



Simultaneous genetic evaluation of simulated mastitis susceptibility and recovery ability using a bivariate threshold sire model

B. G. Welderufael, D. J. de Koning, L. L. G. Janss, J. Franzén & W. F. Fikse

To cite this article: B. G. Welderufael, D. J. de Koning, L. L. G. Janss, J. Franzén & W. F. Fikse (2016) Simultaneous genetic evaluation of simulated mastitis susceptibility and recovery ability using a bivariate threshold sire model, Acta Agriculturae Scandinavica, Section A — Animal Science, 66:3, 125-134, DOI: [10.1080/09064702.2016.1275761](https://doi.org/10.1080/09064702.2016.1275761)

To link to this article: <http://dx.doi.org/10.1080/09064702.2016.1275761>



© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 17 Jan 2017.



Submit your article to this journal [↗](#)



Article views: 365



View related articles [↗](#)



View Crossmark data [↗](#)

Simultaneous genetic evaluation of simulated mastitis susceptibility and recovery ability using a bivariate threshold sire model

B. G. Welderufael^{a,b*}, D. J. de Koning^a, L. L. G. Janss^b, J. Franzén^c and W. F. Fikse^a

^aDepartment of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden; ^bCenter for Quantitative Genetics and Genomics, Department of Molecular Biology and Genetics, Aarhus University, Tjele, Denmark; ^cDepartment of Statistics, Stockholm University, Stockholm, Sweden

ABSTRACT

The aim of this study was to develop a new approach for joint genetic evaluation of mastitis and recovery. Two mastitis incidences (0.28 and 0.95) measured via somatic cell count and three between traits genetic correlations (0.0, 0.2, and -0.2) were simulated for daughter group sizes of 60 and 240. A transition model was applied to model transitions between healthy and disease state. The RJMC package in DMU was used to estimate (co)variances. Heritabilities were consistent with the simulated value (0.039) for susceptibility and a bit upward biased for recovery. Estimates of genetic correlations were -0.055 , 0.205 , and -0.192 for the simulated values of 0.0, 0.2, and -0.2 , respectively. For daughter group size of 60, accuracies of sire EBV ranged from 0.56 to 0.69 for mastitis and from 0.26 to 0.48 for recovery. The study demonstrated that both traits can be modeled jointly and simulated correlations could be correctly reproduced.

ARTICLE HISTORY

Received 10 June 2016
Revised 12 December 2016
Accepted 13 December 2016

KEYWORDS

Dairy cow; health state; transition probability; udder health; recovery

1. Introduction

Mastitis, an inflammation of the mammary gland, is usually caused by bacterial infection. Occasionally, it might arise from chemical, mechanical, or thermal injuries. It is a common and costly disease in modern dairy farms (Halasa et al., 2007; Geary et al., 2012). In addition to its high incidence and economic importance (due to discarded milk, reduced milk production, culling cows, and treatments cost), dairy cows welfare and consumers' demand for antibiotics-free milk and milk products makes mastitis a concern both to the dairy farms and the community. The well-documented (Carlén et al., 2004; Koivula et al., 2005) genetic antagonism between milk production and clinical mastitis is an additional reason to include mastitis in a dairy cattle breeding goal. Developing better models for genetic evaluation of mastitis is among the top priorities for breeding dairy cows with improved mastitis resistance.

Somatic cell counts (SCCs) and recorded cases of clinical mastitis (CM) are the variables that are mostly used for genetic evaluations of mastitis (Franzén et al., 2012). The lack of routine records (Carlén et al., 2006) of mastitis

has led to the use of indirect but related traits, such as: SCC, udder conformation (type traits), milking speed, and electrical conductivity (Schukken et al., 1997). Among these related traits, SCC is widely accepted as a proxy for mastitis and considered as the best measure (Emanuelson et al., 1988; Uhler, 2009) due to its ease of recording and high genotypic and phenotypic correlation with mastitis (Emanuelson et al., 1988; Gernand and König, 2014).

Franzén et al. (2012) developed a model using transition probabilities to analyze SCC changes during lactation to assess genetic merit for mastitis susceptibility in dairy cows. The model was built on the idea that measurements are taken at regular intervals (say, weekly) and that for each measurement an individual cow is classified into one of two possible states (healthy or diseased), whereupon transition probabilities between these states are analyzed. This method models transitions to and from states of infection, i.e. both the disease susceptibility and the recovery process are considered, enhancing the genetic evaluation of mastitis. The SCC-based analysis by Franzén et al. (2012) ignored possible genetic correlation between susceptibility and recovery ability from

CONTACT B. G. Welderufael  berihu.welderufael@slu.se

© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

Table 1. Phenotypic records and selected genetic parameters used in simulations.

