. For example, to the left of the (blue) dot line, balanced design m = 1 is better than the unbalanced design m = 1.2; on the right side of the line, the unbalanced design m = 1.2 is better than balanced design m = 1. Note that the (blue) dot line is to the right of the (red) dash line, thus balanced designs are superior to unbalanced designs in any situation. For the region to the right of the grey (green) solid line, the optimal design is balanced design m = 4; for the region between the black solid line and the grey (green) solid line, the optimal design is balanced design m = 2; for the region to the left of the black solid line, the optimal design is balanced design m = 1.

Fig. 5. ELOD ratio of full model vs. average model for unbalanced design. Setting is scenario 2. Left 4 pairs of bars are for $\sigma_m^2 = 11$ (10% of total variance). Right 4 pairs of bars are for $\sigma_m^2 = 150$ (60% of total variance). Random design: the number of repeated measures follows an exponential distribution. Extreme design: 20% subjects with extreme first measure have an additional measurement.

$CR_{1.2,1}$ and $CR_{2,1,1} < CR_{1,1,1}$, the unbalanced designs can always be outperformed by a balanced design that involves either 2 or 1 measures per subject depending on the cost ratio C_s/C_p. So in terms of cost-effectiveness, balanced designs are always better than unbalanced designs no matter what the cost ratio C_s/C_p. Using the data in table 3, we can draw a similar contour plot as figure 2. This plot is presented in figure 4. The parameter settings are equivalent to figure 2. The plot shows the theoretical result (fig. 2) is consistent with the simulation result (fig. 4).

Comparing Efficiency between the Full Model and the Average Model

Using simulation, we next compared the efficiency between the full model (1) and the average model (2) for unbalanced designs. Both models take into account the
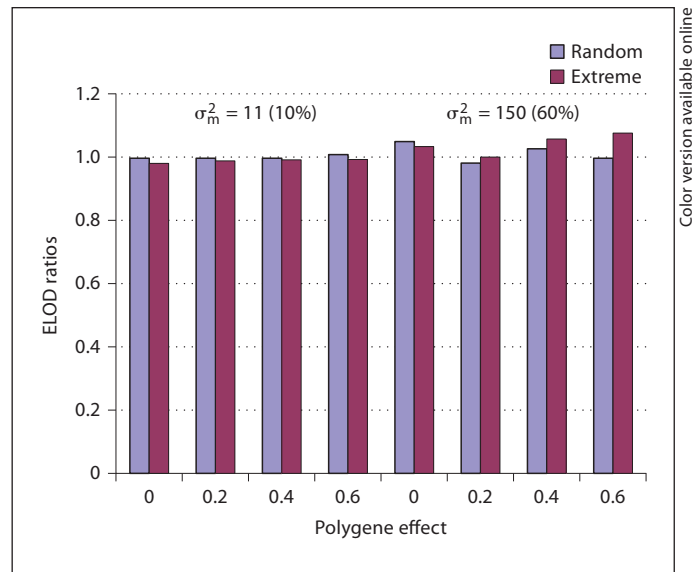
different number of measurements across individuals, give valid likelihood functions and control type I error rate adequately.

Figure 5 shows the ELOD ratio of the full model vs. the average model. For both unbalanced designs, the full model did not provide substantially more efficiency than the average model (only in the extreme design, the full model increases ELOD by 1% on average across all scenarios). The largest increase in ELOD was 9% in settings where the measurement error was large and individuals with an initial extreme measurement were reassessed.

Discussion

When subjects are measured multiple times, it is important for a linkage analysis to appropriately take into account these repeated measures. In this study, we extend the variance components approach to model repeated measures in a quantitative trait linkage study. Our model can explicitly relate the power and cost of different sampling designs. We give the general formulas of optimal sample size and number of repeated measures for a given power or cost. We show for the case of a balanced

