Original article

Longitudinal random effects models for genetic analysis of binary data with application to mastitis in dairy cattle

Romdhane REKAYA[a*], Daniel GIANOLA[b],
Kent WEIGEL[b], George SHOOK[b]

[a] Department of Animal and Dairy Science,
University of Georgia, Athens, GA 30602, USA
[b] Department of Dairy Science, University of Wisconsin,
Madison, WI 53706, USA

Abstract – A Bayesian analysis of longitudinal mastitis records obtained in the course of lactation was undertaken. Data were 3341 test-day binary records from 329 first lactation Holstein cows scored for mastitis at 14 and 30 days of lactation and every 30 days thereafter. First, the conditional probability of a sequence for a given cow was the product of the probabilities at each test-day. The probability of infection at time t for a cow was a normal integral, with its argument being a function of "fixed" and "random" effects and of time. Models for the latent normal variable included effects of: (1) year-month of test + a five-parameter linear regression function ("fixed", within age-season of calving) + genetic value of the cow + environmental effect peculiar to all records of the same cow + residual. (2) As in (1), but with five parameter random genetic regressions for each cow. (3) A hierarchical structure, where each of three parameters of the regression function for each cow followed a mixed effects linear model. Model 1 posterior mean of heritability was 0.05. Model 2 heritabilities were: 0.27, 0.05, 0.03 and 0.07 at days 14, 60, 120 and 305, respectively. Model 3 heritabilities were 0.57, 0.16, 0.06 and 0.18 at days 14, 60, 120 and 305, respectively. Bayes factors were: 0.011 (Model 1/Model 2), 0.017 (Model 1/Model 3) and 1.535 (Model 2/Model 3). The probability of mastitis for an "average" cow, using Model 2, was: 0.06, 0.05, 0.06 and 0.07 at days 14, 60, 120 and 305, respectively. Relaxing the conditional independence assumption *via* an autoregressive process (Model 2) improved the results slightly.

mastitis / longitudinal / threshold model

* Correspondence and reprints
E-mail: rrekaya@arches.uga.edu

1. INTRODUCTION

Longitudinal binary response data arise frequently in many fields of applications ([2,4]). Generally, longitudinal data consist of repeated observations taken over time in a group of individuals. In animal breeding, repeated observations over time on the same animal arise often. Although continuous repeated measurements over time (*e.g.*, test day milk yield) have received a lot of emphasis in the last decade ([8,12,13]), little attention has been devoted towards longitudinal binary responses. Most research done with binary data has focused on cross-sectional data, where only a single response is screened in each animal ([7,9,10,15]). However, a single response is frequently a summary of performance over a period of time. For example, in mastitis studies, a cow is considered infected in lactation if at least one episode of infection has taken place during lactation, disregarding the number of episodes of infection or the timing of the episodes. Furthermore, environmental conditions fluctuate in the course of lactation. A longitudinal analysis of such data gives flexibility by permitting a better modeling of the environmental conditions affecting each episode of infection, and perhaps by allowing a better representation of the covariance structure within and between animals. Also, a longitudinal analysis of binary response data allows developing novel selection criteria other than a single predicted breeding value for liability to disease.

In this study, an approach to the analysis of longitudinal binary responses taken over time is developed in a Bayesian framework. The procedure was applied to longitudinal mastitis records (absence-presence) taken in the course of lactation using three different models for the probability of infection at any time *t* assuming conditional independence. A fourth model allowed for serial correlation between residuals in the underlying scale, thus relaxing the conditional independence assumption.

2. MATERIAL AND METHODS

2.1. Data

The data recorded between July 1982 and June 1989, were provided by the Ohio Agricultural Research and Development Center at Wooster, Ohio, USA. After edits (inconsistent date of control, cows with less than three records), 142 records were eliminated. The final data set consisted of 3341 test-day binary records from 329 first-lactation Holstein cows scored for clinical mastitis at 14 and 30 days of lactation and every 30 d thereafter. Around 88% of the cows had complete lactation and only 4% had 5 test-day records or less. The incidence rate of mastitis (at least one infection in the lactation) was 17.8%. A general description of the original data set can be found in [14]. A summary description is presented in Table I. About 82% the cows did not have mastitis in their first lactation.