Mastitis cases/ lactation	Trait	Genetic variance	Variance of permanent environment (lactation)	Variance of temporary environment (week)	Herd variance	Number of observations
0.28	Mastitis susceptibility	0.036	0.252	0.64	0.072	1,170,000
	Recovery ability	0.036	0.252	0.64	0.072	10,300
0.95	Mastitis susceptibility	0.036	0.252	0.64	0.072	1,120,000
	Recovery ability	0.036	0.252	0.64	0.072	35,000

mastitis. Though both mastitis susceptibility and recovery ability were considered, the traits were analyzed separately with a single trait model. A simple product-moment correlation between estimated breeding values (EBVs) failed to reproduce different values of simulated genetic correlations between mastitis susceptibility and recovery ability. Literature on estimation of genetic correlation between mastitis susceptibility and recovery ability is not available. To be able to investigate whether a genetic correlation exists between mastitis susceptibility and recovery ability a bivariate model had to be developed.

The objectives of the current study were, therefore, (1) to develop a new approach for a joint genetic evaluation of mastitis susceptibility and recovery ability and evaluate its identifiability by estimating genetic correlation between mastitis susceptibility and recovery ability using a bivariate threshold sire model, and (2) to evaluate the effect of daughter group sizes and level of mastitis incidence on breeding value accuracies for given simulated values of genetic correlations between mastitis susceptibility and recovery ability.

2. Material and methods

2.1. Data simulation

Simulation software by Carlén et al. (2006), further developed by Franzén et al. (2012), was used to generate SCC observations for individual cows on the basis of simulated mastitis cases. Mastitis and SCC records were generated on a weekly basis for first lactation cows. A cow's mastitis history for the whole lactation was simulated on the basis of weekly values for susceptibility to become infected with and to recover from mastitis. A cow that was healthy in week $t-1$ was assumed to have mastitis in week t if the mastitis susceptibility for that week was above a certain threshold (that depended on the desired incidence; see below). A cow with mastitis in week $t-1$ was assumed to have recovered from mastitis in week t if the recovery ability for that week was above a certain threshold (that depended on the average recovery rate). The mastitis susceptibility and recovery ability for an individual cow were generated as the sum of a herd effect, the animal's breeding values, permanent environmental cow (lactation) effect, and a (weekly)

temporary environmental effect. The effect of herd on mastitis susceptibility and recovery ability was assumed to be uncorrelated, likewise for permanent and temporal environmental effects. Breeding values for mastitis susceptibility and recovery ability could be correlated (see below). All effects were assumed to follow normal distributions; variances for the random effects are in Table 1.

Weekly SCC observations were generated on the basis of the mastitis history for each cow separately. The level of SCC on non-mastitic test-days followed a baseline SCC lactation curve (denoted as $L(t)$) to which a random noise, generated as the exponential of a random normal deviate ($SD = 0.56$), was added (Figure 1). On mastitic test-days, the SCC level was elevated followed by a successive decline towards the baseline level. The pattern (Figure 2) is a generalization of the effects of *Staphylococcus aureus*, *E. coli*, *Streptococcus dysgalactiae*, and *Streptococcus uberis* on SCC, reported by de Haas et al. (2002).

Two levels of mastitis incidence or cases per lactation were considered: scenario one (0.28 cases/lactation) and scenario two (0.95 cases/lactation). Scenario one was assumed to reflect previous estimates of incidence of mastitis in field data of Swedish first-parity cows (Carlén et al., 2004; Franzén et al., 2012). The second scenario was not based on previous studies but was chosen to take into consideration for possibilities of higher frequencies of mastitis in multiparous cows and for merely theoretical interest to investigate the

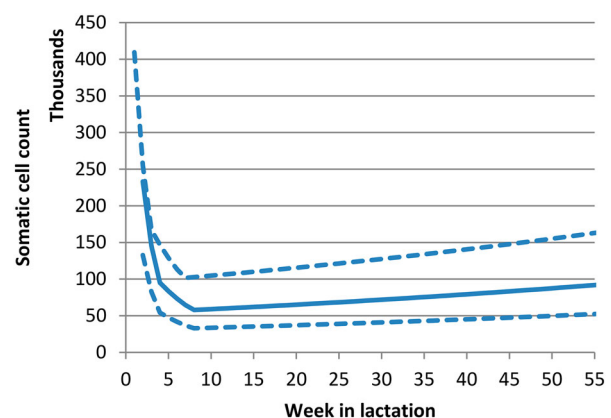


Figure 1. Baseline SCC (solid line) and baseline ± 1 SD (broken lines) for healthy cows.

