Power of quantitative trait locus mapping for polygenic binary traits using generalized and regression interval mapping in multi-family half-sib designs

HAJA N. KADARMIDEEN[1*], LUC L. G. JANSS[2] AND JACK C. M. DEKKERS[3]
[1] *Department of Animal and Poultry Science, University of Guelph, Guelph, ON N1G 2W1, Canada*
[2] *ID-Lelystad, Institute for Animal Science and Health, PO Box 65, 8200 AB Lelystad, The Netherlands*
[3] *Department of Animal Science, Iowa State University, Ames, IA 50011–3150, USA*

(*Received 26 August 1999 and in revised form 9 June 2000*)

Summary

A generalized interval mapping (GIM) method to map quantitative trait loci (QTL) for binary polygenic traits in a multi-family half-sib design is developed based on threshold theory and implemented using a Newton–Raphson algorithm. Statistical power and bias of QTL mapping for binary traits by GIM is compared with linear regression interval mapping (RIM) using simulation. Data on 20 paternal half-sib families were simulated with two genetic markers that bracketed an additive QTL. Data simulated and analysed were: (1) data on the underlying normally distributed liability (NDL) scale, (2) binary data created by truncating NDL data based on three thresholds yielding data sets with three different incidences, and (3) NDL data with polygenic and QTL effects reduced by a proportion equal to the ratio of the heritabilities on the binary versus NDL scale (reduced-NDL). Binary data were simulated with and without systematic environmental (herd) effects in an unbalanced design. GIM and RIM gave similar power to detect the QTL and similar estimates of QTL location, effects and variances. Presence of fixed effects caused differences in bias between RIM and GIM, where GIM showed smaller bias which was affected less by incidence. The original NDL data had higher power and lower bias in QTL parameter estimates than binary and reduced-NDL data. RIM for reduced-NDL and binary data gave similar power and estimates of QTL parameters, indicating that the impact of the binary nature of data on QTL analysis is equivalent to its impact on heritability.

1. Introduction

Statistical methods to map loci affecting quantitative traits (quantitative trait loci, QTL) in animal and plant populations have been thoroughly addressed for traits whose phenotypes follow a continuous distribution (e.g. Haley & Knott, 1992; Knott *et al.*, 1996). In many cases, however, phenotypes are polychotomized into two or more categories, representing binary (e.g. healthy versus diseased) and categorical traits, respectively. The use of linear models, appropriate for continuous traits, has several limitations for the analysis of categorical or binary traits, including the need to restrict predictions to be within the bounds of probability and the presence of heterogeneous (error) variance on the observed scale (Gianola, 1982).

A theoretically appropriate statistical analysis of categorical traits is based on threshold theory (Dempster & Lerner, 1950; Gianola, 1982). Threshold models, or the nearly equivalent logistic regression models in a generalized linear model (GLM) framework, have been used for identification and mapping of QTL affecting categorical traits (Hackett & Weller, 1995; Xu & Atchley, 1996; Visscher *et al.*, 1996; Rebai, 1997; Rao & Xu, 1998). All these studies considered data from line crosses, and generally found negligible to small benefits of the non-linear model over the use of a linear model for analysis of discrete data, either as less bias in parameter estimates (Hackett

* Corresponding author. Current address: Animal Breeding and Genetics Department, Animal Biology Division, Scottish Agricultural College, West Mains Road, Edinburgh EH9 3JG, Scotland, UK. Tel: +44 (0)131 535 3246. Fax: +44 (0)131 535 3121. e-mail: h.kadarmideen@ed.sac.ac.uk

& Weller, 1995) or increased power (Rao & Xu, 1998). Others, however, reported no differences between the two methods (Visscher *et al*., 1996; Rebai, 1997).

Yi & Xu (1999*a*) developed a threshold model with heterogeneous error variance for mapping QTL for binary traits in multiple full-sib families based on Fisher's scoring algorithm. They reported that their method is powerful and provides accurate estimates of QTL variances but did not provide a comparison of their method to a linear model. Greater differences between regression interval mapping (RIM) and generalized interval mapping (GIM) may be expected for multi-family half-sib designs than has been found for inbred line crosses because mean incidences differ between families and the relationship between effects on the underlying liability and observed scale is non-linear. Additional biases could be introduced in the linear model when data are subject to systematic environmental effects and families are cross-classified with fixed effects in an unbalanced manner. The analysis by GIM is on the underlying normal scale and therefore QTL parameter estimates are expected to be robust in the presence of family effects and fixed effects. The evaluation of differences between RIM and GIM for QTL mapping for binary traits in multi-family half-sib designs with or without the presence of fixed effects formed the first objective of this study.

Earlier studies (Xu & Atchley, 1996; Rebai, 1997; Yi & Xu 1999*a*) reported loss of information for QTL mapping for binary data compared with normally distributed data. Robertson & Lerner (1949) derived the amount of genetic information that is present in phenotypes on the binary versus normal scale, by relating heritabilities on the binary and normal scales. The second objective, therefore, was, to investigate whether this same principle applies to QTL effects on the binary versus normal scales.

2. Materials and methods

(i) *The genetic and experimental model*

Consider S unrelated half-sib families. Sire i is randomly mated to n_i unrelated dams resulting in n_i offspring. Let N be the total number of progeny across all sires. Binary phenotypes and marker data are available on the offspring and marker data on the sires. Dams may or may not have marker genotype information. Marker genotype information from two co-dominant polymorphic genetic marker loci, M and N, with known recombination rate θ, is considered. Markers M and N flank an additive QTL with alleles Q_1 and Q_2 and QTL substitution effect (Falconer & Mackay, 1996) denoted as a_n on the underlying normal scale and a_p on the observed binomial scale. The recombination rate of the QTL is r_1 with M and r_2 with N. Haldane's (1919) mapping function (no interference) is assumed. All sires are assumed to be heterozygous for marker loci, with genotypes denoted by M_1N_1/M_2N_2, but can have any one of the three QTL genotypes at the QTL, i.e. Q_1Q_1, Q_1Q_2 or Q_2Q_2. The QTL allele that is in coupling phase with allele M_1 is denoted by QM_1.

(ii) *Threshold model*

The theory for analysis of categorical traits based on a threshold model was addressed by Dempster & Lerner (1950), Gianola (1982), Gianola & Foulley (1983) and Falconer & Mackay (1996). Let y_{ijk} be a binary response variable observed on offspring k in sire family i and herd j. Threshold theory assumes that y_{ijk} results from an underlying continuous variable z_{ijk}, called liability (Dempster & Lerner, 1950), that is normally distributed with mean μ_{ijk} and variance σ^2 and that liability is a linear combination of genetic and systematic and random environmental effects (Dempster & Lerner, 1950; Gianola & Foulley, 1983). There is a fixed category threshold, T, such that $y_{ijk} = 1$ if $z_{ijk} > T$ and $y_{ijk} = 0$ if $z_{ijk} \leqslant T$. Note that T is a 'population parameter' which is the same across all offspring, families and fixed effects.

With the presence of systematic environmental effects (herd), QTL and polygenic effects, a linear model for liability in a multi-family half-sib design can be specified as:

$$z_{ijk} = \beta_{oi} + h_j + \beta_i c_{ijk} + e_{ijk}, \quad (1)$$

with $i = 1, 2, \ldots, s, j = 1, 2, \ldots, r$ and $k = 1, 2, \ldots, m$; where z_{ijk} is the liability of the kth offspring in the jth herd from the ith sire family, β_{oi} is the mean for the ith sire family, h_j is the fixed effect of herd j, β_i is the QTL substitution effect for the ith sire, c_{ijk} is the conditional probability of transmission of allele Q_{M1} from the ith sire to the kth offspring in the jth herd, and e_{ijk} is a residual, with $e_{ijk} \sim \mathrm{N}(0, \sigma^2)$. Liability z_{ijk} is distributed $\mathrm{N}(\mu_{ijk}, \sigma_e^2)$ with $\mu_{ijk} = \beta_{oi} + h_j + \beta_i c_{ijk}$. The conditional probabilities for QTL allele transmission from sire to offspring (c_{ijk}) were assigned based on the method described in Liu & Dekkers (1998). Because σ_e is unidentifiable in threshold models, σ_e can be set at any arbitrary value (here, $\sigma_e = 1$).