Table I. Distribution of cows by the number of mastitis cases.

Test with mastitis	Number of cows	% of total
0	270	82.1
1	32	9.4
2	11	3.3
3	7	2.1
4	3	0.9
5	0	0.0
6	3	0.9
7	1	0.3
8	1	0.3
9	1	0.3
10	1	0.3

2.2. Methods

2.2.1. Cross-sectional binary response

Before describing the longitudinal setting, we describe a basic latent variable model for a cross-sectional binary response. Assume the observed binary response y_i related to a continuous underlying variable l_i satisfying:

$$y_i = \begin{cases} 1 & \text{if } l_i > T \\ 0 & \text{if } l_i \leqslant T, \end{cases} \quad (1)$$

where $l_i \sim N(\mu_i, \sigma^2)$, and T is a threshold value. This is the basic threshold model of quantitative genetics ([5,9,16]). The probability of observing (1) (success) is:

$$P_i = pr(l_i > T|\mu_i) = 1 - pr(l_i > T|\mu_i)$$
$$= 1 - \Phi\left(\frac{T - \mu_i}{\sigma}\right), \quad (2)$$

where Φ is the cumulative standard normal distribution function. It is clear from (2) that it is not possible to infer μ_i, T and σ^2, separately. Hence, some restrictions are placed on two of the three model parameters. A common choice is to set $T = 0$, and $\sigma^2 = 1$, leading to:

$$P_i = pr(l_i > T|\mu_i) = 1 - \Phi(-\mu_i) = \Phi(\mu_i). \quad (3)$$

Furthermore, μ_i can be linearly related to a set of systematic and random effects as:

$$\mu_i = \mathbf{x}_i\beta + \mathbf{z}_i\mathbf{u},$$

where $\mathbf{x}_i$ and $\mathbf{z}_i$ are known incidence row vectors, and β and $\mathbf{u}$ are unknown location vectors corresponding to systematic and random effects, respectively. Implementation of this model in a Bayesian analysis using data augmentation became feasible after Albert and Chib [2] and Sorensen *et al.* [15]. All pertinent regional posterior distributions can be obtained using Markov Chain Monte Carlo procedures.

2.2.2. Longitudinal binary response

Now let $\mathbf{y}_i = (y_{it_1}, y_{it_2}, \ldots, y_{it_{n_i}})'$ be a $n_i x 1$ vector of responses for animal, $(i = 1, 2, \ldots, q)$ observed at times $t_1, t_2, \ldots, t_{n_i}$ test days. As in the case of cross-sectional analysis, the binary response observed at a time t_j related to a normally distributed underlying variable satisfying (1):

$$l_{ij} \sim N(\mu_{ij}, 1),$$

where μ_{ij} is now some function of time. In this study, three functional forms were used to model μ_{ij}, as follows:

Ali-Schaeffer model: **M1**

Here, the Ali and Schaeffer function [1] was fitted as a fixed linear regression within age-season of calving classes. The conditional expected value of liability for a cow, scored at time j in year month of test m calving in age-season r, was assumed to follow the model:

$$\mu_{ijrm} = YM_m + b_{0r} + b_{1r}z_{ij} + b_{2r}z_{ij}^2 + b_{3r}\ln(z_{ij}^{-1}) + b_{4r}[\ln(z_{ij}^{-1})]^2 + u_i + p_i,$$

where $\mu_{ijrm} =$ conditional mean for cow i at time j, year-month m and calving in age-season r; $YM_m =$ effect of year-month of test m ($m = 1, 2, \ldots, 89$); $\mathbf{b}_r = [b_{0r}, b_{1r}, b_{2r}, b_{3r}, b_{4r}] = 5 \times 1$ vector of regressions of liability by age-season (five classes) of calving r; $z_{ij} = \dfrac{\text{days in milk } j \text{ for cow } i}{305}$; $u_i =$ additive genetic value of cow i; $p_i =$ environmental effect peculiar to all n_i records of cow i.

