Heredity **82** (1999) 668–676 *Received 9 November 1998, accepted 9 February 1999*

Mapping quantitative trait loci for complex binary traits in outbred populations

NENGJUN YI & SHIZHONG XU*

Department of Botany and Plant Sciences, University of California, Riverside, CA 92521-0124, U.S.A.

Complex binary traits have a dichotomous phenotypic expression but do not show a simple Mendelian segregation ratio. These traits are considered to be jointly controlled by the actions of several genes and a random environmental effect. The binary phenotype and the underlying factor are assumed to be linked through a threshold model. The underlying factor, referred to as the liability, is treated as a regular but unobservable quantitative character. Mapping quantitative trait loci (QTL) can be performed directly on the liability. Methods of QTL mapping for the liability of a complex binary trait have been well developed in line-crossing experiments. However, such a method is not available in outbred populations which usually consist of many independent pedigrees (families). In this study, we develop a method to analyse jointly multiple families of an outbred population. The method is developed based on a fixed-model approach, i.e. the QTL effects, rather than the variance, are estimated and tested. After the test, the estimated effects are then converted into a single estimate of the QTL variance by taking into consideration errors in the estimated effects. The QTL effects and variance–covariance matrix of the estimates are obtained by a fast Fisher-scoring method. Monte Carlo simulations show that the method is not only powerful but also generates very accurate estimates of QTL variances.

Keywords: binary trait, Fisher-scoring, Monte Carlo simulation, QTL mapping, threshold model.

Introduction

Many characters of biological interest and economic importance vary in a dichotomous form, i.e. presence or absence, but are not inherited in a simple Mendelian fashion. These traits are called complex binary traits. Complex binary traits are presumably controlled by a number of genetic and environmental factors. Because of this, these traits belong to the category of quantitative traits (Falconer & Mackay, 1996).

Complex binary traits are usually analysed using a threshold model, in which it is assumed that the observed category is determined by the value of an underlying unobservable continuous variable (Harville & Mee, 1984; McCulloch, 1994). The underlying continuous variable, called the liability, can be considered as a regular quantitative trait which can be partitioned into genetic and environmental components. The binary phenotype and the continuous liability are linked through a fixed but unknown threshold (Wright, 1934). Existing quantitative genetics theory developed for continuous traits holds exactly for the continuous liability of binary traits. Methods of QTL mapping for binary traits have been well developed in line-crossing experiments (Hackett & Weller, 1995; Visscher *et al*., 1996; Xu & Atchley, 1996; Rebai, 1997; Xu *et al*., 1998) and four-way crosses (Rao & Xu, 1998). The methods are primarily conducted in a single cross or family. The statistical power of QTL mapping with a single family strongly depends on the two parents selected. If the two parents are fixed for the same allele at a putative QTL, the QTL is undetectable, no matter how many offspring are sampled from the family. But on the other hand, even if a QTL is segregating in the family and is detected, the estimated variance of the QTL can not be extrapolated beyond the particular family. To avoid a loss in statistical power as a result of homozygous parents being selected and to increase the statistical inference space of the estimated QTL parameters, one needs to combine data from multiple families. Methods to handle normally distributed data from multiple families in QTL mapping have already been developed. Typically, these methods include maximum likelihood (Knott & Haley, 1992; Grignola *et al*., 1996a,b), simple linear regression (Knott *et al*., 1996) and weighted least squares (Xu, 1998a). However, such methods are lacking for QTL mapping of binary traits.

*Correspondence. E-mail: xu@genetics.ucr.edu

© 1999 The Genetical Society of Great Britain.

In this paper, we develop a fixed-model approach to mapping quantitative trait loci for complex binary traits from multiple families of outbred populations. This approach is based on the threshold model, and describes the liability by a single linear model with a heterogeneous residual variance. We treat the genotypic effects of the QTL for each parent as fixed effects. The Fisher-scoring algorithm is adopted here to estimate these genetic parameters. The method automatically generates an asymptotic variance–covariance matrix for the estimated QTL effects, which are eventually used for hypothesis tests and estimation of QTL variances. The method is tested via analyses of simulated data.

Statistical methods

Threshold model

Consider n independent full-sib families. Let y_{ij} ($i = 1, \ldots, n$; $j = 1, \ldots, n_i$) represent an underlying continuous-response variable associated with the jth individual in the ith family. Denote the genotypes of a putative QTL by $Q^S_{i1}Q^S_{i2}$ and $Q^d_{i1}Q^d_{i2}$ for the two parents of the ith family. The four possible genotypes in the progeny are $Q^S_{i1}Q^d_{i1}$, $Q^S_{i1}Q^d_{i2}$, $Q^S_{i2}Q^d_{i1}$ and $Q^S_{i2}Q^d_{i2}$. We denote the values of the four genotypes by G_{i11}, G_{i12}, G_{i21} and G_{i22}, respectively. The underlying variable y_{ij} can be treated as a usual quantitative character, which is described by the following linear model:

$$y_{ij} = \mathbf{x}_{ij}^{\mathrm{T}}\boldsymbol{\beta} + z_{ij1}G_{i11} + z_{ij2}G_{i12} + z_{ij3}G_{i21} + z_{ij4}G_{i22} + \varepsilon_{ij}, \quad (1)$$

where $\boldsymbol{\beta}$ is a vector of unknown parameters (including the overall mean, common environmental effects shared by family members, polygenic effects and so on), which relates y_{ij} via a known incidence vector $\mathbf{x}_{ij}$, ε_{ij} is the residual error distributed as $N(0, \sigma^2_\varepsilon)$, and $\mathbf{z}_{ij} = (z_{ij1}\ z_{ij2}\ z_{ij3}\ z_{ij4})^{\mathrm{T}}$ are indicators of the four possible genotypes. The variables $z_{ijk} (k = 1, \ldots, 4)$ are defined as follows:

$$z_{ijk} = \begin{cases} 1 & \text{if the } k\text{th genotype is observed} \\ 0 & \text{otherwise.} \end{cases}$$

In quantitative genetic analysis of complex binary traits, y_{ij} itself is not observable. It has been postulated that y_{ij} controls the binary expression of the trait through a threshold model (Wright, 1934). The relationship between the underlying variable y_{ij} and the binary response s_{ij} is assumed to be:

$$s_{ij} = \begin{cases} 1 & \text{if } y_{ij} > t \\ 0 & \text{if } y_{ij} \le t, \end{cases} \quad (2)$$

for some threshold value t. The threshold model is overparameterized so that some constraints must be superimposed. As usual, we set $\sigma^2_\varepsilon = 1$ and $t = 0$ (Harville & Mee, 1984; Sorensen *et al*., 1995).