Table 3. Cost ratios for the comparison between different designs

Measurement error var.	Heritability (% total var.)	$CR_{4,2}$	$CR_{2,1}$	$CR_{2,1,2}$	$CR_{2,1,1}$	$CR_{1,2,1}$	$CR_{1,1,1}$
11 (10%)	0.20 (18%)	31.20	8.38	7.16	7.34	19.00	∞
	0.60 (54%)	26.75	5.67	5.26	5.22	7.57	14.00
25 (20%)	0.20 (16%)	14.44	3.29	2.77	2.97	6.50	9.00
	0.60 (48%)	10.41	2.30	1.91	2.00	4.45	9.00
67 (40%)	0.20 (12%)	4.55	0.85	0.53	0.65	2.75	5.00
	0.60 (36%)	3.33	0.52	0.36	0.38	1.07	2.00
150 (60%)	0.20 (8%)	1.33	0.09	0	0	0.76	1.31
	0.60 (24%)	0.78	0.03	0	0	0.40	0.43

CR_{m_1,m_2} is defined in (5). When $C_s/C_p > CR_{m_1,m_2}$, taking m_1 measures is better than taking m_2 measures per subject. Heritability is defined as $(\sigma^2_{mg} + \sigma^2_{pg})/(\sigma^2_{mg} + \sigma^2_{pg} + \sigma^2_e)$.

design where the same number of measurements is taken for each subject, a standard linkage test that takes the average of measures as the trait of interest is identical to a linkage test based on an appropriate extension of the variance components model.

In our model, the covariance between repeated measures of the same subject follows the compound symmetry structure. This model is valid when measurement errors within a subject are either independent or else equally correlated. It is one of the most commonly used covariance structures in the repeated measures literature. When necessary it should be possible to refine our model to include dominance effects, twin environment or other variance-covariance components or even to incorporate covariate effects into the variance-covariance matrix. In particular, time effects can be introduced into the variance-covariance structure to allow for longitudinal changes in variance components [3].

Through both analytical calculation and simulation, we find that repeated measures provide substantial power improvements across genetic models. The proportional increase in expected LOD score (ELOD Ratio) depends mostly on measurement error and total heritability but not much on marker map or number of alleles per marker. This suggests that the optimal design (in terms of optimal number of repeated measures) will be similar for a range of genotyping strategies (provided they are similar in cost). We give contour plots to help investigators decide on the optimal number of repeated measures for different levels of measurement errors and ratios of genotyping, subject recruitment and phenotyping costs. The R code to help determine the optimal number of repeated measures is available from our website.

Precise trade-offs can be obtained by examining figure 2 and the R package. Still, our results allow us to make some general recommendations. When measurement error is high, accounting for ∼50% of the trait variance, it is typically cost effective to collect 2 or more measures per subject when the ratio of phenotyping to genotyping costs per subject is <16 fold. If genotyping is carried out using a commercially available SNP array that typically costs 100–200 USD per subject, it will almost always be worthwhile to phenotype each individual multiple times, given that most phenotyping assays cost <1600–3200 USD per measurement. When measurement error is small, accounting for ∼10% of the trait variance, it is only cost effective to collect 2 or more measures per subject when phenotyping is relatively inexpensive, costing no more than 0.154 times the cost of genotyping. With the same genotyping costs as above, this would correspond to 15–30 USD per measurement and would only be worthwhile for the most inexpensive phenotypes (such as those that rely on mail-in questionnaires or very simple trait measurements). In other situations, it will be more efficient to collect additional subjects.

For unbalanced designs, a standard linkage test that takes the average measurement as the trait of interest and ignores the number of measures is not valid. A model that uses the average measurement as the trait but takes into account the different number of measures for each subject, i.e. model (2), is a valid alternative to the full model. The advantage of model (2) is that it is less computationally intensive and, typically, only slightly less powerful than the full model. We implemented both the average model and the full model in the MERLIN package [9, 23]. We also assessed the effect of ignoring the imbalance and

taking the average as a single trait. Online suppl. table 2 shows simulations where a random half of subjects were measured 2, 4, or 10 times while the other half were measured only once. The results suggest that ignoring imbalance could lead to approximately correct Type I error but could lose power (at $p < 0.001$) by 2–5% or decrease in ELOD by 3–15%.

In our simulations, parental genotypes were used to help estimate IBD sharing between pairs of relatives. We also investigated the effect of parental phenotypes on power. Online suppl. fig. 1 shows the expected LOD scores with and without using parental phenotypes at a fully informative marker under the same scenario of figure 1. For a simulated trait with relatively low heritability, the additional measures from parents only slightly increase the expected LOD scores, suggesting that phenotyping parents is unlikely to be cost effective. For highly heritable traits, parental phenotypes do substantially increase the expected LOD scores especially for larger number of repeated measures. In this case, there will be a trade-off between phenotyping the parents and collecting more offspring genotypes and phenotypes.