The prior distributions were: $YM_m \sim U[YM_{\min}, YM_{\max}], (m = 1, 2, \ldots, 89)$; $\mathbf{b} \sim U[\mathbf{b}_{\min}, \mathbf{b}_{\max}]$, where $\mathbf{b}$ is a vector containing $\mathbf{b_r}$ vectors; $\mathbf{u} \sim N(\mathbf{0}, A\sigma_u^2)$; $\mathbf{p} \sim N(\mathbf{0}, \sigma_p^2)$; $\sigma_a^2 \sim U[\sigma_{u_{\min}}^2, \sigma_{u_{\max}}^2]$; $\sigma_p^2 \sim U[\sigma_{p_{\min}}^2, \sigma_{p_{\max}}^2]$.

Above, σ_a^2 and σ_p^2 are the additive genetic and permanent environmental components of variances and A is the additive relationship matrix. Bounds were set to large values to avoid truncation of parameter space. The joint prior had the form:

$$p(YM, \mathbf{b}, \mathbf{u}, \mathbf{p}, \sigma_a^2, \sigma_p^2 | \text{hyper-parameters}) = p(YM)p(\mathbf{b})p(\mathbf{u})p(\sigma_a^2)p(\sigma_p^2).$$

Random regression model using the Ali-Schaeffer function: **M2**

This model is similar to the previous one, but the genetic value was modeled *via* five-random regression parameters, leading to:

$$\mu_{ijrm} = YM_m + (b_{0r} + u_{0i}) + (b_{1r} + u_{1i})z_{ij} + (b_{2r} + u_{2i})z_{ij}^2$$
$$+ (b_{3r} + u_{3i})\ln(z_{ij}^{-1}) + (b_{4r} + u_{4i})[\ln(z_{ij}^{-1})]^2 + p_i.$$

The same prior distribution was used for parameters defined earlier. The prior distribution of the genetic regressions was:

$$\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4 | \mathbf{G}_0 \sim N(\mathbf{0}, \mathbf{A} \otimes \mathbf{G}_0),$$

where $\mathbf{G}_0$ is a 5×5 matrix of genetic (co)variances between the random regression parameters. The prior for $\mathbf{G}_0$ was uniform, but with bounds for each non-redundant element (variances and covariances).

The Wilmink hierarchical model: **M3**

A three-stage hierarchical model was implemented using the Wilmink function to relate the mean of the underlying variable to time:

$$\mu_{ijm} = YM_m + \gamma_{0i} + \gamma_{1i}t_{ij} + \gamma_{2i}\exp(-0.05t_{ij}) + p_i,$$

where YM_m and p_i are as defined before, t_{ij} are days in milk on test-day j for cow i and $\gamma_i = (\gamma_{0i}, \gamma_{1i}, \gamma_{2i})'$ a 3×1 vector of the Wilmink's parameters for animal i. At the second stage of the hierarchy, a mixed linear model was imposed to the parameters of the Wilmink function, as follows:

$$\gamma = \mathbf{X}\beta + \mathbf{Z}\mathbf{u} + \mathbf{e},$$

where γ is a vector containing all parameters for all cows, β includes effects of age-season of calving as parameters, $\mathbf{u}$ is a vector of additive genetic values associated with all the Wilmink function parameters, and $\mathbf{e}$ is a second-stage residual term.

To complete the hierarchy, the following prior distributions were assigned to the model parameters:

$$\beta \sim U[\beta_{\min}, \beta_{\max}],$$
$$\mathbf{u}|\mathbf{K}_0 \sim N(\mathbf{0}, \mathbf{A} \otimes \mathbf{K}_0),$$
$$\mathbf{e}|\sum_0 \sim N\left(\mathbf{0}, \mathbf{I} \otimes \sum_0\right),$$

where $\mathbf{K}_0$ and $\sum_0$ are 3×3 matrices of genetic and residual (co)variances between the parameters of the Wilmink function, respectively. For both matrices, flat and bounded priors were adopted.