As a regular quantitative trait, the genotypic values of the liability can be partitioned into additive and dominance effects, i.e.

$$G_{ikl} = \alpha^s_{ik} + \alpha^d_{il} + \delta_{ikl} \quad (k = 1, 2; l = 1, 2), \quad (3)$$

where α^s_{i1} and $\alpha^s_{i2}$are the effects of the two alleles in the sire, α^d_{i1} and $\alpha^d_{i2}$are the effects of the two alleles in the dam, and δ_{ikl} is the dominance deviation. Unfortunately, α^s_{ik}, α^d_{il} and δ_{ikl} are not estimable; some constraints are required. If all the allelic and dominance effects are appropriately scaled (standardized), the following restrictions can be applied

$$\sum_{k=1}^{2}\alpha^s_{ik} = \sum_{l=1}^{2}\alpha^d_{il} = \sum_{k=1}^{2}\delta_{ikl} = \sum_{l=1}^{2}\delta_{ikl} = 0 \ \ \forall i, k, l. \quad (4)$$

Under these constraints, there are only three independent estimable effects, which are α^s_{i1}, α^d_{i1} and δ_{i11}. Denote $\alpha^s_i = 2\alpha^s_{i1}, \alpha^d_i = 2\alpha^d_{i1}, \delta_i = \delta_{i11}$, $\mathbf{G}_i = (G_{i11}\ G_{i12}\ G_{i21}\ G_{i22})^{\mathrm{T}}$ and $\boldsymbol{\gamma}_i = (\alpha^s_i\ \alpha^d_i\ \delta_i)^{\mathrm{T}}$, then $\mathbf{G}_i = \mathbf{H}\boldsymbol{\gamma}_i$, where

$$\mathbf{H} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 1 \\ \frac{1}{2} & \frac{1}{2} & 1 \\ \frac{1}{2} & \frac{1}{2} & 1 \\ \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}.$$

Hence, model (1) can be re-expressed as:

$$y_{ij} = \mathbf{x}_{ij}^{\mathrm{T}}\boldsymbol{\beta} + \mathbf{w}_{ij}^{\mathrm{T}}\boldsymbol{\gamma}_i + \varepsilon_{ij}, \quad (5)$$

where $\mathbf{w}_{ij}^{\mathrm{T}} = \mathbf{z}_{ij}^{\mathrm{T}}\mathbf{H}$.

If the QTL is not at a marker locus, its genotype is unobservable so that $\mathbf{z}_{ij}$ are missing. However, the distribution of $\mathbf{z}_{ij}$ can be inferred from the genotypes of linked markers. Define the probability of $z_{ijk} = 1$ conditional on marker information by

$$p_{ijk} = \Pr(z_{ijk} = 1 | I_M) \quad \text{for } k = 1, 2, 3, 4.$$

These conditional probabilities can be calculated using a multipoint method. Details of the general multipoint

method can be found in Kruglyak & Lander (1995) and Rao & Xu (1998).

Given these conditional probabilities, the expectation and covariance matrices of $\mathbf{z}_{ij}$ are

$$\mathrm{E}(\mathbf{z}_{ij}|I_M) = (p_{ij1} \quad p_{ij2} \quad p_{ij3} \quad p_{ij4})^{\mathrm{T}}$$

and

$$\mathrm{Var}(\mathbf{z}_{ij}|I_M) = \begin{pmatrix} p_{ij1}(1 - p_{ij1}) & p_{ij1}p_{ij2} & p_{ij1}p_{ij3} & p_{ij1}p_{ij4} \\ p_{ij1}p_{ij2} & p_{ij2}(1 - p_{ij2}) & p_{ij2}p_{ij3} & p_{ij2}p_{ij4} \\ p_{ij1}p_{ij3} & p_{ij2}p_{ij3} & p_{ij3}(1 - p_{ij3}) & p_{ij3}p_{ij4} \\ p_{ij1}p_{ij4} & p_{ij2}p_{ij4} & p_{ij3}p_{ij4} & p_{ij4}(1 - p_{ij4}) \end{pmatrix}.$$

Therefore, the conditional expectation and variance of y_{ij} can be derived as follows:

$$\begin{aligned} \mathrm{E}(y_{ij}|I_M, \boldsymbol{\beta}, \boldsymbol{\gamma}_i) &= \mathbf{x}_{ij}^{\mathrm{T}}\boldsymbol{\beta} + \mathrm{E}(\mathbf{w}_{ij}|I_M)^{\mathrm{T}}\boldsymbol{\gamma}_i \\ &= \mathbf{x}_{ij}^{\mathrm{T}}\boldsymbol{\beta} + \mathrm{E}(\mathbf{z}_{ij}|I_M)^{\mathrm{T}}\mathbf{H}\boldsymbol{\gamma}_i = \mu_{ij} \end{aligned}$$

and

$$\begin{aligned} \mathrm{Var}(y_{ij}|I_M, \boldsymbol{\beta}, \boldsymbol{\gamma}_i) &= \boldsymbol{\gamma}_i^{\mathrm{T}}\mathrm{Var}(\mathbf{w}_{ij}|I_M)\boldsymbol{\gamma}_i + \sigma_\varepsilon^2 \\ &= \boldsymbol{\gamma}_i^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathrm{Var}(\mathbf{z}_{ij}|I_M)\mathbf{H}\boldsymbol{\gamma}_i + 1 = V_{ij}. \end{aligned}$$

It should be noted that the conditional distribution of y_{ij} is a mixture of four normal distributions. Nevertheless, when the QTL effects are small relative to σ_ε and the marker information content is high, the conditional distribution of y_{ij} can be close to a normal distribution. Therefore, model (5) can be approximated by the following heterogeneous residual variance model

$$y_{ij} = \mathbf{x}_{ij}^{\mathrm{T}}\boldsymbol{\beta} + \mathrm{E}(\mathbf{w}_{ij}|I_M)^{\mathrm{T}}\boldsymbol{\gamma}_i + e_{ij}, \tag{6}$$

where $e_{ij} \sim \mathrm{N}(0,\ V_{ij})$.

Estimating genetic parameters

Under model (6), i.e. $y_{ij}|(I_M,\ \boldsymbol{\beta},\ \boldsymbol{\gamma}_i) \sim \mathrm{N}(\mu_{ij},\ V_{ij})$, the probability of s_{ij} is

$$P_{ij} = \Pr(s_{ij} = 1) = \Pr(y_{ij} > 0|I_M, \boldsymbol{\beta}, \boldsymbol{\gamma}_i) = \Phi\left(\frac{\mu_{ij}}{\sqrt{V_{ij}}}\right),$$

which leads to $\Pr(s_{ij} = 0) = 1 - \Pr(s_{ij} = 1) = 1 - P_{ij}$, where $\Phi(\cdot)$ is the standardized normal distribution function.