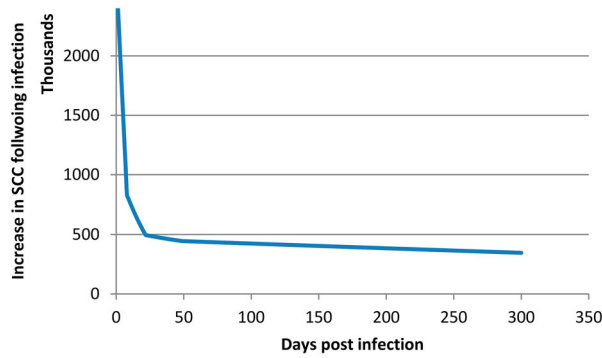


Figure 2. Increase of SCC due to mastitis as a function of the number of days post-infection.

difference in performance of the method for different levels of mastitis incidence (Franzén et al., 2012). Three different genetic correlations between mastitis susceptibility and recovery ability were considered: $r_g = 0.0$, $r_g = 0.2$, and $r_g = -0.2$. We chose these genetic correlations to make our simulations comparable to Franzén et al. (2012). The cow permanent environmental correlation between mastitis and recovery was simulated as zero for all scenarios. As described by Franzén et al. (2012), the simulated heritability for both mastitis susceptibility and recovery ability was 0.036 (within-herd heritability 0.039) on a weekly basis. This is similar to the heritability of mastitis of estimated with survival analysis by Carlén et al. (2005). The cows were daughters of either 400 or 100 unrelated sires distributed over 1200 herds with a fixed herd size of 20, which resulted into daughter group sizes of 60 and 240 respectively. All combinations of mastitis incidences, daughter group sizes, and genetic correlations were produced in 20 replicates. More details on the parameters used to simulate the data can be found in Franzén et al. (2012).

2.2. Transition probabilities

For a healthy, lactating cow, there is a risk of becoming infected and for a mastitic cow there is a possibility of recovering from mastitis, which can be conceptualized into probabilities of mastitis and recovery. For each cow, several transitions from healthy state (H) to diseased state (D), denoted as HD, and from D to H, denoted as DH, can occur within a lactation (Franzén et al., 2012). The probabilities of changing states or remaining in the current state can be summarized in a transition probability matrix, T_i , for cow i , like in Franzén et al. (2012):

$$T_i = \begin{bmatrix} 1 - \pi_i^{\text{HD}} & \pi_i^{\text{HD}} \\ \pi_i^{\text{DH}} & 1 - \pi_i^{\text{DH}} \end{bmatrix}, \quad (1)$$

where $1 - \pi_i^{\text{HD}}$ is the probability of remaining in the H state for cow i , π_i^{HD} the probability of moving from H to D state for cow i , π_i^{DH} the probability of moving from D to H state for cow i , and $1 - \pi_i^{\text{DH}}$ the probability of remaining in the D state for cow i .

In the transition probability matrix in Equation (1), the first row contains the probabilities of being in either of both states at time $t + 1$ for cow i in healthy state (H) at time t , and the second row contains probabilities of being in either of both states at $t + 1$ for cow i in diseased (D) at time t . The transition matrix is desired to have high values of $1 - \pi_i^{\text{HD}}$ (probability of remaining in the H state for cow i) and π_i^{DH} (probability of moving from the D state to H state, i.e. fast recovery if cow i had moved from H to D state), and consequently low values of π_i^{HD} (probability of moving from H to D state) and $1 - \pi_i^{\text{DH}}$ (probability of remaining in the D state) (Franzén et al., 2012).

Thus, a sequence of H's and D's, indicating whether or not a cow had mastitis on subsequent test weeks, can be converted into a new sequence of state changes: 0 if a cow remains in the same state and 1 if the cow changes state. These transitions result into a sequence of longitudinal records (repeated measurements) of mastitis susceptibility or recovery for a cow along the lactation, which can be used to make inferences about π_i^{HD} and π_i^{DH} .

2.3. Trait definition

Weekly SCC observations (generated from the simulated mastitis status) were used to assess whether a cow was in a diseased (D) or healthy (H) state. If a cow's observed weekly SCC exceeded a predefined boundary (denoted as $B(t)$), 2,000,000 cells/mL or $10 \times L(t)$, where $L(t)$ is the SCC level along the lactation curve), the cow was considered mastitic and in the D state; otherwise the cow was in the H state. As defined and explained in Franzén et al. (2012) a multiplier factor of 10 was used to lower the risk of misclassifications. Let SCC_{ijkt} be the observed SCC value for cow i , daughter of sire j , from herd k at time t , for $t = 1, 2, 3, \dots, t_i$ where t_i is the number of observations for cow i within the same lactation. A binary response h_{ijkt} stated whether the t^{th} observation of SCC for cow i was below (H) or above (D) the boundary as indicated in Figure 3. This binary response is formally expressed as follows:

$$h_{ijkt} = \begin{cases} 1 & \text{if } \text{SCC}_{ijkt} > B(t), \\ 0 & \text{if } \text{SCC}_{ijkt} \leq B(t). \end{cases} \quad (2)$$

Following the methods developed by Franzén et al. (2012) and described in the previous section "transition probabilities", two new datasets were created from the binary response in Equation (2). The first