2.2.3. Heterogeneity of the residual variance

Based on model comparison results, the random regression model using the Ali-Schaeffer function (M2) was used to investigate the effect of allowing for dependence of liabilities within a cow in the course of lactation (M4). A first order autoregressive process (AR(1)) was assumed. In this model, a serial residual correlation pattern within-cows was adopted. The resulting residual (co)variance matrix, $\mathbf{R_0}$, has known diagonal elements (equal to 1) and the covariance (correlation) between liability at test-days i and j is given by:

$$\text{cov}(i, j) = \rho^{|i-j|},$$

where $\rho \in [-1, 1]$, so:

$$\mathbf{R_0} = \begin{bmatrix} 1 & \rho & \rho^2 & . & . & . & \rho^{10} \\ & 1 & \rho & \rho^2 & . & . & \rho^9 \\ & & . & . & . & . & . \\ & & & . & . & . & . \\ & & & & . & . & \rho^2 \\ & & & & & . & \rho \\ & & & & & & 1 \end{bmatrix}.$$

The prior distribution of ρ was uniform in $[-1, 1]$.

2.2.4. Implementation

A data augmentation algorithm was implemented. For each of the three models *M1*, *M2* and *M3*, the parameter vector was augmented with 3341 latent variables and with 171 genetic values of known ancestors in each of the three models. The conditional posterior distributions are in known form for all parameters, so Gibbs sampling is straightforward. The needed conditional distributions are normal for systematic, genetic and permanent effects ($\mathbf{b}$, $\boldsymbol{\beta}$, $\mathbf{u}$, $\mathbf{p}$), truncated normal for the 3341 latent variables, scaled inverted chi-square for the genetic variance in the first model (σ_u^2) and for the permanent effect variance in the three models (σ_p^2), and inverted Wishart for the 5×5 genetic (co)variance matrix in *M2* ($\mathbf{G_0}$) and the 3×3 genetic and residual (co)variance in *M3* ($\mathbf{K_0}$, $\sum_0$).

For the autoregressive random regression model (*M4*), all conditional posterior distributions, except for ρ, are in closed form. The sampling of liabilities is more involved because of the non-zero correlation between test-days that induces to truncated multivariate normal distribution. In order to sample from the multi-dimensional posterior density, $p(\mathbf{l}_i|\mathbf{y}_i, \boldsymbol{\beta}, \mathbf{u}, \mathbf{p}, \mathbf{R_0})$, a method of computation consisting of successive sampling from a set of truncated univariate

normal distributions was adopted. The univariate distributions involved in this sampling scheme have the form:

$$p(l_{ij}|\mathbf{l}_{-ij}, \mathbf{y}_i, \beta, \mathbf{u}, \mathbf{p}, \mathbf{R_0}),$$

where l_{ij} is the element j of the vector of liabilities for cow i and $\mathbf{l}_{-ij} = (l_{i1}, l_{i2}, \ldots, l_{i(j-1)})$ are the liabilities for cow i except l_{ij}. Inverse cumulative distribution sampling [6] was used to draw samples from the conditional posterior distribution of the latent variable. This technique is more efficient than a rejection-sampling scheme.

As noted before, the diagonal elements of the residual (co)variance matrix, $\mathbf{R_0}$, are set equal to one and hence, the only element of this matrix to sample in the autoregressive model is ρ. Its conditional distribution is not in a closed form, so a Metropolis-Hastings algorithm was used.

2.2.5. Comparison of Models

The Bayes factor, as defined by Newton and Raftery [11], was used to assess the plausibility of the models postulated. The marginal density of the data under each one of the models was estimated from the harmonic means of likelihood values evaluated at the posterior draws, that is:

$$\hat{p}(\mathbf{y}|M_i) = \left\{ \frac{1}{N} \sum_{j=1}^{N} \left[p\left(\mathbf{y}|\theta^{(j)}, M_i\right) \right]^{-1} \right\}^{-1},$$

where $\mathbf{y}$ is the vector of observed binary responses and $\theta^{(j)}$ is the Gibbs sampling sample j of parameters under model M_i. The estimated Bayes factor between models M_i and M_j is:

$$B_{M_i, M_j} = \frac{\hat{p}(y|M_i)}{\hat{p}(y|M_j)}.$$

2.2.6. Genetic parameters and model selection criteria

For **M2**, the genetic covariance for liability to mastitis between two times, t_i and t_j was defined to be:

$$COV(t_i, t_j) = \mathbf{V}'(t_i)G_0\mathbf{V}(t_j),$$

where $\mathbf{G}_0$ is the 5×5 genetic (co)variances for the random regression parameters and

$$\mathbf{V}'(t_i) = [1 \quad t_i \quad t_i^2 \quad \ln(t_i^{-1}) \quad \ln(t_i^{-1})^2].$$

This definition sets the strong assumption that genetic expression along lactation is completely driven by time, since $\mathbf{G}_0$ is not time-dependent.