In the fixed model, conditional on the genetic effects, $\{s_{ij}\}$ $(i = 1, \ldots, n; j = 1, \ldots, n_i)$ are mutually independent. Therefore, we have the following log-likelihood function

$$L = \sum_{i=1}^{n}\sum_{j=1}^{n_i}\left[s_{ij}\log P_{ij} + (1 \quad s_{ij})\ \log(1 \quad P_{ij})\right]. \tag{7}$$

The maximum likelihood estimates of $\boldsymbol{\theta} = (\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\gamma}_1^{\mathrm{T}} \ldots \boldsymbol{\gamma}_n^{\mathrm{T}})^{\mathrm{T}}$ can be solved using any convenient algorithm. We found that the Fisher-scoring algorithm is easy to derive and also extremely fast; the algorithm is described as follows:

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} + \mathbf{F}^{\ 1}\left(\boldsymbol{\theta}^{(k)}\right)\mathbf{S}\left(\boldsymbol{\theta}^{(k)}\right), \tag{8}$$

where k denotes an iteration index, $\mathbf{S}(\boldsymbol{\theta}) = \frac{\partial L}{\partial \boldsymbol{\theta}} = (\mathbf{S}_{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{S}_1^{\mathrm{T}} \ldots \mathbf{S}_n^{\mathrm{T}})^{\mathrm{T}} = (\frac{\partial L}{\partial \boldsymbol{\beta}^{\mathrm{T}}}\frac{\partial L}{\partial \boldsymbol{\gamma}_1^{\mathrm{T}}} \ldots \frac{\partial L}{\partial \boldsymbol{\gamma}_n^{\mathrm{T}}})^{\mathrm{T}}$ is the score function, and

$$\mathbf{F}(\boldsymbol{\theta}) = \mathrm{E}\left[\mathbf{S}(\boldsymbol{\theta})\mathbf{S}(\boldsymbol{\theta})^{\mathrm{T}}\right] = \begin{pmatrix} \mathrm{E}\left(\frac{\partial L}{\partial \boldsymbol{\beta}}\frac{\partial L}{\partial \boldsymbol{\beta}^{\mathrm{T}}}\right) & \mathrm{E}\left(\frac{\partial L}{\partial \boldsymbol{\beta}}\frac{\partial L}{\partial \boldsymbol{\gamma}_i^{\mathrm{T}}}\right) \\ \mathrm{E}\left(\frac{\partial L}{\partial \boldsymbol{\gamma}_i}\frac{\partial L}{\partial \boldsymbol{\beta}^{\mathrm{T}}}\right) & \mathrm{E}\left(\frac{\partial L}{\partial \boldsymbol{\gamma}_i}\frac{\partial L}{\partial \boldsymbol{\gamma}_i^{\mathrm{T}}}\right) \end{pmatrix}$$

is the Fisher information matrix. The components of the score vector and the Fisher information matrix are given in the Appendix.

Each update of the Fisher-scoring algorithm involves inverting $\mathbf{F}(\boldsymbol{\theta})$, which can be time consuming for a large number of families or a high dimension of $\boldsymbol{\beta}$. These can be avoided by taking advantage of the partitioned structure of $\mathbf{F}(\boldsymbol{\theta})$. Since the lower-right part of $\mathbf{F}(\boldsymbol{\theta})$ is block diagonal, the Fisher-scoring algorithm can be simplified (see Appendix). After obtaining the maximum likelihood estimates of the parameters $\boldsymbol{\beta}$, α_i^s, α_i^d and δ_i, the inverse of $\mathbf{F}(\hat{\boldsymbol{\theta}})$ has to be calculated. However, the inverse of $\mathbf{F}(\boldsymbol{\theta})$ also has a simple form (see Appendix).

Given $\hat{\alpha}_i^s, \hat{\alpha}_i^d$ and $\hat{\delta}_i$, one can estimate the cross-family variances. If parents of the sampled families are randomly sampled from the base population, the cross-family variances provide approximate estimates of the QTL segregation variances in the base population. The additive and dominance variances in the base population are $\sigma_\alpha^2 = \frac{1}{4}\left[\mathrm{Var}\ \alpha_i^s) + \mathrm{Var}\ \alpha_i^d)\right]$ and $\sigma_\delta^2 = \mathrm{Var}(\delta_i)$, respectively.

The additive variance σ_α^2 and dominance variance σ_δ^2 can be estimated by:

$$\begin{aligned} \hat{\sigma}_\alpha^2 = \frac{1}{4n}\sum_{i=1}^{n}\Bigg\{ &\left[\ \hat{\alpha}_i^s)^2 \quad \mathrm{Var}\ \hat{\alpha}_i^s)\right] \\ &+ \left[\ \hat{\alpha}_i^d)^2 \quad \mathrm{Var}\ \hat{\alpha}_i^d)\right]\Bigg\}, \end{aligned} \tag{9}$$

and

$$\hat{\sigma}_\delta^2 = \frac{1}{n}\sum_{i=1}^{n}\left[\hat{\delta}_i^2 \quad \mathrm{Var}\left(\hat{\delta}_i\right)\right]. \tag{10}$$

If $\mathrm{Var}(\hat{\alpha}_i^s)$, $\mathrm{Var}(\hat{\alpha}_i^d)$ and $\mathrm{Var}(\hat{\delta}_i)$ are obtained exactly, formulas (9) and (10) provide unbiased estimates of σ_α^2 and σ_δ^2, respectively. However, $\mathrm{Var}(\hat{\alpha}_i^s)$, $\mathrm{Var}(\hat{\alpha}_i^d)$ and $\mathrm{Var}(\hat{\delta}_i)$ can only be estimated approximately by the inverse of the Fisher information matrix (see the next section), and thus the estimates are only asymptotically unbiased.

Hypothesis testing

A useful property of the Fisher-scoring algorithm is that the variance–covariance matrix of $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^{\mathrm{T}}\, \hat{\boldsymbol{\gamma}}_1^{\mathrm{T}} \ldots \hat{\boldsymbol{\gamma}}_n^{\mathrm{T}})^{\mathrm{T}}$ can be approximated by the inverse of the Fisher information matrix, i.e. $\mathrm{Var}(\hat{\boldsymbol{\theta}}) \approx \mathbf{F}(\hat{\boldsymbol{\theta}})^{-1}$. Because the resulting estimates $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^{\mathrm{T}}\, \hat{\boldsymbol{\gamma}}_1^{\mathrm{T}} \ldots \hat{\boldsymbol{\gamma}}_n^{\mathrm{T}})^{\mathrm{T}}$ are maximum likelihood estimates, they follow a multivariate normal distribution if the family sizes n_i are large enough, i.e.

$$\hat{\boldsymbol{\theta}} \dot{\sim} \mathrm{N}\left(\boldsymbol{\theta}, \mathbf{F}\left(\hat{\boldsymbol{\theta}}\right)^{-1}\right). \tag{11}$$

As a consequence, the following test statistic, w, will follow an approximate chi-squared distribution with m degrees of freedom under the null hypothesis that $\mathbf{C}\boldsymbol{\theta} = 0$:

$$\begin{aligned} w &= \hat{\boldsymbol{\theta}}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\left[\mathbf{C}\mathrm{Var}(\hat{\boldsymbol{\theta}})\mathbf{C}^{\mathrm{T}}\right]^{-1}\mathbf{C}\hat{\boldsymbol{\theta}} \\ &= \hat{\boldsymbol{\theta}}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\left[\mathbf{C}\mathbf{F}(\hat{\boldsymbol{\theta}})^{-1}\mathbf{C}^{\mathrm{T}}\right]^{-1}\mathbf{C}\hat{\boldsymbol{\theta}}, \end{aligned} \tag{12}$$

where m is the rank of matrix $\mathbf{C}$.