In cases of non-normality of the trait distribution and selected sampling, robust statistics such as score statistics [16, 17] or regression-based statistics [18] can help to adequately control the type I error and increase power. Intensive simulations [17, 18] have shown that the regression-based model implemented in MERLIN-REGRESS [18] is robust to violations of normality, selected sampling and population parameter misspecification while achieving high power. Nash et al. 2004 discussed the treatment of average repeated measures in the regression-based model [20]. We take another approach which leads to simpler formulation and hence easier implementation of the software. We show that the regression-based model can be extended to incorporate individual repeated measures as well as average measures [appendix B] and this alternative is implemented in MERLIN-REGRESS [18].

Appendix A: Equivalence between Full Model and Average Measurement Model for Balanced Number of Measurements

When each subject is measured the same number M of times, it can be shown that the full model (1) and average measurement model (2 and 3) are equivalent.

Let vector Y_j represent all measurements for subject j. In full model (1), the variance for vector Y_j is $Var(Y_j) = (\sigma_{mg}^2 + \sigma_{pg}^2 + \sigma_e^2)\mathbf{1}\mathbf{1}' + \sigma_m^2\mathbf{I}$ and the covariance between Y_j for subject j and Y_k for subject k is $Cov(Y_j,Y_k) = (\pi_{jk}\sigma_{mg}^2 + 2\phi_{jk}\sigma_{pg}^2)\mathbf{1}\mathbf{1}'$, where vector $\mathbf{1}$ consists of all 1's and $\mathbf{I}$ is the identity matrix.

We first apply a linear transformation $\mathbf{T}$ on multiple measurements Y_j

$$Y_j^* = \mathbf{T}(Y_{j1}, Y_{j2}, ..., Y_{jM})' \tag{A1}$$

where

$$\mathbf{T} = \begin{pmatrix} \dfrac{1}{M} & \dfrac{1}{M} & \cdots & \dfrac{1}{M} & \dfrac{1}{M} \\ -\dfrac{1}{2} & \dfrac{1}{2} & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \dfrac{1}{2} & 0 \\ 0 & 0 & \cdots & -\dfrac{1}{2} & \dfrac{1}{2} \end{pmatrix}.$$

Thus the covariance matrix for the transformed vector Y_j^* is

$$\begin{aligned} Var(Y_j^*) &= \mathbf{T}Var(Y_j)\mathbf{T}' \\ &= (\sigma_{mg}^2 + \sigma_{pg}^2 + \sigma_e^2)\mathbf{T}\mathbf{1}\mathbf{1}'\mathbf{T}' + \sigma_m^2\mathbf{T}\mathbf{T}' \end{aligned} \tag{A2}$$

$$\begin{aligned} Cov(Y_j^*,Y_k^*) &= \mathbf{T}Cov(Y_j,Y_k)\mathbf{T}' \\ &= (\pi_{jk}\sigma_{mg}^2 + 2\phi_{jk}\sigma_{pg}^2)\mathbf{T}\mathbf{1}\mathbf{1}'\mathbf{T}' \end{aligned} \tag{A3}$$

Simple algebra gives

$$\mathbf{T}\mathbf{1}\mathbf{1}'\mathbf{T}' = \begin{pmatrix} 1 & 0 \\ 0 & 0_{(M-1)\times(M-1)} \end{pmatrix}, \qquad \mathbf{T}\mathbf{T}' = \begin{pmatrix} 1/M & 0 \\ 0 & \mathbf{A}_{(M-1)\times(M-1)} \end{pmatrix}$$

where $\mathbf{A}$ is some $(M-1)$ by $(M-1)$ matrix. Let Z_j and Y_{jm}^* denote the first and the rest of the elements in the transformed vector Y_j^* respectively. Then,