The probability for the ijk^{th} observation being scored as $y_{ijk} = 1$, given the mean, herd effect and QTL coefficient c_{ijk}, is

$$(\pi_{ijk} \mid \mu_{ijk}) = \int_{z_{ijk=T}}^{\infty} \varphi(z_{ijk} \mid \mu_{ijk})\, \partial z_{ijk} = 1 - \Phi(T - \mu_{ijk}) = 1 - \Phi(t) = \Phi(\mu_{ijk}), \quad (2)$$

where $t = (T - \mu_{ijk})$ is the standardized threshold point for a standard normal distribution, $\varphi(.)$ is the normal probability density function and $\Phi(.)$ is the normal cumulative density function. Equation (2) holds when

https://doi.org/10.1017/S001667230000481X Published online by Cambridge University Press

T and the variance of z_{ijk} are arbitrarily set to 0 and 1, respectively.

(iii) *Generalized interval mapping* (*GIM*)

(*a*) *Model and likelihood.* The liability model for binary observations (eq. 1) written in matrix notation is

$$\mathbf{z} = \mathbf{Xb} + \mathbf{e},$$

where $\mathbf{z}_{N\times 1}$ is a vector of liabilities for the observations, $\mathbf{X}_{N\times(2S+r)}$ is a design matrix and $\mathbf{b}_{(2S+r)\times 1}$ is the solution vector and $\mathbf{e}_{N\times 1}$ is a vector of random error terms.

Let $\mathbf{b}' = (\beta_{o1}, \beta_{o2}, \ldots, \beta_{oS}, h_1, h_2, \ldots, h_r, \beta_1, \beta_2, \ldots, \beta_s)'$. The log-likelihood for the binomial distribution of observations from S unrelated sire families, L, can be written, based on principles outlined in Gianola & Foulley (1983), as the sum of the log-likelihood for each offspring over all families and herds:

$$L = \sum_{i=1}^{S}\sum_{j=1}^{r}\sum_{k=1}^{m}[y_{ijk}\ln(\pi_{ijk}) + (1-y_{ijk})\ln(1-\pi_{ijk})]. \quad (3)$$

Here π_{ijk} is the probability that $y_{ijk} = 1$ and 1-π_{ijk} is the probability that $y_{ijk} = 0$ given the parameter vector $\mathbf{b}$ (β_{oi}, h_j, β_i). From equation (2), π_{ijk} is $\Phi(\mu_{ijk})$ with $\mu_{ijk} = E(z_{ijk}\,|\,\mathbf{b}) = \beta_{oi} + h_j + \beta_i c_{ijk}$. Note that the contribution of the ijk^{th} observation to the log-likelihood involves only one of the two terms in equation (3), depending on whether y_{ijk} is equal to 1 or 0. Also, note that the QTL position, r_1, enters into the computation of the probability of $y_{ijk} = 1$ (π_{ijk}; equation 2) and, hence, into the log-likelihood (eq. 3), through the conditional probability of QTL allele transmission c_{ijk}, which is a function of assumed QTL position r_1.

(*b*) *Parameter estimation.* Parameters were estimated as the joint *maximum a posteriori* likelihood (MAP; see Gianola & Foulley, 1983) of β_{oi}'s, h_j's, β_i's and r_1. The maximum is found in a two-step procedure: first, given a value for r_1, the maximum is located with respect to β_{oi}'s, h_j's and β_i's using a Newton–Raphson algorithm (see Appendix) and the corresponding likelihood (equation 3) is computed. Then a one-dimensional grid-search is performed with regard to r_1 (based on varying QTL positions at 1 cM intervals). The position with the highest global likelihood provides the estimate of QTL location. Estimates of β_i and h_j at this position are then taken as the best estimates of QTL substitution and herd effects, respectively.

An estimate of QTL variance (σ^2_Q) was computed as

$$\sigma^2_Q = \sigma^2_\beta - PEV_\beta, \quad (4)$$

where σ^2_β is the variance of 'best' estimates of QTL substitution effects (β_i) across sires and PEV_β is the average prediction error variance of QTL substitution effects (see Appendix) for the best fitting model.

(*c*) *Test of significance for presence of a QTL.* Under the null hypothesis of $\beta_i = 0$ for all i, the log-likelihood of a reduced model (L_{red}) with only family means, β_{oi}, and herd effects, h_j, is maximized with $\pi_{ijk} = \Phi(\beta_{oi} + h_j)$:

$$L_{\text{red}} = \sum_{i=1}^{S}\sum_{j=1}^{r}\sum_{k=1}^{m}[y_{ijk}\ln(\pi_{ijk}) + (1-y_{ijk})\ln(1-\pi_{ijk})].$$

The likelihood ratio (LR) test statistic for testing for presence of a QTL in the marker bracket is:

$$\text{LR} = 2[L_{\text{full}} - L_{\text{red}}],$$

where L_{full} is the likelihood under the alternative hypothesis ($\beta_i \neq 0$ for at least one i), which is given in equation (3).

(iv) *Regression interval mapping* (*RIM*)

The linear regression model fitted to binary observations y_{ijk} with fixed herd effects, was:

$$y_{ijk} = \beta_{oi} + h_j + \beta_i c_{ijk} + e_{ijk}, \quad (5)$$

with all terms as described earlier and implemented using interval mapping as described by Knott *et al.* (1996). An estimate of QTL variance was computed based on equation (4) with PEV_β computed as shown in the Appendix. To test for significance of the presence of a QTL in the marker bracket, the LR test statistic was computed as:

$$\text{LR} = N\ln(RSS_{\text{red}}/RSS_{\text{full}}),$$

where RSS_{red} was obtained from fitting a model with family means and herd effects under the null hypothesis, $\beta_i = 0$ for all i, and RSS_{full} was obtained from fitting the full model (equation 5) under the alternative hypothesis, $\beta_i \neq 0$ for at least one i.

(v) *Simulation*

(*a*) *Genetic data.* Twenty paternal half-sib families were simulated. All sires were heterozygous for markers M and N, which had a spacing of 20 cM. The QTL was 15 and 5 cM from M and N. Sires were either homozygous or heterozygous at the QTL based on allele frequencies equal to 0·5. Marker-QTL (MQTL) haplotypes produced by the sires were sampled according to their expected frequencies of transmission. Maternal marker haplotypes were sampled based on population frequencies (0·5 for all alleles).

(*b*) *Phenotypic data: Binary data with fixed effects.* Phenotypic values of offspring were first generated on

https://doi.org/10.1017/S001667230000481X Published online by Cambridge University Press

the NDL scale according to model (1) with c_{ijk} sampled from a binomial distribution with frequency 0·5. Residuals were sampled from $N[0, \sigma_p^2 - (0{\cdot}25\sigma_u^2 + \sigma_h^2 + 0{\cdot}5\sigma_{QTL}^2)]$, where σ_p^2 is the phenotypic variance, σ_u^2 is the polygenic variance, σ_h^2 is the herd variance and σ_{QTL}^2 is the population QTL variance, assuming equal frequency for QTL alleles. A total heritability, including QTL effects, of 0·25 on the NDL scale was used in all cases. Polygenic sire effects were sampled from $N[0, 0{\cdot}25\sigma_u^2]$. One fixed herd effect with five levels was simulated. Herd effects were sampled from $N[0, \sigma_h^2]$ where σ_h^2 is the herd variance which was set equal to 40 % of the total phenotypic variance on the NDL scale. The distribution of progeny across the five herds was non-random for 5 sires and random for 15 sires. Non-random distribution of progeny was simulated by assigning offspring of a given sire family to one of the 5 herds based on a multinomial distribution with frequencies given in Table 1. The remaining 15 sire families were randomly distributed across 5 herds.