The exact form of matrix $\mathbf{C}$ determines the type of hypothesis test. To test the overall null hypothesis $\mathrm{H}_0 : \alpha_i^s = \alpha_i^d = \delta_i = 0 \ \ \forall i$, then $\mathbf{C} = (\mathbf{0} \quad \mathbf{I}_{3n})$. Theoretically, there are many other hypotheses to test. In this study, we carry out only the overall test that no QTL at the locus of interest is segregating.

Simulation studies

Design of simulations

Properties of the proposed method were investigated numerically via Monte Carlo simulations. The following properties were examined: the bias, the standard error of the parameter estimates, and the statistical power of QTL detection. We considered the following factors on the performance of the mapping procedure: (i) the variance explained by the QTL; (ii) the sampling strategy (number of families vs. family size); and (iii) the trait incidence (proportion of affected individuals). We simulated one chromosome of length 100 cM with 11 codominant markers evenly spaced along the chromosome. A single QTL was simulated at 25 cM. The simulation was repeated 100 times for each situation. The total number of individuals was set at 750 in all simulations. The standard error of the parameter estimates was calculated from the standard deviation of the estimates among 100 replicates. The statistical power was determined by counting the number of runs (over the 100 replicates) that have test statistics greater than an empirical critical value. The empirical critical value under each condition is obtained by choosing the 95th and 99th percentile of the highest test statistic over 1000 additional runs under the null model (no QTL segregating).

Five equally frequent alleles were simulated for each marker locus. This setting allows each parent to have a 20% chance of being homozygous at each marker locus. In all situations, the residual error was assumed to be normally distributed, with a variance set at $\sigma_\varepsilon^2 = 1.0$. The broad-sense heritability of QTL, $h_q^2 (= \sigma_g^2 / (\sigma_g^2 + \sigma_\varepsilon^2) = (\sigma_\alpha^2 + \sigma_\delta^2)/(\sigma_\alpha^2 + \sigma_\delta^2 + \sigma_\varepsilon^2))$, was examined at four levels: 0.4, 0.3, 0.2 and 0.1. Only a mixed mode of QTL inheritance was considered, i.e. $\sigma_\alpha^2 = \sigma_\delta^2$. Therefore, $h_q^2 = 0.1$ corresponds to $\sigma_\alpha^2 = \sigma_\delta^2 = 0.06$, $h_q^2 = 0.2$ corresponds to $\sigma_\alpha^2 = \sigma_\delta^2 = 0.125$, $h_q^2 = 0.3$ corresponds to $\sigma_\alpha^2 = \sigma_\delta^2 = 0.215$, and $h_q^2 = 0.4$ corresponds to $\sigma_\alpha^2 = \sigma_\delta^2 = 0.33$. The sampling strategy was simulated at three levels: 5×150, 10×75 and 15×50 (number of families $\times$ family size). The trait incidence was set at 50% and 20%.

QTL allelic effects were considered to be normally distributed with preassigned additive and dominance variances. Each parent of a family was made of two alleles: the first allele was assigned a value sampled from a standard normal distribution, and the second assigned the negative value of the first allele. The dominance effect was the interaction effect between any two sampled alleles, which was assigned a value sampled from a N(0,1) distribution, independent of the allelic effects. When offspring were generated, their genetic values at the QTL were re-scaled so that they had the appropriate assigned variances. The liability of each offspring was the sum of its genetic value, the overall mean μ, and a residual error sampled from the N(0,1) distribution. The observable binary phenotype was set to 1 if the corresponding liability exceeded 0, and 0 otherwise. The overall mean of the liability, contained in $\boldsymbol{\beta}$, determines the proportion of the trait incidence. What we did was to select the appropriate mean so that a preassigned level of incidence was obtained.

Results of simulation

The empirical critical values at Type I error rates of 0.05 and 0.01 for situations with disease incidences of 50% and 20% are given in Table 1. The trait incidence has a

small effect on the empirical critical values. The empirical critical values are slightly higher than the critical values of the χ^2 distribution with corresponding degrees of freedom ($3 \times n$). As the number of families increases, the empirical critical value increases dramatically. This is expected because increasing the number of families increases the number of parameters tested.

Table 1 Empirical critical values for the significance test at $\alpha = 0.05$ and $\alpha = 0.01$, where α is the type I error rate

Trait incidence	Sampling strategy	Empirical critical value	
		$\alpha = 0.05$	$\alpha = 0.01$
50%	5 × 150†	31.2454	36.5079
	10 × 75	49.1508	54.9654
	15 × 50	67.9501	74.6404
20%	5 × 150	30.0567	34.8417
	10 × 75	47.2046	51.2633
	15 × 50	61.6487	66.3394

†Number of families × number of individuals per family.

The estimates of QTL parameters (QTL position, additive and dominance variances, and heritability) and the empirical power are summarized in Tables 2, 3 and 4. The proposed method successfully locates the QTL position and estimates the additive and dominance variance components as well as the heritability with negligible biases. However, the QTL heritability, the sampling strategy and the trait incidence have strong impacts on the performance of the mapping procedure.

The performance of the proposed method is strongly affected by the trait incidence. In all cases, the estimates of the QTL parameters are more accurate and the statistical power is higher with the trait incidence of 50% than those with the trait incidence of 20% (see Tables 2 and 3). Under the threshold model, some information will be lost because of the translation from the underlying liability into the observed binary phenotype. The trait incidence determines the amount of lost information. The closer the trait incidence is to 50%, the less information is lost. This explains why the proposed method performs better in terms of the accuracy of parametric estimation and statistical power when the trait incidence is 50%.