For *M3*, the genetic covariance for liability to mastitis between two times was:

$$\text{cov}(t_i, t_j) = \mathbf{w}'(t_i)\mathbf{K}_0\mathbf{w}(t_j),$$

where $\mathbf{K}_0$ is a 3×3 genetic (co)variances matrix between the Wilmink function parameters and:

$$\mathbf{w}'(t_i) = [1 \quad t_i \quad \exp(-0.05t_i)].$$

Heritabilities of liability to mastitis using the three models were defined as:

$$h^2 = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_p^2 + 1}(M1),$$

$$h_t^2 = \frac{\mathbf{V}'(t)G_0\mathbf{V}(t)}{\mathbf{V}'(t)G_0\mathbf{V}(t) + \sigma_p^2 + 1}(M2),$$

$$h_t^2 = \frac{\mathbf{w}'(t)G_0\mathbf{w}(t)}{\mathbf{w}'(t)[\mathbf{K}_0 + \sum_0]\mathbf{w}(t) + \sigma_p^2 + 1}(M3).$$

For the model with the autoregressive structure, heritability is computed as for *M2*.

Longitudinal data allows developing new selection criteria other than a single predicted additive genetic value for a cow. Potential selection criteria from these models include, for example, the probability of no mastitis during lactation, the probability of at most a certain number of episodes, and the expected number of days a cow has mastitis. In this study, the following arbitrary criteria were used:

(a) Expected number of days with mastitis (MD):

$$MD_i = \sum_{j=14}^{300} \Phi(\mu_{ij})$$

$$0 \leqslant MD_i \leqslant 287.$$

(b) Probability of mastitis during lactation:

$$p_i(1) = 1 - \prod_{j=1}^{n_i}[1 - \Phi(\mu_{ij})].$$

(c) Probability of no mastitis during lactation:

$$p_i(0) = 1 - p_i(1).$$

(d) Expected fraction of days without mastitis (NMD):

$$NMD_i = \sum_{j=14}^{300}[1 - \Phi(\mu_{ij})] = 287 - MD_i.$$

(e) Probability of no mastitis at 30, 150 and 300 days:

$$p(y_{i30} = 0, y_{i150} = 0, y_{i300} = 0) = p_{i30}(0)p_{i150}(0)p_{i300}(0).$$

Some of the selection criteria are a function of the others. However, for demonstration purposes and to illustrate the flexibility of the models, all these quantities were inferred from their posterior means.

Computations were by Gibbs sampling and Metropolis-Hastings algorithms, with a burn-in period of 20 000 samples. Analysis was based on 50 000 additional samples, drawn without thinning.

3. RESULTS AND DISCUSSION

The posterior mean for heritability of liability to mastitis was 0.05 using the first model, where the breeding value was assumed constant along lactation. Even though the data set used in this analysis was too small to draw a definite conclusion on genetic parameters, this estimate was similar to the values found by [14]. Under Models 2, 3 and 4, heritability is a function of time. Figure 1 shows the variation of heritability throughout lactation using Models 2, 3 and 4. For Model 2, heritability was high at the beginning of lactation (0.27 at day 14), dropped quickly to reach values close to 0.05 in the middle of lactation, and then increased by the end of lactation. A similar pattern was observed using the random regression model for continuous test-day data (milk yield). A possible explanation for such behavior was attributed in part to the heterogeneity of the residual variances, and by assuming a constant permanent environmental effect along lactation. Given the small amount of information in our data set, we did not test the effect of applying a random regression for the permanent effect. The same pattern was observed for heritability throughout lactation using Model 3. However, this time, the heritability at the beginning of lactation was much higher (0.57 at day 14) compared with Model 2. Also, the lowest values for heritability were in the middle of lactation. With Model 4, the heritability was lower at the beginning of lactation (0.21 at day 14) compared with M2, although it is still high for the trait under analysis, indicating the necessity of treating the permanent effect as a function of time. For all four models, the posterior standard deviation of heritability was high and ranged from 0.006 to 0.18. The posterior mean and standard deviation of the correlation parameter were 0.19 and 0.07, respectively. This estimate indicates a low correlation between two successive test-days. The correlation declines quickly as the interval between test-days increases; it drops to 0.04 and 0.007 when the interval between test-days is around 60 and 90 days, respectively.