$$Z_j = \sum_{m=1}^{M} Y_{jm}/M,$$

and

$$Y_{jm}^* = \frac{Y_{jm} - Y_{j(m-1)}}{2}$$

for $m = 2, …, M$. Covariances (A2) and (A3) imply

$$Var(Z_j) = \sigma_{mg}^2 + \sigma_{pg}^2 + (\sigma_e^2 + \sigma_m^2/M)$$
$$Cov(Z_j, Z_k) = \pi_{jk}\sigma_{mg}^2 + 2\phi_{jk}\sigma_{pg}^2, \text{ for } j \neq k$$
$$Var(Y_{jm}^*) = \sigma_m^2/2, \text{ for } m = 2, ..., M$$
$$Cov(Y_{jm}^*, Y_{j(m+1)}^*) = \sigma_m^2/4, \text{ for } m = 2, ..., M-1 \qquad (A4)$$

and all other covariances are 0. Thus, the full model (1) implies model (A4). The reverse is also true since the transformation (A1) is not singular. Now we assume in the average model (3), $\sigma_e^{2*} = \sigma_e^2 + \sigma_m^2/M$. By comparing model (A4) and model (3), we can see model (A4) implies model (3).

Let $\mathbf{\Omega}_Z$ denote the variance-covariance matrix of vector $z = (Z_1, ..., Z_J)'$, and $\mathbf{\Omega}_*$ denotes the variance-covariance matrix of $y^* = (y_2^*, ..., y_M^*)'$, where $y_m^* = (Y_{1m}^*, ..., Y_{Jm}^*)'$ for $m = 2, …, M$. Model (A4) shows vector z and y_m^* are orthogonal, indicating the variance-covariance matrix of $(z, y^*)'$ with form

$$\begin{pmatrix} \mathbf{\Omega}_Z & 0 \\ 0 & \mathbf{\Omega}_* \end{pmatrix}.$$

Thus, the likelihood of a family is

$$L = \frac{1}{(\sqrt{2\pi})^{M*J}} \begin{vmatrix} \mathbf{\Omega}_Z & 0 \\ 0 & \mathbf{\Omega}_* \end{vmatrix}^{-1} \exp\left\{-\frac{1}{2}(z', y^{*\prime})' \begin{pmatrix} \mathbf{\Omega}_Z^{-1} & 0 \\ 0 & \mathbf{\Omega}_*^{-1} \end{pmatrix} \begin{pmatrix} z \\ y^* \end{pmatrix}\right\}$$

$$= \frac{1}{(\sqrt{2\pi})^J} \frac{1}{|\mathbf{\Omega}_Z|} \exp\left\{-\frac{1}{2} z' \, \mathbf{\Omega}_Z^{-1} z\right\}$$

$$\times \frac{1}{(\sqrt{2\pi})^{(M-1)*J}} \frac{1}{|\mathbf{\Omega}_*|} \exp\left\{-\frac{1}{2} y^{*\prime}\mathbf{\Omega}_*^{-1} y^*\right\}.$$

The first part of the above likelihood is exactly the likelihood in model (3). Since the second part in the last expression of the likelihood only contains information about σ_m^2 and does not carry any information about σ_{mg}^2, σ_{pg}^2 and σ_e^{2*}, the maximum likelihood estimates about $(\sigma_{mg}^2, \sigma_{pg}^2, \sigma_e^2)$ in the average measurement model are identical to those for $(\sigma_{mg}^2, \sigma_{pg}^2, \sigma_e^{2*})$ in the full model. Therefore, for balanced data, the average of the repeated measurements can be treated as the actual trait, and the standard variance components analysis is the equivalent to the full model. When the number of measurements is not balanced, the equivalence between the above two models does not hold anymore and the full model uses more information than the average measurement model.

Appendix B: Extension of the Regression Model for Linkage Analysis in Sham et al. 2002 [18] to Accommodate Repeated Measures

To incorporate repeated measures into the regression model, we only need to re-specify the form for the expectation and covariance that involves the squared sum S and squared difference D, other terms in the model will be identical to Sham et al. 2002. In fact, the regression model can be extended to model individual measures as well as the average measures and the relative performance of models using all available measurements, the average measurement and the count of measurements for each subject, or just the average measurement is analogous to the performance of formulas (1)–(3) in the variance component model.