Liability values of offspring (z_{ijk}'s) were standardized by $z^*_{ijk} = (z_{ijk} - m_z)/s_z$ where m_z is the mean and s_z is the standard deviation of the liability data for a given replicate. Standardized liability values were transformed into observable 0–1 binary phenotypes (y_{ijk}) using the same standardized threshold point (t) for all offspring, families and herds. Differences in incidence between sires and between herds were, therefore, generated by applying the same 't' for all N offspring across all sire by herd sub-classes but letting polygenic and herd effects differ across families and herds.

(*c*) *Phenotypic data: NDL, reduced-NDL and binary data without fixed effects.* All data were simulated *without* fixed effects to test for prediction of loss of information from NDL to binary data, and to be able to compare binary data with and without fixed effects. Let subscript *ij* denote the j^{th} progeny in the i^{th} family. First, the underlying NDL data (z_{ij}) without fixed effects were simulated as described previously except that the herd effect was not simulated. Then the same underlying NDL data (z_{ij}) were used to generate two types of data, viz. reduced-NDL and binary data (y_{ij}). The reduced-NDL data were generated by reducing the variance contributed by polygenic and QTL effects by a factor R, which, following Robertson & Lerner (1949), was set equal to

$$R = \frac{\eta^2}{p(1-p)}, \tag{6}$$

where η is the height of the ordinate of the standard normal distribution at threshold t corresponding to population incidence p. This was accomplished by multiplying polygenic (u_i) and QTL (q_{ij}) effects in the liability model, $z_{ij} = u_i + q_{ij} + e_{ij}$, by $\sqrt{R}$ and by sampling environmental effects from $N[0, \sigma_p^2 - (0{\cdot}25\sigma_u^2 . R + 0{\cdot}5\sigma_{QTL}^2 . R)]$.

(*d*) *Parameters for simulations.* Binary datasets were simulated for population incidences p, equal to 0·15, 0·25 and 0·50, corresponding to standardized thresholds t equal to 1·04, 0·67 and 0·00 and for the number of offspring per sire n, equal to 100 and 500. For the reduced-NDL data, the values of R were 0·42, 0·53 and 0·63 for incidences of 0·15, 0·25 and 0·50, respectively. The QTL effect on the underlying NDL scale was equal to 0·30 phenotypic standard deviation units ($\sigma = 1$) for all data and combinations.

(*e*) *Parameter estimation and significance threshold values.* The empirical mean and standard deviation of parameter estimates were obtained by averaging estimates over 1000 replicates. Comparison of RIM and GIM methods for binary data was based on power to detect a QTL and bias and accuracy of QTL parameter estimates. Statistical power was calculated as the proportion of replicates in which the LR test statistic was higher than the significance threshold value. Significance threshold values were determined empirically from data generated under the null hypothesis $\beta_i = 0$ for all i based on 10000 replicates of data for each combination of parameters.

(*f*) *Test of unbiasedness of estimates of QTL effects.* Estimates of QTL substitution effects by sire were regressed on their true values to test for unbiasedness, as described in Liu & Dekkers (1998). A regression coefficient equal to 1 indicates unbiasedness. For binary data with and without fixed effects, the true QTL effect of the i^{th} sire on the probability scale (a_{ip}) was computed based on the sire's true polygenic mean (u_i) and the true QTL effect on the standardized NDL scale (a_n) as

$$a_{ip} = \Phi\left[t_i^* + \frac{a_n}{2}\right] - \Phi\left[t_i^* - \frac{a_n}{2}\right], \tag{7}$$

where $t_i^* = (t - u_i)$.

RIM estimates for binary data with and without fixed effects were also tested for unbiasedness on the normal scale by first transforming estimates of QTL substitution effects for each sire (β_{pi}) to the normal scale (β_{ni}). This transformation was based on the method described for backcross designs by Visscher *et al.* (1996):

$$\beta_{ni} = \Phi^{-1}[1 - \beta_{opi}] - \Phi^{-1}[1 - (\beta_{opi} + \beta_{pi})], \tag{8}$$

where β_{opi} is the estimated mean on the probability scale in family i and $\Phi^{-1}(\Psi)$ is the inverse normal cumulative density function of the argument Ψ. In case of fixed effects, the transformation of QTL

https://doi.org/10.1017/S001667230000481X Published online by Cambridge University Press

Table 1. *Frequencies of offspring per herd used for non-random distribution of five sire families across herds for QTL interval mapping for a binary trait with fixed herd effects*

	Sire family				
Herd	1	4	7	14	20
1	**0·40**	0·10	0·10	0·15	0·15
2	0·20	**0·50**	0·20	0·15	0·10
3	0·10	0·10	**0·35**	0·10	0·10
4	0·10	0·10	0·20	**0·45**	0·10
5	0·20	0·20	0·15	0·15	**0·55**

Diagonals (in bold) indicate which sire family is predominant.

substitution effects on the probability scale (β_{pi}) to the NDL scale (β_{ni}) varied between herds. Hence mean incidence within a family was used to transform estimates of β_{pi} to β_{ni}. The RIM estimates of QTL substitution effects for reduced-NDL data for each sire (β_{ri}) were regressed on their true values on the reduced-NDL scale (based on $a_n \sqrt{R}$).

3. Results

(i) *Comparison of GIM with RIM for binary data with fixed effects*

(*a*) *Significance thresholds*. For each incidence (*p*) by progeny group size (*n*) combination, empirical threshold values are given in Table 2. Significance threshold values were not significantly different for RIM and GIM for any combination. For both methods, significance threshold values tended to be similar across incidences and progeny group sizes. The probabilities of committing a type I error (*P* values) when χ^2_{20} table significance values (taken from Snedecor & Cochran, 1982) are used instead of empirical significance threshold values are also given in parentheses in Table 2. Empirical threshold values were higher than χ^2_{20} table values for all combinations. For a given significance level (1%, 5% or 10%), the *P* values obtained with χ^2_{20} table threshold values were much higher than those for empirical threshold values

Table 2. *Empirical significance threshold values for the LR test for QTL interval mapping for a binary trait in 20 sire families based on linear regression* (*RIM*) *and threshold* (*GIM*) *models for different incidences and progeny group sizes* (*n*)[1]. *Results are based on 10000 replicates of binary data with fixed effects*

		Incidence = 0·15		Incidence = 0·25		Incidence = 0·50		
Significance level	*n*	RIM	GIM	RIM	GIM	RIM	GIM	χ^2_{20}
1%	100	42.4[a] (0·021)	41·9[a] (0·021)	40·5[a] (0·023)	40·5[a] (0·021)	40·9[a] (0·024)	40·8[a] (0·024)	37·57
	500	41·1[a] (0·022)	39·4[a] (0·021)	40·0[a] (0·019)	40·2[a] (0·019)	40·6[a] (0·021)	40·4[a] (0·021)	
5%	100	35·2[a] (0·095)	35·4[a] (0·095)	35·1[a] (0·101)	35·2[a] (0·101)	34·8[a] (0·101)	34·1[a] (0·100)	31·41
	500	34·1[a] (0·093)	33·6[a] (0·091)	33·6[a] (0·085)	33·6[a] (0·085)	33·9[a] (0·087)	33·8[a] (0·084)	
10%	100	31·9[a] (0·171)	32·0[a] (0·171)	31·4[a] (0·173)	31·5[a] (0·174)	31·2[a] (0·174)	31·3[a] (0·174)	28·41
	500	31·0[a] (0·166)	30·4[a] (0·165)	30·7[a] (0·159)	30·7[a] (0·160)	30·8[a] (0·165)	30·8[a] (0·165)	

χ^2_{20} is a chi-square with 20 degrees of freedom. Values in parentheses are expected probabilities of type I error when χ^2_{20} table significance values are used instead of empirical significance threshold values[2].