Table 2 Estimates of QTL parameters and empirical power ($\alpha = 0.05, 0.01$) under different levels of heritability of the QTL and different sampling strategies when the trait incidence is 50%. Standard errors of the estimates, given in parentheses, are calculated by the standard deviations among 100 replicated simulations. σ_α^2, σ_δ^2, σ_g^2 and h_q^2 are additive, dominance, genetic variances and heritability of the QTL, respectively

Heritability	Sampling strategy	cM_A	σ_α^2	σ_δ^2	σ_g^2	h_q^2	Power (%)	
							$\alpha = 0.05$	$\alpha = 0.01$
0.10	Parametric value	25	0.06	0.06	0.12			
	5 × 150†	24.91	0.0612	0.0719	0.1331	0.1149	92	87
		(5.9561)	(0.0377)	(0.0537)	(0.0618)	(0.0469)		
	10 × 75	25.50	0.0676	0.0606	0.1282	0.1112	90	84
		(8.7270)	(0.0385)	(0.0416)	(0.0590)	(0.0459)		
	15 × 50	26.03	0.0759	0.0657	0.1416	0.1217	83	65
		(8.8831)	(0.0420)	(0.0446)	(0.0611)	(0.0454)		
0.20	Parametric value	25	0.125	0.125	0.25			
	5 × 150	24.83	0.1275	0.1178	0.2452	0.1918	100	98
		(4.2665)	(0.0551)	(0.0774)	(0.0635)	(0.0635)		
	10 × 75	24.98	0.1362	0.1211	0.2572	0.2003	100	98
		(4.4563)	(0.0684)	(0.0672)	(0.0937)	(0.0584)		
	15 × 50	24.64	0.1409	0.1169	0.2579	0.2022	100	98
		(4.9927)	(0.0534)	(0.0502)	(0.0765)	(0.0478)		
0.30	Parametric value	25	0.215	0.215	0.43			
	5 × 150	24.41	0.1979	0.1727	0.3706	0.2628	100	100
		(4.3649)	(0.0953)	(0.1219)	(0.1437)	(0.0736)		
	10 × 75	24.69	0.2029	0.1870	0.3899	0.2763	100	100
		(5.007)	(0.0971)	(0.0865)	(0.1074)	(0.0593)		
	15 × 50	25.89	0.2138	0.1939	0.4077	0.2848	100	100
		(5.0767)	(0.0767)	(0.0967)	(0.1181)	(0.0581)		

†Number of families × number of individuals per family.

Table 3 Estimates of QTL parameters and empirical power ($\alpha = 0.05, 0.01$) under different levels of heritability of the QTL and different sampling strategies when the trait incidence is 20%. Standard errors of the estimates, given in parentheses, are calculated by the standard deviations among 100 replicated simulations. σ_α^2, σ_δ^2, σ_g^2 and h_q^2 are additive, dominance, genetic variances and heritability of the QTL, respectively

							Power (%)	
Heritability	Sampling strategy	cM_A	σ_α^2	σ_δ^2	σ_g^2	h_q^2	$\alpha = 0.05$	$\alpha = 0.01$
0.10	Parametric value	25	0.06	0.06	0.12			
	5 × 150†	27.03	0.0667	0.0711	0.1379	0.1169	87	72
		(12.792)	(0.0502)	(0.0564)	(0.0832)	(0.0596)		
	10 × 75	27.72	0.0801	0.0645	0.1447	0.1225	70	56
		(15.393)	(0.0508)	(0.0468)	(0.0791)	(0.0576)		
	15 × 50	28.43	0.1063	0.0857	0.1919	0.1556	62	47
		(16.066)	(0.0691)	(0.0613)	(0.1008)	(0.0649)		
0.20	Parametric value	25	0.125	0.125	0.25			
	5 × 150	25.45	0.1419	0.1348	0.2767	0.2059	96	92
		(4.7426)	(0.0880)	(0.1214)	(0.1579)	(0.0894)		
	10 × 75	24.84	0.1443	0.1472	0.2915	0.2196	96	92
		(5.8753)	(0.0729)	(0.0839)	(0.1177)	(0.0684)		
	15 × 50	26.03	0.1898	0.1654	0.3552	0.2481	85	75
		(8.3090)	(0.1299)	(0.1120)	(0.2018)	(0.0975)		
0.30	Parametric value	25	0.215	0.215	0.43			
	5 × 150	24.91	0.1964	0.1912	0.3877	0.2717	100	98
		(4.5662)	(0.1147)	(0.1355)	(0.1453)	(0.0749)		
	10 × 75	26.16	0.2362	0.2069	0.4432	0.2971	99	97
		(4.7667)	(0.1148)	(0.1306)	(0.1744)	(0.0842)		
	15 × 50	25.61	0.2895	0.2145	0.5040	0.3245	99	99
		(8.2998)	(0.1379)	(0.1175)	(0.1957)	(0.0835)		

†Number of families × number of individuals per family.

The proportion of the phenotypic variance explained by the QTL, i.e. the QTL heritability h_q^2, has an effect on the accuracy of the estimated parameters and the statistical power (see Tables 2, 3 and 4). As expected, a lower QTL heritability ($h_q^2 = 0.1$) can increase the standard deviation of the estimated QTL position, and decrease the statistical power. Higher heritability levels tend to be associated with a slightly larger standard deviation in the estimated variances and heritabilities, because of the scaling effect, i.e. the standard deviation is correlated to the mean. In the case of high QTL heritability ($h_q^2 = 0.4$), the genetic variances and heritability are underestimated, although the estimated QTL position is accurate and the statistical power is high

Table 4 Estimates of QTL parameters and empirical power ($\alpha = 0.05, 0.01$) under different sampling strategies when the QTL heritability h_q^2 is 0.4 and the trait incidence is 50%. Standard errors of the estimates, given in parentheses, are calculated by the standard deviations among 100 replicated simulations. Parametric values not listed in the table are: QTL position (cM_A) = 25 cM, additive variance $\sigma_\alpha^2 = 0.33$, dominance variance $\sigma_\delta^2 = 0.33$ and genetic variance $\sigma_g^2 = 0.66$

						Power (100%)	
Sampling strategy	cM_A	σ_α^2	σ_δ^2	σ_g^2	h_q^2	$\alpha = 0.05$	$\alpha = 0.01$
5 × 150†	24.95	0.2797	0.2524	0.5321	0.3385	100	100
	(4.4298)	(0.1381)	(0.1502)	(0.1802)	(0.0762)		
10 × 75	24.39	0.2783	0.2655	0.5438	0.3456	100	100
	(4.7769)	(0.1012)	(0.1367)	(0.1637)	(0.0644)		
15 × 50	24.84	0.2843	0.2484	0.5326	0.3416	100	100
	(5.2051)	(0.1009)	(0.1155)	(0.1497)	(0.0618)		

†Number of families × number of individuals per family.

(Table 4). This is expected because the mixture of normal distributions can not be approximated by a single normal distribution when the QTL heritability is high.