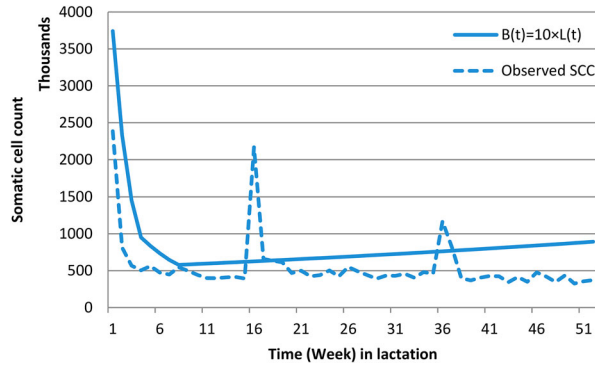


Figure 3. SCC level based boundary ($B(t)$, solid line) and observed SCC (dashed line) as a function of time in lactation (t , weeks after calving) for a given primiparous cow. According to the observed SCC, the cow in this figure has made a transition from healthy (H) to disease (D) state at week 15 and stayed being diseased for 2 weeks and recovered at week 18 to stay in H state until week 35 and moved to D state in the next two weeks and finally recovered to remain in H state throughout the remaining lactation period.

dataset contained transitions from healthy to diseased states (HD) and was used to make inference about mastitis susceptibility. The second dataset contained transitions from diseased to healthy states (DH) and was used to make inference about recovery ability. The HD dataset was much larger than the DH dataset because it included all cows, while the DH dataset contained only cows with at least one case of mastitis that thus had an opportunity to express recovery ability. All the transitions between states were recorded as a binary variable which indicates whether or not a transition took place between two consecutive observations, at time t and $t+1$. Thus, for each cow and lactation, two series of binary transition indicators were created according to:

$$y_{ijkt}^{HD} = \begin{cases} 1 & \text{if } h_{ijkt} = 0 \text{ and } h_{ijk(t+1)} = 1, \\ 0 & \text{if } h_{ijkt} = 0 \text{ and } h_{ijk(t+1)} = 0, \\ \text{missing} & \text{if } h_{ijkt} = 1 \text{ and } h_{ijk(t+1)} = 0 \text{ or } 1, \end{cases} \quad (3)$$

for the H to D transitions and

$$y_{ijkt}^{DH} = \begin{cases} 1 & \text{if } h_{ijkt} = 1 \text{ and } h_{ijk(t+1)} = 0, \\ 0 & \text{if } h_{ijkt} = 1 \text{ and } h_{ijk(t+1)} = 1, \\ \text{missing} & \text{if } h_{ijkt} = 0 \text{ and } h_{ijk(t+1)} = 0 \text{ or } 1, \end{cases} \quad (4)$$

for D to H transitions for $t = 1, 2, 3, \dots, t_i - 1$.

A cow without a case of classified mastitis did therefore not have any data for transitions in the D to H

direction. At the same time, the duration of a diseased state was usually much shorter than that of a healthy state. This resulted in a much smaller dataset for D to H transitions in comparison to the data in the H to D direction. Cow, sire, and herd indicators were repeated as many times as there were binary responses in the lactation in both transitions. Each sequence was just a replicate of the same indicator duplicated as many times as the length of y_{ijkt}^{HD} and y_{ijkt}^{DH} for susceptibility and recovery, respectively.

2.4. Statistical model

A bivariate threshold sire model was fitted to the longitudinal series of binary transitions indicators. When a trait is defined as a binary variable, the sire model is preferred over the animal model for convergence reasons (Carlén et al., 2005). The threshold model assumes the binary data are the outcome of a continuous variable on an unobserved scale called liability (Gianola and Foulley, 1983). The observed binary transitions take the value 1 if the liability is larger or smaller than the defined threshold and 0 otherwise as explained above. The variable for the observed transitions (y_{ijkt}^{HD} and y_{ijkt}^{DH}) was modeled with a bivariate probit threshold sire model describing susceptibility to and recovery from mastitis:

$$\Pr(y_{ijkt}^{HD} = 1) = \pi_{ijkt}^{HD} = \Pr(l_{ijkt}^{HD} < 0), \text{ and}$$

$$l_{ijkt}^{HD} = \beta^{HD} + H_k^{HD} + S_j^{HD} + C_{ij}^{HD} + e_{ijkt}^{HD}, \quad (5)$$

$$\Pr(y_{ijkt}^{DH} = 1) = \pi_{ijkt}^{DH} = \Pr(l_{ijkt}^{DH} < 0), \text{ and}$$

$$l_{ijkt}^{DH} = \beta^{DH} + H_k^{DH} + S_j^{DH} + C_{ij}^{DH} + e_{ijkt}^{DH}, \quad (6)$$

where

$y_{ijkt}^{HD} = 1$ for H to D transition, 0 for H to H transition, and missing otherwise for observation t .