[1] Comparisons are based on a chi-square test for test of significance of a binomial proportion (Snedecor & Cochran, 1982) using RIM significance values for GIM and counting number of replicates of GIM falling above or below the significance threshold values of the RIM and comparing these with expected numbers at a given significance level.

[2] Computed as the number of replicates (under the null hypothesis) with test statistics greater than the χ^2_{20} table significance values divided by the total number of replicates (10000) at a given significance level.

[a] Values with the same superscript within a combination of progeny group size by incidence parameters at a given significance level are not significantly different ($P > 0{\cdot}005$).

https://doi.org/10.1017/S001667230000481X Published online by Cambridge University Press

Table 3. *Empirical power at 1% and 5% levels of significance for QTL interval mapping for a binary trait with a QTL effect of 0·3*[1] *based on linear regression (RIM) and threshold models (GIM) for different incidences and progeny group sizes (n)*[2]*. Correlations between LR test statistics of two methods are also given. Results are based on 1000 replicates of binary data with fixed effects*

		Incidence = 0·15		Incidence = 0·25		Incidence = 0·50	
n	Significance level	RIM	GIM	RIM	GIM	RIM	GIM
100	1%	15·1[a]	15·3[a]	20·7[a]	20·8[a]	24·5[a]	24·2[a]
	5%	27·4[a]	28·8[a]	35·1[a]	34·9[a]	48·0[a]	48·9[a]
500	1%	87·0[a]	87·3[a]	94·9[a]	94·2[a]	98·2[a]	98·5[a]
	5%	95·6[a]	95·9[a]	98·6[a]	98·4[a]	99·3[a]	99·0[a]
		Correlations between LR test statistics					
100		0·934		0·967		0·996	
500		0·959		0·974		0·998	

[1] QTL effect on the underlying normal scale in phenotypic standard deviations.
[2] Comparisons are based on a chi-square test for comparison of proportions in paired samples (Snedecor & Cochran, 1982)
[a] Values with the same superscript within a combination of progeny group size by incidence parameters at a given significance level are not significantly different; $P > 0{\cdot}005$.

Table 4. *Mean and standard deviation (in parentheses) of estimates of QTL location* (in centimorgans from the left marker locus) for QTL interval mapping for a binary trait with a QTL effect of 0·3*[1] *based on linear regression (RIM) and threshold models (GIM) for different incidences and progeny group sizes (n). Results are based on 1000 replicates of binary data with fixed effects*

	Incidence = 0·15		Incidence = 0·15		Incidence = 0·15	
n	RIM	GIM	RIM	GIM	RIM	GIM
100	10·9[a] (8·7)	10·9[a] (8·7)	12·1[a] (8·1)	12·4[a] (8·0)	12·6[a] (7·9)	12·6[a] (7·9)
500	14·2[a] (5·7)	14·1[a] (5·6)	14·3[a] (5·0)	14·3[a] (5·0)	14·5[a] (4·4)	14·5[a] (4·4)

The true location of the QTL was 15 cM from the left marker locus. Distance between marker loci was 20 cM.
[1] QTL effect on the underlying normal scale in phenotypic standard deviations.
* All estimates significantly different from the true QTL location; $P < 0{\cdot}001$.
[a] Values with the same superscript within a combination of progeny group size by incidence parameters are not significantly different ($P > 0{\cdot}005$) and comparisons of methods are based on pairwise *t*-test.

in all situations. This indicates that the use of table values in this application would lead to too liberal a test.

(*b*) *Power and LR test statistics*. Empirical power to detect the QTL and correlations between LR test statistics under the alternative hypothesis for RIM and GIM are given in Table 3. None of the parameter combinations revealed significant differences in power between RIM and GIM. Power increased with progeny group size and incidence from low to intermediate for both methods. Correlations between the LR test statistics for RIM and GIM were close to unity. RIM and GIM produced not only very similar significance threshold values under the null hypothesis (Table 2) but also similar LR test statistics under the alternative hypothesis. Correlations increased with progeny group size and with an increase in incidence from low to intermediate.

(*c*) *QTL location*. The empirical means and standard deviations of estimates of QTL location are in Table 4. There were no significant differences between estimates of QTL location from RIM and GIM for any situation investigated. RIM and GIM estimates of QTL location were significantly biased towards the

https://doi.org/10.1017/S001667230000481X Published online by Cambridge University Press

Table 5. *Empirical mean estimates of the QTL variance* ($\times 10$), *with empirical standard deviations in parentheses, for QTL interval mapping for a binary trait with a QTL effect of 0·3*[1] *based on linear regression* (*RIM*) *and threshold models* (*GIM*) *for different incidences and progeny group sizes* (*n*). *Results are based on 1000 replicates of binary data with fixed effects*

n	Incidence	RIM	True value[2]	GIM	True value[3]
100	0·15	0·034 (0·036)*	0·025	0·98 (1·02)*	0·45
500		0·031 (0·015)*	0·025	0·71 (0·37)*	0·45
100	0·25	0·061 (0·055)*	0·045	0·94 (0·82)*	0·45
500		0·054 (0·025)*	0·045	0·68 (0·29)*	0·45
100	0·50	0·092 (0·069)*	0·072	0·89 (0·61)*	0·45
500		0·081 (0·040)*	0·072	0·60 (0·24)*	0·45

[1] QTL effect on the underlying normal scale in phenotypic standard deviations.
[2] True amount of phenotypic variance contributed by the QTL on the probability scale for RIM.
[3] True amount of phenotypic variance contributed by the QTL on the NDL scale for GIM.
* Estimate significantly different from the true QTL variance; $P < 0{\cdot}001$.

Table 6. *Regression of estimates on true values for QTL effects of individual sires for QTL interval mapping for a binary trait with a QTL effect of 0·3*[1] *based on linear regression* (*RIM*) *and threshold models* (*GIM*) *for different progeny group sizes* (*n*) *and incidences. Results are based on 1000 replicates of binary data with fixed effects*

n	Incidence	RIM (probability scale)	RIM (NDL scale)	GIM (NDL scale)
100	0·15	0·83*	1·09*	1·06*
500		0·96*	1·34*	1·04*
100	0·25	0·92*	1·36*	1·05*
500		1·39*	1·44*	1·03*
100	0·50	0·96*	1·21*	1·00
500		1·01*	1·48*	1·02*

[1] QTL effect on the underlying normal scale in phenotypic standard deviations.
* Significantly different from 1 at 0·001 significance.

middle of the bracket by 0·5 cM to 4·1 cM. Large progeny group size and higher incidence, up to 50%, generally reduced the bias in estimates of QTL location.

Frequency distributions of estimates of QTL location within a marker bracket were similar for RIM and GIM (results not shown). The QTL was positioned more frequently at the right marker locus than at the left marker locus for all situations. In general, the QTL was positioned at markers more frequently when the power was low.

(*d*) *QTL variance*. Empirical means of estimates of QTL variances and their empirical standard deviations across 1000 replicates are given in Table 5. True QTL variance on the NDL scale was computed based on true QTL effects on the NDL scale as $0{\cdot}5a_n^2$. True QTL variance on the probability scale was computed based on true QTL effects on the probability scale as $0{\cdot}5a_p^2$ with a_p derived from equation (7). In general, both RIM and GIM significantly overestimated the QTL variance. Bias and standard deviations of estimates of QTL variances were lower when family size was large and when incidence approached 50%. Estimates of QTL variances obtained from RIM and GIM were on different scales and could not be compared.