The sampling strategy also has an impact on the estimation of the QTL position and the statistical power (see Tables 2, 3 and 4). When the number of families increases, the standard deviation of the estimated QTL position increases. In the case of low QTL heritability ($h_q^2 = 0.1$), the statistical power decreases as the number of families increases. When the power is already very high (e.g. $h_q^2 > 0.1$), the effect of the sampling strategy is expected to be negligible. However, the sampling strategy has a small effect on the precision of the QTL variances and estimated heritabilities. Overall, QTL mapping performs well with a few large families.

Discussion

We have developed a general framework of QTL mapping for complex binary traits by combining data from multiple families. Instead of analysing each family separately, this method carries out a joint statistical inference for multiple families. The link among families is reflected by the common fixed effect vector $\boldsymbol{\beta}$ (see models 5 and 6), which includes the overall mean μ and some common genetic and nongenetic factors, e.g. common environmental effects shared by these families and maternal effects. Because of these common effects, the estimates and tests of genetic parameters in different families are correlated. Therefore, a joint test for multiple families is more powerful than a test considering each family separately (Rebai & Goffinet, 1993). In addition, there are other reasons to justify the use of the proposed consensus method. Like most human diseases, complex binary traits in animal and plant populations undoubtedly have a complex genetic basis. Some QTLs controlling complex binary traits may be homozygous in any single individual and different individuals may be heterozygous for different QTLs. Ideally, several families should be selected and analysed jointly. As a result of using multiple families, the method has a wider statistical inference space than using a single family. Theoretically, the variance attributable to the QTL is better estimated with a large number of families. However, the number of parameters dramatically increases as the number of families sampled increases, which undoubtedly reduces the statistical efficiency of the proposed method. Therefore, with a fixed number of individuals, there is an optimal allocation between the number of families and the number of individuals per family where QTL mapping reaches its maximum power and minimum estimation error. Limited investigations have shown that a mating with several parents (5–10) should give a good sample of variance and allow the detection of QTL with reasonable power (Muranty, 1996; Xu, 1998a).

A typical problem in QTL mapping comes from missing QTL genotypes. In QTL mapping using outbred line crosses, when the putative QTL is not at a marker, the liability is actually a mixture of four normal distributions. Theoretically, the optimal treatment of the unobservable genotype is the mixture model maximum likelihood method, which uses all information contained in the data (Lander & Botstein, 1989; Zeng, 1994). The heterogeneous residual variance model proposed here uses a single distribution to approximate the four distributions, assuming that the residual is normally distributed. This approximation is feasible only when the QTL effects are small relative to the residual variance. Some comparisons between the heterogeneous residual variance model and the mixture model have been made in QTL mapping, showing that the two methods are virtually identical for normally distributed traits, even when the QTL effects are large (Xu, 1998b). Our simulations show that the proposed method performs well when the QTL heritability is not overly high. In the analysis of real data, the effect of any individual QTL being tested is usually small for most polygenic traits, which makes the proposed method valid for most situations. There are some advantages of the heterogeneous residual variance model over the mixture model. First, a simple Fisher-scoring algorithm is available for the heterogeneous residual variance model, which provides, as a by-product, an estimate of the variance–covariance matrix of the estimated parameters. Therefore, it is straightforward to conduct hypothesis tests. The Fisher-scoring method is difficult to derive for the mixture model. As a result, computing the variance–covariance matrix of the estimated parameters is difficult with the mixture model. Secondly, the heterogeneous residual variance model implemented via the Fisher-scoring algorithm is fast, which allows a multiple sampling technique, e.g. the permutation test, to be used more conveniently.

The results presented in this study are based on known linkage phases. Therefore, implementation of this method requires the knowledge of marker linkage phases in the parents. There are several ways to deduce the linkage phase in outbred pedigrees (e.g. Maliepaard *et al*., 1997).

The proposed method is a fixed-model approach because the genetic effects are treated as fixed. As observed in the simulation studies, the method is efficient when there is a small number of families with large family sizes. However, as the number of families increases, the substantial number of parameters to be estimated often generates some statistical problems, in particular when the family sizes are small. The random

model approach, on the other hand, estimates only a few parameters, because only the variances are estimated and tested. With the random model approach, statistical analysis can be carried out even when the family sizes are small. The random model approach, though, is as-of-yet undeveloped for complex binary traits, and as such deserves further investigation.

Acknowledgements

The authors thank Drs D. Gessler and C. Xie for their helpful comments on the manuscript. This research was supported by the National Institutes of Health Grant GM55321-01 and the USDA National Research Initiative Competitive Grants Program 97-35205-5075.

References

FAHRMEIR, L. AND TUTZ, G. 1994. *Multivariate Statistical Modeling Based on Generalized Linear Models*. Springer-Verlag, New York.

FALCONER, D. S. AND MACKAY, T. F. C. 1996. *Introduction to Quantitative Genetics*, 4th edn. Longman, London.

GRIGNOLA, F. E., HOESCHELE, I. AND TIER, B. 1996a. Mapping quantitative trait loci via residual maximum likelihood: I. Methodology. *Génét. Sél. Évol.*, **28**, 479–490.

GRIGNOLA, F. E., HOESCHELE, I. AND TIER, B. 1996b. Mapping quantitative trait loci via residual maximum likelihood: II. A simulation study. *Génét. Sél. Évol.*, **28**, 491–504.

HACKETT, C. A. AND WELLER, J. I. 1995. Genetic mapping of quantitative trait loci for traits with ordinal distributions. *Biometrics*, **51**, 1252–1263.

HARVILLE, D. A. AND MEE, R. W. 1984. A mixed-model procedure for analyzing ordered categorical data. *Biometrics*, **40**, 393–408.

KNOTT, S. A. AND HALEY, C. S. 1992. Maximum likelihood mapping of quantitative trait loci using full-sib families. *Genetics*, **132**, 1211–1222.

KNOTT, S. A., ELSEN, J. M. AND HALEY, C. S. 1996. Methods for multiple-marker mapping of quantitative trait loci in half-sib populations. *Theor. Appl. Genet.*, **93**, 71–80.

KRUGLYAK, E. S. AND LANDER, E. S. 1995. Complete multipoint sib-pair analysis of qualitative and quantitative traits. *Am. J. Hum. Genet.*, **57**, 439–454.

LANDER, E. S. AND BOTSTEIN, D. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, **121**, 185–199.

MALIEPAARD, C., JANSEN, J. AND VAN OOIJEN, J. W. 1997. Linkage analysis in a full-sib family of an outbreeding plant species: overview and consequence for applications. *Genet. Res.*, **70**, 237–250.

McCULLOCH, C. E. 1994. Maximum likelihood variance components estimation for binary data. *J. Am. Stat. Ass.*, **89**, 330–335.

MURANTY, H. 1996. Power of tests for quantitative trait loci detection using full-sib families in different schemes. *Heredity*, **76**, 156–165.