$y_{ijkt}^{DH} = 1$ for D to H transition, 0 for D to D transition, and missing otherwise for observation t .

β^{HD} , β^{DH} = mean mastitis susceptibility and recovery ability respectively for an average cow.

H_k^{HD} , H_k^{DH} = random effect of herd k for mastitis susceptibility and recovery ability respectively.

C_{ij}^{HD} , C_{ij}^{DH} = random permanent effect of cow i of sire j for mastitis susceptibility and recovery ability.

S_j^{HD} , S_j^{DH} = random sire effect of sire j for mastitis susceptibility and recovery ability respectively.

e_{ijkt}^{HD} , e_{ijkt}^{DH} = random residual effect of cow i of sire j from herd k for the t th observation

The mean mastitis susceptibility and recovery ability were fitted as fixed effects. The random effects were assumed to be normally distributed with zero means

and (co)variances in a matrix notation:

$$\text{Var} \begin{bmatrix} H_k^{\text{HD}} \\ H_k^{\text{DH}} \\ C_{ij}^{\text{HD}} \\ C_{ij}^{\text{DH}} \\ S_j^{\text{HD}} \\ S_j^{\text{DH}} \\ e_{ijkt}^{\text{HD}} \\ e_{ijkt}^{\text{DH}} \end{bmatrix} = \begin{bmatrix} \mathbf{I}\sigma_{\text{HHD}}^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{I}\sigma_{\text{HDH}}^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{I}\sigma_{\text{CHD}}^2 & \mathbf{I}\sigma_{\text{CHD,DH}} & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{I}\sigma_{\text{CDH,DH}} & \mathbf{I}\sigma_{\text{CDH}}^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{I}\sigma_{\text{SHD}}^2 & \mathbf{I}\sigma_{\text{SHD,DH}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{I}\sigma_{\text{SDH,DH}} & \mathbf{I}\sigma_{\text{SDH}}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{I}\sigma_{\text{eHD}}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{I}\sigma_{\text{eDH}}^2 \end{bmatrix},$$

where $\mathbf{I}$ are identity matrices of appropriate size. The σ_{HHD}^2 , σ_{CHD}^2 , σ_{SHD}^2 , and σ_{eHD}^2 are variances for herd, cow permanent environmental effect, sire additive genetic, and residual effects, respectively, for mastitis susceptibility. The σ_{HDH}^2 , σ_{CDH}^2 , σ_{SDH}^2 , and σ_{eDH}^2 are variances for herd, cow permanent environmental effect, sire additive genetic, and residual effects, respectively, for recovery ability. The $\sigma_{\text{CHD,DH}}$ and $\sigma_{\text{SHD,DH}}$ stand for the permanent environmental and sire covariances, respectively, between mastitis susceptibility and recovery ability. However, as the residual variances are not identifiable for binary traits for both traits, it was fixed to one. The two traits were never observed simultaneously and the residual covariance between them was therefore constrained to zero.

2.5. Analysis, sampling, and Bayesian inference

Bayesian analysis with the RJMC package in DMU (Madsen and Jensen, 2013), which implements MCMC methods and Gibbs sampling, was performed to obtain the posterior distribution of all model parameters. Flat prior distributions were employed for all (co)variance components in the model. We ran single chains of 50,000 iterations with the first 10,000 iterations discarded as burn-in and a sampling interval of 25 to produce a posterior distribution of sample size of 2000 from which point estimates of parameters were derived. Trace and density plots as well as autocorrelations were (visually) inspected using the BOA software package (Smith, 2007) in R (R Core Team, 2015), to evaluate mixing and convergence of MCMC chains to the posterior distribution.

The bivariate model provides an option to estimate a genetic ($\hat{r}_g$) and permanent environmental ($\hat{r}_c$) correlations between the two traits. The correlations between mastitis susceptibility and recovery ability were calculated as follows:

$$\hat{r}_g = \frac{\hat{\sigma}_{\text{SHD,DH}}}{\sqrt{\hat{\sigma}_{\text{SHD}}^2 \times \hat{\sigma}_{\text{SDH}}^2}} \quad \text{for genetic correlation} \quad (7)$$

and

$$\hat{r}_c = \frac{\hat{\sigma}_{\text{CHD,DH}}}{\sqrt{\hat{\sigma}_{\text{CHD}}^2 \times \hat{\sigma}_{\text{CDH}}^2}} \quad (8)$$

for permanent environmental correlation

where

$\hat{\sigma}_{\text{SHD,DH}}$ and $\hat{\sigma}_{\text{CHD,DH}}$ are the posterior mean estimate of genetic and permanent environmental covariance, respectively, between the traits.

$\hat{\sigma}_{\text{SHD}}^2$ and $\hat{\sigma}_{\text{CHD}}^2$ are the posterior mean estimates of genetic and permanent environmental variances, respectively, for mastitis susceptibility.