(*e*) *Test of unbiasedness of estimates of QTL effects*. Estimates of QTL effects were significantly biased (Table 6). For RIM, per cent bias ranged from −17% to +39% on the probability scale and from 9% to 48% on the NDL scale. For the GIM method, per

https://doi.org/10.1017/S001667230000481X Published online by Cambridge University Press

Table 7. *Empirical significance threshold values for the LR test for QTL interval mapping for a binary trait in 20 sire families based on linear regression (RIM) and threshold (GIM) models for different incidences and progeny group sizes (n)*[1]. *Results are based on 10000 replicates of binary data without fixed effects*

Significance level	*n*	Incidence = 0·15				Incidence = 0·25			Incidence = 0·50			
		RIM_n[2]	RIM_r[3]	RIM_b[4]	GIM	RIM_r	RIM_b	GIM	RIM_r	RIM_b	GIM	χ^2_{20}
1 %	100	$40{\cdot}6^a$	$40{\cdot}7^a$	$42{\cdot}2^b$	$42{\cdot}0^{bc}$	$40{\cdot}3^{ac}$	$40{\cdot}2^{ab}$	$40{\cdot}5^{bc}$	$40{\cdot}7^{ac}$	$41{\cdot}0^{ab}$	$40{\cdot}9^{bc}$	37·57
		(0·021)	(0·021)	(0·021)	(0·021)	(0·021)	(0·025)	(0·025)	(0·021)	(0·026)	(0·026)	
	500	$40{\cdot}0^a$	$40{\cdot}0^a$	$41{\cdot}7^b$	$37{\cdot}9^c$	$40{\cdot}0^{ac}$	$41{\cdot}5^{ab}$	$39{\cdot}3^{bc}$	$40{\cdot}0^{ac}$	$40{\cdot}3^{ab}$	$40{\cdot}0^{bc}$	
		(0·023)	(0·023)	(0·023)	(0·020)	(0·023)	(0·023)	(0·023)	(0·023)	(0·025)	(0·024)	
5 %	100	$34{\cdot}1^a$	$34{\cdot}4^a$	$35{\cdot}1^b$	$35{\cdot}0^{bc}$	$34{\cdot}4^{ac}$	$34{\cdot}6^{ab}$	$34{\cdot}4^{bc}$	$34{\cdot}4^{ac}$	$34{\cdot}2^{ab}$	$34{\cdot}5^{bc}$	31·41
		(0·085)	(0·085)	(0·088)	(0·088)	(0·085)	(0·096)	(0·092)	(0·085)	(0·102)	(0·097)	
	500	$33{\cdot}6^a$	$33{\cdot}6^a$	$34{\cdot}8^{ab}$	$31{\cdot}7^c$	$33{\cdot}6^{ac}$	$34{\cdot}6^{ab}$	$33{\cdot}5^{bc}$	$33{\cdot}6^{ac}$	$33{\cdot}9^{ab}$	$33{\cdot}7^{bc}$	
		(0·092)	(0·092)	(0·095)	(0·087)	(0·092)	(0·095)	(0·092)	(0·092)	(0·096)	(0·093)	
10 %	100	$31{\cdot}1^a$	$31{\cdot}1^a$	$31{\cdot}7^{ab}$	$31{\cdot}9^{bc}$	$31{\cdot}3^{ac}$	$31{\cdot}2^{ab}$	$31{\cdot}1^{bc}$	$31{\cdot}1^{ac}$	$31{\cdot}2^{ab}$	$31{\cdot}2^{bc}$	28·41
		(0·159)	(0·159)	(0·162)	(0·162)	(0·159)	(0·171)	(0·169)	(0·159)	(0·179)	(0·179)	
	500	$30{\cdot}7^a$	$30{\cdot}6^a$	$31{\cdot}5^b$	$28{\cdot}6^c$	$30{\cdot}6^{ac}$	$31{\cdot}2^{ab}$	$30{\cdot}5^{bc}$	$30{\cdot}6^{ac}$	$30{\cdot}9^{ab}$	$30{\cdot}6^{bc}$	
		(0·170)	(0·170)	(0·178)	(0·161)	(0·170)	(0·176)	(0·171)	(0·170)	(0·174)	(0·172)	

χ^2_{20} is a chi-square with 20 degrees of freedom. Values in parentheses are expected probabilities of type I error when χ^2_{20} table significance values are used instead of empirical significance threshold values[5].

[a, b, c] Values with the same superscript within a combination of progeny group size by incidence parameters at given a significance level are not significantly different; $P > 0{\cdot}005$.

[1] Comparisons between methods are as explained in Table 2.

[2] RIM_n RIM applied to NDL data.

[3] RIM_r RIM applied to reduced-NDL data.

[4] RIM_b RIM applied to binary data.

[5] Computation of values is explained in Table 2.

Table 8. *Empirical power at 1% and 5% levels of significance for QTL interval mapping for a binary trait with a QTL effect of 0·3*[1] *based on linear regression (RIM) and threshold models (GIM) for different incidences and progeny group sizes (n)*[2]. *Results are based on 1000 replicates of binary data without fixed effects*

n	Significance level	Incidence = 0·15				Incidence = 0·25			Incidence = 0·50		
		RIM_n	RIM_r	RIM_b	GIM	RIM_r	RIM_b	GIM	RIM_r	RIM_b	GIM
100	1 %	$30{\cdot}6^a$	$6{\cdot}1^b$	$6{\cdot}0^b$	$5{\cdot}10^b$	$12{\cdot}3^b$	$11{\cdot}8^b$	$11{\cdot}9^b$	$12{\cdot}3^b$	$12{\cdot}0^b$	$11{\cdot}9^b$
	5 %	$52{\cdot}7^a$	$18{\cdot}4^b$	$18{\cdot}2^b$	$18{\cdot}3^b$	$24{\cdot}5^b$	$25{\cdot}1^b$	$25{\cdot}8^b$	$32{\cdot}0^b$	$31{\cdot}1^b$	$31{\cdot}2^b$
500	1 %	$99{\cdot}6^a$	$72{\cdot}5^b$	$71{\cdot}1^b$	$71{\cdot}3^b$	$85{\cdot}4^b$	$84{\cdot}9^b$	$84{\cdot}8^b$	$90{\cdot}8^b$	$90{\cdot}8^b$	$90{\cdot}9^b$
	5 %	$100{\cdot}0^a$	$86{\cdot}7^b$	$88{\cdot}4^b$	$87{\cdot}7^b$	$93{\cdot}4^b$	$94{\cdot}7^b$	$95{\cdot}0^b$	$96{\cdot}7^b$	$97{\cdot}3^b$	$98{\cdot}0^b$

RIM_n, RIM_r, RIM_b: abbreviations explained in Table 7.

[1] QTL effect on the underlying normal scale in phenotypic standard deviations.

[2] Comparisons between methods are as explained in Table 3.

[a, b] Values with the same superscript within a combination of progeny group size by incidence parameters at a given significance level are not significantly different; $P > 0{\cdot}005$.

cent bias ranged from 2 % to 6 %. GIM, therefore, has smaller and less variable bias, and also is consistent in always showing a (slight) overestimate.

(ii) *Comparison of NDL, reduced-NDL and binary data based on RIM* (*all without fixed effects*)

(*a*) *Significance threshold values*. Empirical threshold values for data without fixed effects are given in Table 7. Significance threshold values were not significantly different between reduced-NDL and binary data, except for $p = 0{\cdot}15$. Significance threshold values did not differ significantly between NDL and reduced-NDL data in any situation. Similar to data with fixed effects, the *P*-values (in parentheses) obtained when χ^2_{20} table threshold values were applied, were much higher than those expected for empirical threshold values in all situations.