The Bayes factor was 0.011 between Model 1 and Model 2, 0.017 between Model 1 and Model 3 and 1.537 between Model 2 and Model 3. These results
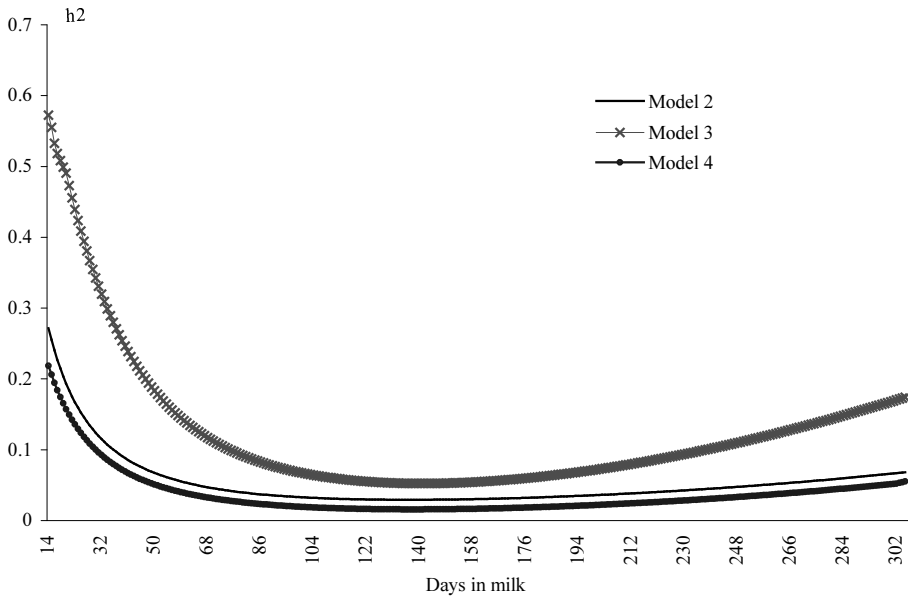
Figure 1. Heritabilities (posterior means) along lactation using Model 2, Model 3 and Model 4.

show that the data favored models with a genetic component in the regression (Model 2 and 3). Although, the results were not conclusive, Model 2 received more support by the data than Model 3. Comparing Model 2 and 4, the Bayes factor was 1.10 in favor of the latter, indicating that the heterogeneity of residual variance has to be postulated by the statistical model.

New selection criteria, including the expected number of days with mastitis, the probability of no mastitis during lactation and the probability of at most a certain number of episodes with mastitis, were computed for the best and worst cows using Model 2. Table II presents the posterior mode and high posterior density interval at 95% for the five best and worst cows for the expected number of days with mastitis, respectively. For the five best cows, the expected number of days with mastitis ranged between 8 and 9 days. At a phenotypic level, these cows did not experience any episodes of mastitis during their lactation. However, for the five worst cows, the expected number of days with mastitis ranged from 151 days for a cow having six episodes of mastitis during lactation to 223 days for a cow having 10 episodes of mastitis. Table III presents the probability of no mastitis at 30, 150 and 300 days for the best and worst five cows. Such probability was higher than 0.92 for the best cows, and, in fact, they never got mastitis on the mentioned dates. For the worst cow, the probability was lower than 0.01.

Table II. Posterior mode and HPD (95%) for the expected number of days with mastitis for the five best and worst cows using Model 2.

Cow	Mode	HPD (95%)
5 best cows		
12	7.9	3.7, 12.4
126	8.1	4.6, 12.9
314	8.5	4.8, 13.4
278	8.8	4.5, 12.9
181	9.0	5.2, 14.1
5 worst cows		
271	151.3	130.2, 176.6
115	165.9	141.1, 183.2
224	171.5	153.9, 195.6
143	177.5	155.2, 192.3
109	222.9	196.5, 251.8

Table III. Posterior mode and HPD (95%) for the probability of no mastitis at 30, 150 and 300 days of lactation.