$\hat{\sigma}_{\text{SDH}}^2$ and $\hat{\sigma}_{\text{CDH}}^2$ are the posterior mean estimates of genetic and permanent environmental variances, respectively, for recovery ability.

Heritability for each trait was calculated taking into account the sire model. Under this model heritability for mastitis susceptibility ($\hat{h}_{\text{HD}}^2$) was estimated using posterior means of variances as follows:

$$\hat{h}_{\text{HD}}^2 = \frac{4 \times \hat{\sigma}_{\text{SHD}}^2}{\hat{\sigma}_{\text{SHD}}^2 + \hat{\sigma}_{\text{CHD}}^2 + \hat{\sigma}_{\text{eHD}}^2}, \quad (9)$$

where

$\hat{h}_{\text{HD}}^2$ is the posterior mean heritability of mastitis susceptibility, $\hat{\sigma}_{\text{SHD}}^2$ is the posterior mean sire additive genetic variance for mastitis susceptibility, $\hat{\sigma}_{\text{CHD}}^2$ is the posterior mean permanent environmental variance of cow for mastitis susceptibility and $\hat{\sigma}_{\text{eHD}}^2$ the residual variance for mastitis susceptibility (fixed to 1).

The posterior mean heritability for recovery ability ($\hat{h}_{\text{DH}}^2$) was calculated as follows:

$$\hat{h}_{\text{DH}}^2 = \frac{4 \times \hat{\sigma}_{\text{SDH}}^2}{\hat{\sigma}_{\text{SDH}}^2 + \hat{\sigma}_{\text{CDH}}^2 + \hat{\sigma}_{\text{eDH}}^2}, \quad (10)$$

where

$\hat{h}_{\text{DH}}^2$ is the posterior mean heritability of recovery ability, $\hat{\sigma}_{\text{SDH}}^2$ is the posterior mean sire additive genetic variance for recovery ability, $\hat{\sigma}_{\text{CDH}}^2$ is the posterior mean permanent environmental variance of cow for recovery

ability and $\sigma_{e_{DH}}^2$ the residual variance for recovery ability (fixed to one).

Correlations between true breeding values (TBVs), generated during the simulation process, and the estimated breeding values (EBVs, posterior means) from the MCMC analysis were calculated to quantify the accuracy for both directions of the disease, that is susceptibility to mastitis and recovery ability. The estimation accuracy for the health to disease direction ($\hat{r}_{HD}$) was defined as the Pearson correlation between sire TBV and EBV as follows:

$$\hat{r}_{HD} = \frac{\hat{\sigma}_{TBV_{HD}, EBV_{HD}}}{\sqrt{\hat{\sigma}_{TBV_{HD}}^2 \times \hat{\sigma}_{EBV_{HD}}^2}}, \quad (11)$$

where $\hat{\sigma}_{TBV_{HD}, EBV_{HD}}$ is the covariance between TBV and EBV, $\hat{\sigma}_{TBV_{HD}}^2$ and $\hat{\sigma}_{EBV_{HD}}^2$ are the TBV and EBV variances for mastitis susceptibility.

The estimation accuracy for the disease to health direction ($\hat{r}_{DH}$) of the trait was defined as the Pearson correlation between sire TBV and EBV and calculated as follows:

$$\hat{r}_{DH} = \frac{\hat{\sigma}_{TBV_{DH}, EBV_{DH}}}{\sqrt{\hat{\sigma}_{TBV_{DH}}^2 \times \hat{\sigma}_{EBV_{DH}}^2}}, \quad (12)$$

where $\hat{\sigma}_{TBV_{DH}, EBV_{DH}}$ is the covariance between TBV and EBV, $\hat{\sigma}_{TBV_{DH}}^2$ and $\hat{\sigma}_{EBV_{DH}}^2$ are the TBV and EBV variances for recovery ability.

3. Results

3.1. Genetic correlation and heritability

In scenario one (low mastitis incidence), the mean (of 20 replicates) of estimated heritabilities for mastitis

susceptibility ranged from 0.035 to 0.037 and from 0.041 to 0.044 for the daughter group size of 60 and 240, respectively (Table II). In scenario two (high mastitis incidence), estimated heritabilities for mastitis susceptibility ranged from 0.036 to 0.042 over all daughter groups. The mean of the estimated heritabilities for recovery ability ranged from 0.036 to 0.066 over all cases of mastitis incidence and daughter groups (Table II). While the heritabilities for recovery ability were upward biased, the posterior modes of heritabilities were closer to the simulated (true) value (Table II).