(*b*) *Power and LR test statistics*. Empirical power to detect QTL for all data is given in Table 8. Power was not significantly different between binary and reduced-

https://doi.org/10.1017/S001667230000481X Published online by Cambridge University Press

Table 9. *Mean and standard deviation* (*in parentheses*) *of estimates of QTL location** (*in centimorgans from the left marker locus*) *for QTL interval mapping with a QTL effect of 0·3*[1] *based on linear regression* (*RIM*) *and threshold models* (*GIM*) *for different incidences and progeny group sizes* (*n*). *Results are based on 1000 replicates of binary data without fixed effects*

	Incidence = 0·15				Incidence = 0·25			Incidence = 0·50		
n	RIM_n	RIM_r	RIM_b	GIM	RIM_r	RIM_b	GIM	RIM_r	RIM_b	GIM
100	12·8	11·0[b]	11·4[b]	11·0[b]	11·3[b]	11·6[b]	11·7[b]	11·5[b]	11·6[b]	11·8[b]
	(7·5)	(8·7)	(8·6)	(8·6)	(8·6)	(8·7)	(8·5)	(8·4)	(8·4)	(8·3)
500	14·7	13·5[b]	13·5[b]	13·5[b]	13·9[b]	13·9[b]	13·9[b]	14·2[b]	14·1[b]	14·2[b]
	(4·0)	(6·4)	(6·3)	(6·3)	(5·7)	(5·7)	(5·7)	(5·3)	(5·2)	(5·2)

The true location of the QTL was 15 cM from the left marker locus. Distance between marker loci was 20 cM.
RIM_n, RIM_r, RIM_b: abbreviations explained in Table 7.
[1] 1QTL effect on the underlying normal scale in phenotypic standard deviations.
* All estimates were significantly different from the true QTL location; $P < 0{\cdot}001$, except for RIM_n at $n = 100$.
[a,b] Values with the same superscript within a combination of progeny group size by incidence parameters are not significantly different ($P > 0{\cdot}005$) and comparisons of methods are based on a pairwise *t*-test.

Table 10. *Empirical mean estimates of the QTL variance* ($\times 10$), *with empirical standard deviations in parentheses*, *for QTL interval mapping with a QTL effect of 0·3*[1] *based on linear regression* (*RIM*) *and threshold models* (*GIM*) *for different incidences* (*p*) *and progeny group sizes* (*n*). *Results are based on 1000 replicates of binary data without fixed effects*

n	RIM_n	*P*	RIM_r	True value[2]	RIM_b	True value[3]	GIM	True value[4]
100	0·50 (0·36)*	0·15	0·24 (0·27)*	0·19	0·033 (0·039)*	0·025	0·72 (0·84)*	0·45
500	0·46 (0·16)		0·20 (0·10)*	0·19	0·027 (0·013)*	0·025	0·41 (0·21)*	0·45
100		0·25	0·29 (0·28)*	0·24	0·054 (0·057)*	0·045	0·64 (0·60)*	0·45
500			0·25 (0·12)*	0·24	0·048 (0·021)*	0·045	0·49 (0·20)*	0·45
100		0·50	0·33 (0·29)*	0·29	0·084 (0·076)*	0·072	0·63 (0·56)*	0·45
500			0·29 (0·13)	0·29	0·074 (0·031)	0·072	0·49 (0·18)*	0·45

RIM_n, RIM_r, RIM_b: abbreviations explained in Table 7.
[1] QTL effect on the underlying normal scale in phenotypic standard deviations.
[2] True amount of phenotypic variance contributed by the QTL on the reduced-NDL scale for RIM_r.
[3] True amount of phenotypic variance contributed by the QTL on the probability scale for RIM_b.
[4] True amount of phenotypic variance contributed by the QTL on the NDL scale for RIM_n and GIM.
[5] Estimate significantly different from the true QTL variance; $P < 0{\cdot}001$.

NDL data in any situation. However, power was significantly higher for NDL data for all situations. Power increased with progeny group size and incidence from low to intermediate for all data. Mean LR test statistics under the alternative hypothesis for binary and reduced-NDL data were very similar and the correlations between LR test statistics for binary and reduced-NDL data were close to unity in all situations (results not shown). Correlations between LR test statistics for both binary and reduced-NDL data with that of NDL data were low but increased with progeny group size and with an increase in incidence from low to intermediate (results not shown). Binary and reduced-NDL data produced not only very similar significance threshold values under the null hypothesis (Table 7) but also similar LR test statistics under the alternative hypothesis.

(*c*) *QTL location*. The empirical means and standard deviations of estimates of QTL location are given in Table 9. There were no significant differences between estimates of QTL location from binary and reduced-NDL data for any situation investigated. Mean estimates of QTL location for binary and reduced-NDL data were both significantly different (and biased) from estimates obtained from NDL data, in all situations. For NDL data, RIM estimates of QTL location were unbiased, except for when progeny group size was small. For large progeny group size and intermediate incidence, mean estimates of QTL location were close to the true location (15 cM) for all data. The NDL data had a higher frequency of estimates at or near the simulated location (15 cM) than the reduced-NDL and binary data (results not shown) for all combinations.

https://doi.org/10.1017/S001667230000481X Published online by Cambridge University Press

Table 11. *Regression of estimates on true values for QTL effects of individual sires for QTL interval mapping with a QTL effect of 0·3*[1] *based on linear regression (RIM) and threshold models (GIM) for different progeny group sizes (n) and incidences (P)*

n	RIM_n (NDL scale)	*p*	RIM_r (reduced-NDL scale)	RIM_b (probability scale)	RIM_b (NDL scale)	GIM (NDL scale)
100	1·00	0·15	0·97*	0·94*	1·13*	1·04
500	1·00		0·99*	0·95*	1·06*	0·95*
100		0·25	0·98*	0·99	1·07*	1·01
500			0·99*	1·04*	1·05*	1·03*
100		0·50	0·98*	1·06*	1·05*	1·04
500			1·00	1·10*	1·05*	1·05*

RIM_n, RIM_r, RIM_b: abbreviations explained in Table 7.
[1] QTL effect on the underlying normal scale in phenotypic standard deviations.
[2] Significantly different from 1 at 0·001 significance.

(*d*) *QTL variance*. Empirical means of estimates of QTL variances and their empirical standard deviations across 1000 replicates are given in Table 10. True QTL variances on the reduced-NDL, probability and the NDL scale are also in Table 10. True QTL variances for the reduced-NDL scale were computed based on the true QTL effects on the reduced-NDL scale as $0{\cdot}5(a_n\sqrt{R})^2$. In general, estimates of QTL variance obtained for reduced-NDL and binary data were both significantly biased upwards more than estimates for NDL data. Estimates of QTL variances obtained from different data were on different scales and could not be compared.

(*e*) *Test of unbiasedness of estimates of QTL effects*. Results for the test of unbiasedness of estimates of QTL effects for all data are given in Table 11. Estimates of QTL effects were significantly biased for binary and reduced-NDL data. For RIM for binary data, per cent bias ranged from −6% to +10% on the probability scale and from 5% to 13% on the NDL scale. For RIM for reduced-NDL data, estimates were biased downwards by 3%. For RIM for NDL data, estimates were unbiased. Also here, GIM, compared with RIM for binary data, had smaller and less variable bias (from −5% to +5%), confirming the results from Table 6. However, differences between the two approaches observed here are much lesser than those observed in Table 6 and, therefore, must have been largely caused by the inclusion of fixed effects.

4. Discussion

In this study, methods to map a QTL affecting a binary polygenic trait in multi-family half-sib designs based on linear and threshold models were applied to simulated data. Biologically and statistically, a threshold model is more appropriate for traits that are recorded on a binary scale and that have a polygenic basis for their manifestation.

The threshold model of Gianola & Foulley (1983), adopted here for QTL mapping, is closely related to the GLM of McCullagh & Nelder (1989), which was applied by Visscher *et al*. (1996) and Rao & Xu (1998) for QTL mapping. The GIM method used here differs from the above two studies in that parameters were estimated as the joint *maximum a posteriori* (MAP) and the Newton–Raphson (N-R) method was used to find the MAP estimates (see Appendix) at each putative QTL location. In our application with fixed effects only, MAP equals maximum likelihood, but in this Bayesian framework one or more random effects can also be easily handled, as often incorporated in animal breeding models. The algorithms used here (and in Yi & Xu, 1999*a*, *b*) differ from the common GLM (McCullagh & Nelder, 1989) in that the maximization algorithms are based on second-order derivatives, which converge quicker than the first-order derivatives used in GLM, and also provide standard errors for parameter estimates.