Let c be the within-subject correlation

$$\frac{\sigma_{mg}^2 + \sigma_{pg}^2 + \sigma_e^2}{\sigma_{mg}^2 + \sigma_{pg}^2 + \sigma_e^2 + \sigma_m^2}$$

and H^2 be the total heritability

$$\frac{\sigma_{mg}^2 + \sigma_{pg}^2}{\sigma_{mg}^2 + \sigma_{pg}^2 + \sigma_e^2 + \sigma_m^2}.$$

Assuming the full model (1), all pairs of individual measures standardized by their population mean μ and variance $\sigma^2 = \sigma_{mg}^2 + \sigma_{pg}^2 + \sigma_e^2 + \sigma_m^2$ are considered. The vector of squared sums is

$$\left[S_{i_1 j_1, i_2 j_2} = \left(\frac{Y_{i_1 j_1} - \mu}{\sigma} + \frac{Y_{i_2 j_2} - \mu}{\sigma}\right)^2\right]$$

and similarly the vector of squared differences is

$$\left[D_{i_1 j_1, i_2 j_2} = \left(\frac{Y_{i_1 j_1} - \mu}{\sigma} - \frac{Y_{i_2 j_2} - \mu}{\sigma}\right)^2\right].$$

In the expectation and covariance of the squared sums S and squared differences D, only the form of correlation needs to be changed and it is equal to:

$$r_{i_1 j_1, i_2 j_2} = cov\left(\frac{Y_{i_1 j_1} - \mu}{\sigma}, \frac{Y_{i_2 j_2} - \mu}{\sigma}\right) = \begin{cases} c & \text{if } i_1 = i_2 \\ 2\phi_{i_1 i_2} H^2 & \text{otherwise} \end{cases}$$

The parameter c as well as the population mean, μ, variance, σ^2, and total heritability H^2 will need to be specified by the user.

The remaining terms that need to be considered are the covariance between S, D and $\hat{\pi}$: $\{Cov_1(S_{i_1 j_1, i_2 j_2}, \hat{\pi}_{k_1 l_1, k_2 l_2})\}$ and $\{Cov_1(D_{i_1 j_1, i_2 j_2}, \hat{\pi}_{k_1 l_1, k_2 l_2})\}$. For $i_1 \neq i_2$ and $k_1 \neq k_2$, these terms remain unchanged. For $k_1 = k_2$, $\hat{\pi}_{k_1 l_1, k_2 l_2} = 1$ so the covariance is 0. For $i_1 = i_2$ and $k_1 \neq k_2$, since the joint distribution of $(Y_{i_1 j_1}, Y_{i_2 j_2})$ does not involve π, the covariance is again 0. This suggests that we only need to include the pair of measures that involve different subjects; greatly reducing the dimension of mean vectors and covariance matrixes. More importantly, all formulas in Sham et al. 2002 can be directly applied if we only include pairs of measures that are from different subjects.

Assuming model (2) for average measures under unbalanced designs, the variance for each average measure will be different. Unlike the treatment in [20], we propose to standardize the average measures $\{\bar{Y}_i\}$ by the population mean μ and their own variances $\{\sigma_i^2 = \sigma_{mg}^2 + \sigma_{pg}^2 + \sigma_e^2 + \sigma_m^2/m_i = c\sigma^2 + (1 - c)\sigma^2/m_i\}$ so that they are multivariate normal with mean 0 and variance 1 and results in Appendix A of Sham et al. 2002 can apply. Hence the formulae for covariances of the squared sums S and squared differences D remain unchanged. Only the correlation between a pair of standardized average measures needs to be changed to:

$$r_{ij} = cov\left(\frac{\bar{Y}_i - \mu}{\sigma_i}, \frac{\bar{Y}_j - \mu}{\sigma_j}\right) = 2\phi_{ij} H^2 \frac{\sigma^2}{\sigma_i \sigma_j}$$

For covariance between S, D and $\hat{\pi}$, following a similar derivation to Drigalenko 1998 [19], we have:

$$cov_1\left(S_{ij}, \hat{\pi}_{kl}\right) = \frac{2Q\,cov_1\left(\hat{\pi}_{ij}, \hat{\pi}_{kl}\right)}{\sigma_i \sigma_j}$$

and

$$\mathrm{cov}_1\left(D_{ij}, \hat{\pi}_{kl}\right) = \frac{-2Q\,\mathrm{cov}_1\left(\hat{\pi}_{ij}, \hat{\pi}_{kl}\right)}{\sigma_i \sigma_j}$$

Other equations will be identical to Sham et al. 2002.