Estimated genetic correlations between mastitis susceptibility and recovery ability ($\hat{r}_g$) were mostly in agreement with the simulated correlations (r_g), except in scenario one for the daughter group size of 60 (Table II). The mean (of 20 replicates) estimate of genetic correlation deviated by -0.055 , 0.005 , and 0.008 for the alternative with simulated values of $r_g = 0.0$, $r_g = 0.2$, and $r_g = -0.2$, respectively, over all daughter group sizes and mastitis incidences (Table II). For the daughter group size of 60, however, the genetic correlations were imprecise and deviated considerably, but not significantly from the simulated values. The width of the 95% HPD for $\hat{r}_g$ was larger than one for the alternatives with low mastitis incidence for the daughter group size of 60, and even for those (high mastitis incidence) alternatives with more information about recovery the 95% HPD for $\hat{r}_g$ was large with a mean of 0.69 (Table III). For 92.5%, 95%, and 96.25% of the replicates the true value for the genetic correlations between the traits was within the 95% HPD interval for $r_g = 0.0$, $r_g = 0.2$, and $r_g = -0.2$, respectively, over all daughter groups and mastitis incidences (Table IV).

Table II. Means and modes (SE in subscript) of genetic ($\hat{r}_g$) and permanent environmental ($\hat{r}_c$) correlations between mastitis susceptibility and recovery ability, heritability for mastitis susceptibility ($\hat{h}_{HD}^2$) and recovery ability ($\hat{h}_{DH}^2$).

	60 daughters/sire				240 daughters/sire			
	0.28 mastitis cases/lactation		0.95 mastitis cases/lactation		0.28 mastitis cases/lactation		0.95 mastitis cases/lactation	
	Mean	Mode	Mean	Mode	Mean	Mode	Mean	Mode
$\hat{r}_g$								
$r_g = 0.0$	-0.113 _{0.060}	-0.131 _{0.069}	-0.010 _{0.031}	-0.009 _{0.032}	-0.045 _{0.040}	-0.060 _{0.048}	-0.052 _{0.035}	-0.059 _{0.038}
$r_g = 0.2$	0.122 _{0.058}	0.160 _{0.067}	0.210 _{0.046}	0.217 _{0.048}	0.231 _{0.039}	0.262 _{0.044}	0.256 _{0.030}	0.264 _{0.031}
$r_g = -0.2$	-0.073 _{0.054}	-0.099 _{0.068}	-0.244 _{0.026}	-0.248 _{0.026}	-0.216 _{0.062}	-0.252 _{0.069}	-0.235 _{0.038}	-0.249 _{0.040}
$\hat{r}_c$								
$r_g = 0.0$	-0.047 _{0.017}	-0.021 _{0.013}	0.008 _{0.006}	0.008 _{0.006}	-0.144 _{0.061}	-0.123 _{0.068}	0.002 _{0.007}	0.002 _{0.007}
$r_g = 0.2$	-0.059 _{0.029}	-0.064 _{0.051}	0.023 _{0.006}	0.022 _{0.006}	-0.011 _{0.011}	-0.006 _{0.011}	0.018 _{0.006}	0.018 _{0.006}
$r_g = -0.2$	-0.085 _{0.014}	-0.050 _{0.009}	-0.014 _{0.006}	-0.015 _{0.006}	-0.056 _{0.013}	-0.035 _{0.011}	-0.010 _{0.007}	-0.010 _{0.007}
$\hat{h}_{HD}^2$								
$r_g = 0.0$	0.037 _{0.001}	0.036 _{0.001}	0.039 _{0.001}	0.038 _{0.001}	0.044 _{0.002}	0.041 _{0.002}	0.042 _{0.002}	0.040 _{0.002}
$r_g = 0.2$	0.037 _{0.002}	0.036 _{0.002}	0.038 _{0.001}	0.037 _{0.001}	0.041 _{0.002}	0.039 _{0.002}	0.041 _{0.002}	0.039 _{0.002}
$r_g = -0.2$	0.035 _{0.002}	0.034 _{0.002}	0.036 _{0.001}	0.036 _{0.001}	0.041 _{0.002}	0.038 _{0.002}	0.041 _{0.001}	0.039 _{0.001}
$\hat{h}_{DH}^2$								
$r_g = 0.0$	0.057 _{0.006}	0.042 _{0.006}	0.037 _{0.003}	0.034 _{0.003}	0.054 _{0.006}	0.046 _{0.006}	0.042 _{0.003}	0.039 _{0.003}
$r_g = 0.2$	0.057 _{0.006}	0.034 _{0.005}	0.036 _{0.003}	0.034 _{0.003}	0.050 _{0.004}	0.043 _{0.005}	0.041 _{0.002}	0.038 _{0.002}
$r_g = -0.2$	0.066 _{0.008}	0.048 _{0.008}	0.042 _{0.002}	0.040 _{0.003}	0.054 _{0.004}	0.044 _{0.003}	0.046 _{0.003}	0.044 _{0.003}