RAO, S. AND XU, S. 1998. Mapping quantitative trait loci for ordered categorical traits in four-way crosses. *Heredity*, **81**, 214–224.

REBAI, A. 1997. Comparison of methods for regression interval mapping in QTL analysis with non-normal traits. *Genet. Res.*, **69**, 69–74.

REBAI, A. AND GOFFINET, B. 1993. Power of tests for QTL detection using replicated progenies derived from a diallel cross. *Theor. Appl. Genet.*, **86**, 1014–1022.

SORENSEN, D. A., ANDERSON, D., GIANOLA, D. AND KORSGAARD, I. 1995. Bayesian inference in threshold models using Gibbs sampling. *Génét. Sél. Évol.*, **27**, 229–249.

VISSCHER, P. M., HALEY, C. S. AND KNOTT, S. A. 1996. Mapping QTLs for binary traits in backcross and F_2 populations. *Genet. Res.*, **68**, 55–63.

WRIGHT, S. 1934. An analysis of variability in number of digits in an inbred strain of guinea pigs. *Genetics*, **19**, 506–536.

XU, S. 1998a. Mapping quantitative trait loci using multiple families of line crosses. *Genetics*, **148**, 517–524.

XU, S. 1998b. Iteratively reweighted least squares mapping of quantitative trait loci. *Behav. Genet.*, **28**, 341–355.

XU, S. AND ATCHLEY, W. R. 1996. Mapping quantitative trait loci for complex binary diseases using line crosses. *Genetics*, **143**, 1417–1424.

XU, S., YONASH, N., VALLEJO, R. L. AND CHENG, H. H. 1998. Mapping quantitative trait loci for binary traits using a heterogeneous residual variance model: an application to Marek's disease susceptibility in chickens. *Genetica*, **104**, 171–178.

ZENG, Z. B. 1994. Precision mapping of quantitative trait loci. *Genetics*, **136**, 1457–1468.

Appendix

The components of the score vector and the Fisher information matrix

The components of the score vector and the Fisher information matrix are given by

$$\frac{\partial L}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{n}\sum_{j=1}^{n_i} \frac{s_{ij} \quad P_{ij}}{P_{ij}(1 \quad P_{ij})}\frac{\partial P_{ij}}{\partial \boldsymbol{\beta}}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n_i} \frac{s_{ij} \quad P_{ij}\varphi\left(\frac{\mu_{ij}}{\sqrt{V_{ij}}}\right)}{P_{ij} \;\; 1 \quad P_{ij})\sqrt{V_{ij}}}\mathbf{x}_{ij},$$

$$\frac{\partial L}{\partial \boldsymbol{\gamma}_i} = \sum_{j=1}^{n_i} \frac{s_{ij} \quad P_{ij}}{P_{ij}(1 \quad P_{ij})}\frac{\partial P_{ij}}{\partial \boldsymbol{\gamma}_i}$$

$$= \sum_{j=1}^{n_i} \frac{(s_{ij} \quad P_{ij})\varphi\left(\frac{\mu_{ij}}{\sqrt{V_{ij}}}\right)}{P_{ij} \;\; 1 \quad P_{ij})\sqrt{V_{ij}}}\left[\mathrm{E}(\mathbf{w}_{ij}|I_M) \quad \frac{\mu_{ij}\mathrm{Var}(\mathbf{w}_{ij}|I_M)\boldsymbol{\gamma}_i}{V_{ij}}\right],$$

$$\mathrm{E}\left(\frac{\partial L}{\partial \boldsymbol{\beta}}\frac{\partial L}{\partial \boldsymbol{\beta}^{\mathrm{T}}}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n_i} \frac{\left[\varphi\left(\frac{\mu_{ij}}{\sqrt{V_{ij}}}\right)\right]^2}{P_{ij}(1 \quad P_{ij})V_{ij}}\mathbf{x}_{ij}\mathbf{x}_{ij}^{\mathrm{T}},$$

$$\mathrm{E}\left(\frac{\partial L}{\partial \boldsymbol{\beta}}\frac{\partial L}{\partial \boldsymbol{\gamma}_i^{\mathrm{T}}}\right) = \mathrm{E}\left(\frac{\partial L}{\partial \boldsymbol{\gamma}_i}\frac{\partial L}{\partial \boldsymbol{\beta}^{\mathrm{T}}}\right)^{\mathrm{T}}$$
$$= \sum_{j=1}^{n_i} \frac{\left[\varphi\left(\frac{\mu_{ij}}{\sqrt{V_{ij}}}\right)\right]^2}{P_{ij}(1 \quad P_{ij})V_{ij}}\mathbf{x}_{ij}$$
$$\times \left[\mathrm{E}(\mathbf{w}_{ij}|I_M) \quad \frac{\mu_{ij}\mathrm{Var}(\mathbf{w}_{ij}|I_M)\boldsymbol{\gamma}_i}{V_{ij}}\right]^{\mathrm{T}},$$

$$\mathrm{E}\left(\frac{\partial L}{\partial \boldsymbol{\gamma}_i}\frac{\partial L}{\partial \boldsymbol{\gamma}_i^{T}}\right) = \sum_{j=1}^{n_i} \frac{\left[\varphi\left(\frac{\mu_{ij}}{\sqrt{V_{ij}}}\right)\right]^2}{P_{ij}(1 \quad P_{ij})V_{ij}}$$
$$\times \left[\mathrm{E}(\mathbf{w}_{ij}|I_M) \quad \frac{\mu_{ij}\mathrm{Var}(\mathbf{w}_{ij}|I_M)\boldsymbol{\gamma}_i}{V_{ij}}\right]$$
$$\times \left[\mathrm{E}(\mathbf{w}_{ij}|I_M) \quad \frac{\mu_{ij}\mathrm{Var}(\mathbf{w}_{ij}|I_M)\boldsymbol{\gamma}_i}{V_{ij}}\right]^{\mathrm{T}}$$

and

$$\mathrm{E}\left(\frac{\partial L}{\partial \boldsymbol{\gamma}_k}\frac{\partial L}{\partial \boldsymbol{\gamma}_l^{\mathrm{T}}}\right) = 0 \quad \text{if } k \neq l,$$

where $\varphi(\cdot)$ is the probability density of the standardized normal distribution.