Cow	Mode	HPD (95%)
5 best cows		
324	0.952	0.894, 0.979
9	0.941	0.871, 0.983
283	0.936	0.918, 0.992
131	0.932	0.833, 0.974
126	0.923	0.783, 0.957
5 worst cows		
224	0.011	0.003, 0.029
271	0.009	0.0016, 0.027
14	0.008	0.0014, 0.023
111	0.006	0.0022, 0.018
143	0.004	0.0008, 0.015

4. CONCLUSION

This study demonstrates the feasibility and advantage of a longitudinal analysis of sequential binary responses. A random regression model proved to be superior than a model with constant genetic value over time based on the Bayes factor results. The proposed model allowed not only a better modeling of systematic effects associated with each episode of mastitis, but also defining new selection criteria other than a single predicted additive genetic value for a cow. Furthermore, longitudinal data analysis accounts for the number of episodes of mastitis along lactation as well as their timing. Relaxing the

conditional independence assumption *via* an autoregressive process helped improve the results.

The methodology presented in this study is being applied to large data sets in Norway and Denmark.

REFERENCES

[1] Agresti A., A model for repeated measurements of a multivariate binary response, J. R. Statist. Assoc. 92 (1997) 315–322.

[2] Albert J., Chib S., Bayesian analysis of binary and polychotomous response data, J. R. Statist. Assoc. 88 (1993) 669–679.

[3] Ali T.E., Schaeffer L.R., Accounting for covariances among test days milk yield in dairy cows, Can. J. Anim. Sci. 67 (1987) 637.

[4] Coull B.A., Agresti A., Random effects modeling of multiple binary responses using the multivariate binomial logit-normal distribution, Biometrics 56 (2000) 73–80.

[5] Dempster E.R., Lerner I.M., Heritability of threshold characters, Genetics 35 (1950) 212–236.

[6] Devroy L., Non-uniform random variate generation, Springer, New York, 1986.

[7] Foulley J.L., Gianola D., Im S., Genetic evaluation for discrete polygenic traits in animal breeding, in: Gianola D., Hammond K. (Eds.), Advances in Statistical Methods for Genetic Improvement of Livestock, Springer-Verlag, 1990, pp. 361–396.

[8] Jamrozik J., Schaeffer L.R., Dekkers J.C.M., Random regression models for production traits in Canadian Holsteins. Proc. Open Session INTERBULL Annu. Mtg., Veldhoven, The Netherlands, Int. Bull Eval. Serv. Bull. No. 14 (1996) 124, Dep. Anim. Breed. Genet., SLU, Uppsala, Sweden.

[9] Gianola D., Theory and analysis of threshold characters, J. Anim. Sci. 54 (1982) 1079–1096.

[10] Gianola D., Foulley J.L., Sire evaluation for ordered categorical data with a threshold model, Génét. Sél. Évol. 14 (1983) 201–224.

[11] Newton A.M., Raftery A.E., Approximate Bayesian inference with the weighted likelihood Bootstrap, J. Roy. Stat. Soc. B 56 (1994) 3–48.

[12] Ptak E., Schaeffer L.R., Use of test day yields for genetic evaluation of dairy sires and cows, Livest. Prod. Sci. 34 (1993) 23–34.

[13] Rekaya R., Carabano M.J., Toro M.A., Use of test day yields for the genetic evaluation of production traits in Holstein-Friesian cattle, Livest. Prod. Sci. 57 (1999) 203–217.

[14] Rodriguez-Zas S.L., Gianola D., Shook G.E., Factors affecting susceptibility to intramammary infection and mastitis: An Approximate Bayesian Analysis, J. Dairy Sci. 80 (1997) 75–85.

[15] Sorensen D.A., Andersen S., Gianola D., Korsgaard I., Bayesian inference in threshold models using Gibbs sampling, Genet. Sel. Evol. 17 (1995) 229–249.

[16] Wright S., An analysis of variability in number of digits in an inbred strain of guinea pigs, Genetics 19 (1934) 506–536.