Analogous to model (3) for average measures with balanced designs, the average measures $\{\bar{Y}_i, i = 1 \dots n\}$ can be treated as an actual trait and standardized by the population mean μ and the variance $\sigma_{mg}^2 + \sigma_{pg}^2 + \sigma_e^2 + \sigma_m^2/m = c\sigma^2 + (1-c)\sigma^2/m$. So the model in Sham et al. 2002 can apply.

References

1 Boomsma DI, Dolan CV: A comparison of power to detect a QTL in sib-pair data using multivariate phenotypes, mean phenotypes, and factor scores. Behav Genet 1998;28:329–340.

2 Levy D, DeStefano AL, Larson MG, O'Donnell CJ, Lifton RP, Gavras H, Cupples LA, Myers RH: Evidence for a gene influencing blood pressure on chromosome 17. Genome scan linkage results for longitudinal blood pressure phenotypes in subjects from the Framingham heart study. Hypertension 2000;36:477–483.

3 de Andrade M, Gueguen R, Visvikis S, Sass C, Siest G, Amos CI: Extension of variance components approach to incorporate temporal trends and longitudinal pedigree data analysis. Genet Epidemiol 2002;22:221–232.

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

5 Almasy L, Blangero J: Multipoint quantitative trait linkage analysis in general pedigree. Am J Hum Genet 1998;62:1198–2211.

6 Gauderman WJ, Macgregor S, Briollais L, Scurrah K, Tobin M, Park T, Wang D, Rao S, John S, Bull S: Longitudinal data analysis in pedigree analysis. Genet Epidemiol 2003;25:S18–S28.

7 Schmitz S, Cherny SS, Fulker DW: Increase in power through multivariate analysis. Behav Genet 1998;28:357–363.

8 Boomsma DI, Molenaar PCM: The genetic analysis of repeated measures. I. Simplex models. Behav Genet 1987;17:111–123.

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

10 Lange K, Westlake J, Spence MA: Extensions to pedigree analysis. III. Variance components by the scoring method. Ann Hum Genet 1976;39:485–491.

11 Jacquard A. Genetic information given by a relative. Biometrics 1972;28:1101–1114.

12 Maxwell SE, Delaney HD: Designing Experiments and Analyzing Data: A Model Comparison Perspective, ed 2. Mahwah, NJ, Lawrence Erlbaum, 2003, pp 525–572.

13 Wald A: Tests of statistical hypotheses concerning several parameters when the number of observations is large. Trans Am Math Soc 1943;54:426–482.

14 Rijsdijk FV, Hewitt JK, Sham PC: Analytic power calculation for QTL linkage analysis of small pedigrees. Eur J Hum Genet. 2001;9:335–340.

15 Chen WM, Abecasis GR: Estimating the power of variance component linkage analysis in large pedigrees. Genet Epidemiol 2006;30:471–484.

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

17 Bhattacharjee S, Kuo CL, Mukhopadhyay N, Brock GN, Weeks DE, Feingold E: Robust score statistics for QTL linkage analysis. Am J Hum Genet 2008;82:1–16.

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

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

20 Nash MW, Huezo-Diaz P, Williamson RJ, Sterne A, Purcell S, Hoda F, Cherny SS, Abecasis GR, Prince M, Gray JA, Ball D, Asherson P, Mann A, Goldberg D, McGuffin P, Farmer A, Plomin R, Craig IW, Sham PC: Genome-wide linkage analysis of a composite index of neuroticism and mood-related scales in extreme selected sibships. Hum Mol Genet 2004;13:2173–2182.

21 Sham PC, Cherny SS, Purcell S, Hewitt JK: Power of Linkage versus Association Analysis of Quantitative Traits, by Use of Variance-Components Models, for Sibship Data. Am J Hum Genet 2000;66:1616–1630.

22 Bauman LE, Almasy L, Blangero J, Duggirala R, Sinsheimer JS, Lange K: Fishing for pleiotropic QTLs in a polygenic sea. Ann Hum Genet 2005;69:590–611.

23 Abecasis GR, Wigginton JE: Handling marker-marker linkage disequilibrium: pedigree analysis with clustered markers. Am J Hum Genet 2005;77:754–767.

24 Fullerton J, et al: Linkage analysis of extremely discordant and concordant sibling pairs identifies quantitative-trait loci that influence variation in the human personality trait neuroticism. Am J Hum Genet 2003;72:879–890.