A simple algorithm for estimating parameters

Denote $\mathbf{F}_{\boldsymbol{\beta}\boldsymbol{\beta}} = \mathrm{E}\left(\frac{\partial L}{\partial \boldsymbol{\beta}}\frac{\partial L}{\partial \boldsymbol{\beta}^{\mathrm{T}}}\right)$, $\mathbf{F}_{\boldsymbol{\beta}i} = \mathbf{F}_{i\boldsymbol{\beta}}^{\mathrm{T}} = \mathrm{E}\left(\frac{\partial L}{\partial \boldsymbol{\beta}}\frac{\partial L}{\partial \boldsymbol{\gamma}_i^{\mathrm{T}}}\right)$ and $\mathbf{F}_{ii} = \mathrm{E}\left(\frac{\partial L}{\partial \boldsymbol{\gamma}_i}\frac{\partial L}{\partial \boldsymbol{\gamma}_i^{\mathrm{T}}}\right)$. The Fisher information matrix is partitioned into

$$\mathbf{F}(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{F}_{\boldsymbol{\beta}\boldsymbol{\beta}} & \mathbf{F}_{\boldsymbol{\beta}1} & \mathbf{F}_{\boldsymbol{\beta}2} & \cdots & \mathbf{F}_{\boldsymbol{\beta}n} \\ \mathbf{F}_{1\boldsymbol{\beta}} & \mathbf{F}_{11} & 0 & \cdots & 0 \\ \mathbf{F}_{2\boldsymbol{\beta}} & 0 & \mathbf{F}_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{F}_{n\boldsymbol{\beta}} & 0 & 0 & \cdots & \mathbf{F}_{nn} \end{pmatrix}.$$

.

Since the lower-right part of $\mathbf{F}(\boldsymbol{\theta})$ is block diagonal, algorithm (8) can be re-expressed more simply as

$$\mathbf{F}_{\boldsymbol{\beta}\boldsymbol{\beta}}^{(k)}\Delta\boldsymbol{\beta}^{(k)} + \sum_{i=1}^{n}\mathbf{F}_{\boldsymbol{\beta}i}^{(k)}\Delta\boldsymbol{\gamma}_i^{(k)} = \mathbf{S}_{\boldsymbol{\beta}}^{(k)},$$
$$\mathbf{F}_{i\boldsymbol{\beta}}^{(k)}\Delta\boldsymbol{\beta}^{(k)} + \mathbf{F}_{ii}^{(k)}\Delta\boldsymbol{\gamma}_i^{(k)} = \mathbf{S}_i^{(k)}, \quad i = 1, \ldots, n,$$

where $\Delta\boldsymbol{\beta}^{(k)} = \boldsymbol{\beta}^{(k+1)} \quad \boldsymbol{\beta}^{(k)}$ and $\Delta\boldsymbol{\gamma}_i^{(k)} = \boldsymbol{\gamma}_i^{(k+1)} \quad \boldsymbol{\gamma}_i^{(k)}$.

After some transformations, the following algorithm is obtained, where each iteration step implies working off the data twice to obtain first the corrections (Fahrmeir & Tutz, 1994):

$$\Delta\boldsymbol{\beta}_i^{(k)} = \left[\mathbf{F}_{\boldsymbol{\beta}\boldsymbol{\beta}}^{(k)} \quad \sum_{i=1}^{n}\mathbf{F}_{\boldsymbol{\beta}i}^{(k)}\left(\mathbf{F}_{ii}^{(k)}\right)^{\ 1}\mathbf{F}_{i\boldsymbol{\beta}}^{(k)}\right]^{\ 1}$$
$$\times \left[\mathbf{S}_{\boldsymbol{\beta}}^{(k)} \quad \sum_{i=1}^{n}\mathbf{F}_{\boldsymbol{\beta}i}^{(k)}\left(\mathbf{F}_{ii}^{(k)}\right)^{\ 1}\mathbf{S}_{\boldsymbol{\beta}}^{(k)}\right]$$

and then

$$\Delta\boldsymbol{\gamma}_i^{(k)} = \left(\mathbf{F}_{ii}^{(k)}\right)^{\ 1}\left[\mathbf{S}_i^{(k)} \quad \mathbf{F}_{i\boldsymbol{\beta}}^{(k)}\Delta\boldsymbol{\beta}^{(k)}\right], \ i = 1, \ldots, n.$$

A simple method for calculating $\mathbf{F}(\boldsymbol{\theta})^{\ 1}$

Since the lower-right part of $\mathbf{F}(\boldsymbol{\theta})$ is block diagonal, $\mathbf{F}(\boldsymbol{\theta})^{\ 1}$ is obtained using standard formulae for inverting partitioned matrices (Fahrmeir & Tutz, 1994). The result is summarised as follows:

$$\mathbf{F}(\boldsymbol{\theta})^{\ 1} = \begin{pmatrix} \mathbf{V}_{\boldsymbol{\beta}\boldsymbol{\beta}} & \mathbf{V}_{\boldsymbol{\beta}1} & \mathbf{V}_{\boldsymbol{\beta}2} & \cdots & \mathbf{V}_{\boldsymbol{\beta}n} \\ \mathbf{V}_{1\boldsymbol{\beta}} & \mathbf{V}_{11} & \mathbf{V}_{12} & \cdots & \mathbf{V}_{1n} \\ \mathbf{V}_{2\boldsymbol{\beta}} & \mathbf{V}_{21} & \mathbf{v}_{22} & \cdots & \mathbf{V}_{2n} \\ \vdots & & & & \vdots \\ \mathbf{V}_{n\boldsymbol{\beta}} & \mathbf{V}_{n1} & \mathbf{V}_{n2} & \cdots & \mathbf{V}_{nn} \end{pmatrix},$$

where

$$\mathbf{V}_{\boldsymbol{\beta}\boldsymbol{\beta}} = \left(\mathbf{F}_{\boldsymbol{\beta}\boldsymbol{\beta}} \quad \sum_{i=1}^{n}\mathbf{F}_{\boldsymbol{\beta}i}\mathbf{F}_{ii}^{\ 1}\mathbf{F}_{i\boldsymbol{\beta}}\right)^{\ 1},$$

$$\mathbf{V}_{\boldsymbol{\beta}i} = \mathbf{V}_{i\boldsymbol{\beta}}^{\mathrm{T}} = \quad \mathbf{V}_{\boldsymbol{\beta}\boldsymbol{\beta}}\mathbf{F}_{\boldsymbol{\beta}i}\mathbf{F}_{ii}^{\ 1},$$

$$\mathbf{V}_{ii} = \mathbf{F}_{ii}^{\ 1} + \mathbf{F}_{ii}^{\ 1}\mathbf{F}_{i\boldsymbol{\beta}}\mathbf{V}_{\boldsymbol{\beta}\boldsymbol{\beta}}\mathbf{F}_{\boldsymbol{\beta}i}\mathbf{F}_{ii}^{\ 1}$$

and

$$\mathbf{V}_{ij} = \mathbf{V}_{ji}^{\mathrm{T}} = \mathbf{F}_{ii}^{\ 1}\mathbf{F}_{i\boldsymbol{\beta}}\mathbf{V}_{\boldsymbol{\beta}\boldsymbol{\beta}}\mathbf{F}_{\boldsymbol{\beta}j}\mathbf{F}_{ii}^{\ 1}, \ i \neq j.$$