(a) *Comparison of RIM and GIM methods for binary data with and without fixed effects*

Earlier studies (eg. Visscher *et al*., 1996; Rebai, 1997) reported similarity of RIM and GIM for populations derived from F1 intercross and backcross designs. A backcross design is similar to a single half-sib family. For multi-family half-sib designs, however, differences between RIM and GIM were expected because of differences in mean incidences between families. Varying incidences among families were expected to cause no bias for GIM because analysis is on the underlying normal scale. In addition, the presence of fixed effects in binary data was expected to introduce additional biases with RIM because QTL mapping is done within family but *not* within fixed effects. An

https://doi.org/10.1017/S001667230000481X Published online by Cambridge University Press

unbalanced design was therefore considered in the simulation to reflect, for example, the practice of using better sires in better herds. However, although unbalancedness was introduced (for 5 of the 20 sires), no such association between QTL status and herd mean was added.

Results showed that there were no significant differences between RIM and GIM for data with or without fixed effects, in terms of significance thresholds, LR test statistics, power and bias of QTL location estimates. For estimates of QTL effects, differences in bias were found especially when fixed effects were present. This suggests that benefits of GIM (and threshold models in general) may be found especially in the more complicated designs with unbalanced and/or confounded effects. In our designs, however, GIM was not able to convert this speculative benefit into larger power. For most aspects, therefore, our results expand the similarity of GIM and RIM reported by Visscher *et al.* (1996), Xu & Atchley (1996) and Rebai (1997) to multi-family half-sib designs and presence of fixed effects.

Recently, Yi & Xu (1999*b*) proposed a random model for QTL mapping for binary traits in multiple full-sib families, in which they found the threshold model to result in greater power to detect QTL than the linear model. Yi & Xu (1999*b*) fitted families as random rather than fixed effects which, along with a number of other differences (e.g. design, genetic model and likelihood, number and nature of genetic parameters, tests of hypothesis, and method of estimation and approximation), may have contributed to the better performance of the threshold model compared with our analysis. The exact reasons for the differences in results between Yi & Xu (1999*b*) and our study can not be determined easily, but it is clear that other situations may be found in which the threshold model could be superior to the linear model, as would be expected from theory.

Mean estimates of QTL location were found biased towards the centre of the marker bracket in this study. The magnitude of bias was related to power to detect QTL (Tables 3, 8); bias was greater when the power was low. Other studies have also found similar bias (Visscher *et al.*, 1996; Knott *et al.*, 1996; Walling *et al.*, 1998). Using simulation, Visscher *et al.* (1996) obtained unbiased estimates of QTL location for both normal and binary data by averaging LR statistics for each position across replicates and choosing the position with the highest average LR statistics. With interval mapping, an estimate of QTL location is obtained from each replicate based on the position with the highest LR for that replicate. When QTL position was evaluated based on this criterion in our simulation, an unbiased estimate (15 cM) was obtained also for all parameter combinations and for all data (results not shown). This implies that the bias towards the centre of the marker bracket is due to the non-linear relationship between QTL position and the LR statistic. In practice, however, multiple replicates are not available and the position with the largest and significant test statistic in a given replicate must be chosen.

(b) *Comparison of NDL, reduced-NDL and binary data* (*all without fixed effects*)

The main interest in considering *reduced-NDL* data was to test whether loss of information in QTL mapping for binary data is predictable by the proportion given by Robertson & Lerner (1949). The results showed that loss of information as a result of truncating NDL data to binary form is equivalent to the extent that heritability is reduced in binary versus normal data. The results for *original NDL* data were also presented to show its contrast with results for *reduced-NDL* data in terms of power and bias of QTL parameters.

The higher power and better accuracy for NDL compared with binary data are expected because of loss of information when underlying continuous data are truncated to binary data. Earlier studies have also shown similar results for crosses between inbred lines (Xu & Atchley, 1996; Rebai, 1997). However, comparison of results for reduced-NDL and binary data showed that there are no significant differences between the two types of data (or RIM_r versus RIM_b) in terms of significance threshold values (Table 7), power to detect QTL (Table 8) and estimates of QTL location (Table 9). Estimates of QTL variances were on different scales and, hence, could not be compared, but the test of unbiasedness (Table 11) showed that both types of data tended to have similar bias in estimates of QTL effect. Although the prediction based on the formula of Robertson & Lerner (1949) has been used to transform heritabilities to alternative scale, none of the studies has shown that it can be used for prediction of loss of power to detect QTL and bias in QTL parameter estimates for binary traits. Hence these results will be useful in designing QTL mapping experiments.

(c) *General remarks*

In this study, we considered a single marker interval flanked by two informative markers. In practice, there would be many marker intervals of different information content. In this situation, the GIM method shown here could be directly extended to include multiple markers of different information content as described for RIM by Knott *et al.* (1996). The results from comparison of GIM and RIM would still be applicable in this situation because the effect of adding more markers of different information content is expected to be the same for both the GIM and RIM.

https://doi.org/10.1017/S001667230000481X Published online by Cambridge University Press

One widely known problem in QTL mapping with outbred populations is the uncertainty of marker and QTL allele transmission that warrants the use of probabilities to indicate the QTL allele transmission status (e.g. Knott *et al*., 1996; Kadarmideen & Dekkers, 1999). With the true model, the independent variable (QTL allele transmission status) would be known with certainty (0 or 1). Uncertainty, however, causes at least two problems. First, in our model on the liability scale, we assumed that residual error is uniform and normally distributed. These assumptions are indeed approximations because the residual error has a mixed distribution and is heterogeneous due to errors in independent variables. Secondly, *PEV*s are unbiased only if independent variables are known without error; otherwise they are expected to be biased downwards. This underestimation in *PEV*s results in overestimation of QTL variances (Eq. 4).

Many discontinuous traits can be classified as a binary trait although there are discontinuous traits of interest that have more than two categories. For such traits, GIM could be extended based on methods described by Gianola & Foulley (1983). RIM is not suitable for analysis of traits with multiple categories, especially when the category probabilities must be known accurately for their use in breeding programmes (Rebai, 1997). It is known that interactions found on the underlying scale are expected to be less than those found on the observed scale (confirmed in real calving ease data from American Simmental cattle by Quaas *et al*., 1988), in which case threshold models may be preferred over the linear model. The comparison between threshold and linear models in more complicated designs is an important area of research that needs further investigation.

A binomial distribution of the trait results in a much greater violation of normality than e.g. Poisson or gamma distributions. Therefore it could be expected that similarities between RIM and GIM in the efficiency of QTL mapping, observed in this study, would also be applicable for traits with other non-normal distributions.

(d) *Conclusions*

We showed that, for most practical purposes, the RIM and GIM are equivalent methods for QTL mapping for binary traits for multi-family half-sib designs and in the presence of fixed effects. The similarity in results for RIM and GIM for binary data may be due to the specific situations considered in this study. Specifically, further analysis of more unbalanced data structures, associations between sire QTL status and herd mean, smaller herd and progeny group sizes, and other QTL mapping designs is required. We have also shown that the impact of the binary nature of data on power and QTL parameters can be predicted on the basis of its impact on heritability.

This research was funded by the Guelph-Wageningen Exchange Agreement and Going-Global program in the Ministry of Foreign Affairs in Canada, Natural Sciences and Engineering Research Council of Canada (NSERC) and the Iowa Agriculture and Home Economics Experiment Station of Iowa State University, USA. The authors thank Prof. Bill Hill for constructive suggestions, in particular for suggesting the comparison of reduced NDL data, and Prof. Johan Van Arendonk for help with the Guelph-Wageningen exchange.

Appendix

Newton–Raphson algorithm

Given recombination rate r_1, covariates c_{ijk} can be computed and used as covariates in a (non-linear) regression model. The resulting non-linear regression model was fitted using a Newton–Raphson algorithm of Gianola & Foulley (1983). The variables used in the iterative maximization were:

$$v_{ijk} = \frac{-\varphi(\mu_{ijk})}{1-\Phi(\mu_{ijk})} \quad \text{if } y_{ijk} = 0$$

$$v_{ijk} = \frac{\varphi(\mu_{ijk})}{\Phi(\mu_{ijk})} \quad \text{if } y_{ijk} = 1$$

$$w_{ijk} = v_{ijk}^2 + \mu_{ijk} \cdot \nu_{ijk},$$

where elements, ν_{ijk} and w_{ijk} are first and second differentials of the log-likelihood, L (equation 3), which can be derived based on principles outlined by Gianola & Foulley (1983). Elements ν_{ijk} and w_{ijk} can be accumulated in a vector $\mathbf{v}$ and a diagonal matrix $\mathbf{W}$, respectively. Further, y_{ijk} is replaced by a 'working' dependent variable λ_{ijk}, defined as: $\lambda_{ijk} = \mu_{ijk} + \nu_{ijk} \cdot w_{ijk}^{-1}$. Then solution for $\mathbf{b}$ in iteration number $k+1$, $(\mathbf{b})_{k+1}$, at a given location r_1 in the marker interval is obtained from:

$$\begin{bmatrix} \mathbf{x}_1'\mathbf{w}\mathbf{x}_1 & \mathbf{x}_1'\mathbf{w}\mathbf{x}_2 & \mathbf{x}_1'\mathbf{w}\mathbf{x}_3 \\ \mathbf{x}_1'\mathbf{w}\mathbf{x}_2 & \mathbf{x}_2'\mathbf{w}\mathbf{x}_2 & \mathbf{x}_2'\mathbf{w}\mathbf{x}_3 \\ \mathbf{x}_1'\mathbf{w}\mathbf{x}_3 & \mathbf{x}_2'\mathbf{w}\mathbf{x}_3 & \mathbf{x}_3'\mathbf{w}\mathbf{x}_3 \end{bmatrix}^{k} \begin{bmatrix} \boldsymbol{\beta}_o \\ \mathbf{h} \\ \boldsymbol{\beta} \end{bmatrix}^{k+1} = \begin{bmatrix} \mathbf{x}_1'\mathbf{w}\boldsymbol{\lambda} \\ \mathbf{x}_2'\mathbf{w}\boldsymbol{\lambda} \\ \mathbf{x}_3'\mathbf{w}\boldsymbol{\lambda} \end{bmatrix}^{k}$$

where the elements of the left- and right-hand sides of the equations are sub-matrices and sub-vectors:

$$\mathbf{x}_1'\mathbf{w}\mathbf{x}_1 = \text{diag}\{w_{i..}\}, \mathbf{x}_1'\mathbf{w}\mathbf{x}_2\{w_{ij.}\}, \mathbf{x}_1'\mathbf{w}\mathbf{x}_3 = \text{diag}\{w_{i...}c_{i...}\},$$
$$\mathbf{x}_2'\mathbf{w}\mathbf{x}_2 = \text{diag}\{w_{.j.}\}, \mathbf{x}_2'\mathbf{w}\mathbf{x}_3 = \{w_{ij.}c_{ij.}\}, \mathbf{x}_3'\mathbf{w}\mathbf{x}_3 = \{w_{i..}\, c_{i..}^2\},$$
$$\mathbf{x}_1'\mathbf{w}\boldsymbol{\lambda} = \{w_{i..}\, \lambda_{i..}\}, \mathbf{x}_2'\mathbf{w}\boldsymbol{\lambda} = \{w_{.j.}\, \lambda_{.j.}\}$$
$$\mathbf{x}_3'\mathbf{w}\boldsymbol{\lambda} = [w_{i..}\, c_{i..}\, \lambda_{i..}\}.$$

Because of a linear dependency in the equations, one of the herd effects was set to zero. Note that for each round of iteration, weights w_{ijk} and working dependent variables λ_{ijk} are computed from solutions from the previous round. Iterations were continued until

$$1/q(\Delta_{[k+1]}'\Delta_{[k+1]}) \leqslant \epsilon,$$

https://doi.org/10.1017/S001667230000481X Published online by Cambridge University Press

where $\Delta_{[k+1]} = |\mathbf{b_k} - \mathbf{b_{k+1}}|$, q is the number of parameters fitted and ϵ is an arbitrary small number (10^{-5}). Then, the log-likelihood L in equation (3) is computed using the converged solutions for a given QTL location.

Prediction Error Variance of QTL substitution effects (PEV_β)

GIM. The average PEV_β across S sire families was computed as:

$$PEV_\beta = \frac{1}{S}\sum_{i=1}^{S} \text{diag}\{(\mathbf{X'WX})^{-1}\}_i.$$

Note that σ_e^2 is not involved in the computation of PEV_β because $\mathbf{X'WX}$ is a Hessian matrix and the inverse of $\mathbf{X'WX}$ is the information matrix.

RIM. For RIM, average PEV_β was computed for the best fitting model as:

$$PEV_\beta = \frac{1}{S}\sum_{i=1}^{S} \text{diag}\{(\mathbf{X'X})^{-1}\}_i \sigma_e^2.$$

References

Dempster, E. R. & Lerner, I. M. (1950). Heritability of threshold characters. *Genetics* **35**, 212–235.

Falconer, D. S. & Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*. 4th edn. Harlow, Essex: Longman

Gianola, D. (1982). Theory and analysis of threshold characters. *Journal of Animal Science* **56**, 1079–1096.

Gianola, D. & Foulley, J. L. (1983). Sire evaluation for ordered categorical data with a threshold model. *Genetics*, *Selection*, *Evolution* **15**, 201–224.

Hackett, C. A. & Weller, J. I. (1995). Genetic mapping of quantitative trait loci for traits with ordinal distribution. *Biometrics* **51**, 1252–1263.

Haldane, J. B. S. (1919). The combination of linkage values, and the calculation of distances between the loci of linked factors. *Journal of Genetics* **8**, 299–309.

Haley, C. S. & Knott, S. A. (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* **69**, 315–324.

Kadarmideen, H. N. & Dekkers, J. C. M. (1999). Regression on markers with uncertain allele transmission for QTL mapping in half-sib designs. *Genetics*, *Selection*, *Evolution* **31**, 437–455.

Knott, S. A., Elsen, J. M. & Haley, C. S. (1996). Methods of multiple-marker mapping of quantitative trait loci in half-sib populations. *Theoretical Applied Genetics* **93**, 71–80.

Liu, Z. & Dekkers, J. C. M. (1998). Least squares interval mapping of quantitative trait loci under the infinitesimal genetic model in outbred populations. *Genetics* **148**, 495–505.

McCullagh, P. & Nelder, J. A. (1989). *Generalised Linear Models*. London: Chapman and Hall.

Quaas, R. L., Zhao, Y. & Pollack, E. J. (1988). Describing interactions in dystocia scores with a threshold model. *Journal of Animal Science* **66**, 396–399.

Rao, S & 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. *Genetical Research* **69**, 69–74.

Robertson, A. & Lerner, I. M. (1949). The heritability of all-or-none traits: liability of poultry. *Genetics* **34**, 395–411.

Snedecor, G. W. & Cochran, W. G. (1982). *Statistical Methods*. 7th edn., Ames, Iowa: Iowa State University Press.

Visscher, P. M., Haley, C. S. & Knott, S. A. (1996). Mapping QTLs for binary traits in backcross and F2 populations. *Genetical Research* **68**, 55–63.

Walling, G. A., Visscher, P. M., & Haley, C. S. (1998). A comparison of bootstrap methods to construct confidence intervals in QTL mapping. *Genetical Research* **71**, 171–180.

Xu, S. & Atchley, W. R. (1996). Mapping quantitative trait loci for complex binary diseases using line crosses. *Genetics* **143**, 1417–1424.

Yi, N. & Xu, S. (1999*a*). Mapping quantitative trait loci for complex binary traits in outbred populations. *Heredity* **82**, 668–676.

Yi, N. & Xu, S. (1999*b*). A random model approach to mapping quantitative trait loci for complex binary traits in outbred populations. *Genetics* **153**, 1029–1040.

https://doi.org/10.1017/S001667230000481X Published online by Cambridge University Press