Table III. Length of 95% highest posterior density interval (SE in subscript) for genetic correlations between mastitis susceptibility and recovery ability for each simulated genetic correlation (r_g).

r_g	60 daughters/sire						240 daughters/sire					
	0.28 mastitis cases/lactation			0.95 mastitis cases/lactation			0.28 mastitis cases/lactation			0.95 mastitis cases/lactation		
	Width	Lower	Upper	Width	Lower	Upper	Width	Lower	Upper	Width	Lower	Upper
0.0	1.130 _{0.056}	−0.690 _{0.070}	0.440 _{0.075}	0.697 _{0.028}	−0.360 _{0.036}	0.337 _{0.038}	0.986 _{0.055}	−0.545 _{0.052}	0.441 _{0.048}	0.628 _{0.019}	−0.366 _{0.035}	0.262 _{0.036}
0.2	1.186 _{0.047}	−0.472 _{0.080}	0.714 _{0.062}	0.703 _{0.034}	−0.132 _{0.054}	0.570 _{0.050}	0.945 _{0.047}	−0.223 _{0.054}	0.722 _{0.036}	0.593 _{0.012}	−0.041 _{0.032}	0.551 _{0.027}
−0.2	1.123 _{0.061}	−0.649 _{0.058}	0.474 _{0.074}	0.663 _{0.023}	−0.590 _{0.030}	0.074 _{0.027}	0.857 _{0.044}	−0.650 _{0.059}	0.207 _{0.074}	0.574 _{0.015}	−0.521 _{0.033}	0.053 _{0.042}

Table IV. Proportion of replicates for which the 95% highest posterior density interval included the simulated value for genetic correlations between mastitis susceptibility and recovery ability ($\hat{r}_g$), heritability for mastitis susceptibility ($\hat{h}_{HD}^2$) and recovery ability ($\hat{h}_{DH}^2$) for each simulated genetic correlation (r_g).

	60 daughters/sire		240 daughters/sire	
	0.28 mastitis cases/lactation	0.95 mastitis cases/lactation	0.28 mastitis cases/lactation	0.95 mastitis cases/lactation
$\hat{r}_g$				
$r_g = 0$	0.80	1.00	0.95	0.95
$r_g = 0.2$	0.90	0.95	0.95	1.00
$r_g = -0.2$	0.90	1.00	0.85	0.90
$\hat{h}_{HD}^2$				
$r_g = 0$	0.90	1.00	0.95	1.00
$r_g = 0.2$	0.90	1.00	1.00	0.95
$r_g = -0.2$	0.80	0.90	1.00	1.00
$\hat{h}_{DH}^2$				
$r_g = 0$	0.95	0.95	0.90	0.85
$r_g = 0.2$	1.00	0.90	1.00	0.95
$r_g = -0.2$	0.95	1.00	1.00	1.00

Estimates of permanent environmental effect correlations ($\hat{r}_c$) between mastitis susceptibility and recovery ability were mostly negative. The permanent environmental correlations were very weak (close to 0) with the exception of scenario one for the daughter group size of 240 (Table II). A negative $\hat{r}_c$ (−0.144) was observed in scenario one for the daughter group size of 240.

3.2. Accuracies of breeding values

The mean of accuracies of breeding values for mastitis susceptibility, in scenario one, was 0.57 for the daughter

group size of 60 (Table V). For the daughter group size of 240, the accuracies ranged from 0.81 to 0.83. Similarly, in scenario two, the accuracy of breeding values for mastitis susceptibility markedly improved from 0.69 for the daughter group size of 60 to 0.89 for the daughter group size of 240 (Table V).

The accuracies of breeding values for recovery ability were, in all cases, low compared to those for mastitis susceptibility. Accuracies in scenario one ranged from 0.26 to 0.27 for the daughter group size of 60 and from 0.49 to 0.50 for the daughter group size of 240. In scenario two, the accuracy for recovery ability improved from a minimum value of 0.44 for the daughter group size of

Table V. Mean (SE in subscript) correlations between true and estimated breeding values of sires for mastitis susceptibility ($\hat{r}_{HD}$) and recovery ability ($\hat{r}_{DH}$) for each simulated genetic correlation (r_g).

	60 daughters/sire		240 daughters/sire	
	0.28 mastitis cases/lactation	0.95 mastitis cases/lactation	0.28 mastitis cases/lactation	0.95 mastitis cases/lactation
$\hat{r}_{HD}$				
$r_g = 0.0$	0.57 _{0.01}	0.69 _{0.01}	0.81 _{0.01}	0.89 _{0.00}
$r_g = 0.2$	0.56 _{0.01}	0.69 _{0.01}	0.83 _{0.01}	0.89 _{0.01}
$r_g = -0.2$	0.57 _{0.01}	0.69 _{0.01}	0.81 _{0.01}	0.89 _{0.00}
$\hat{r}_{DH}$				
$r_g = 0.0$	0.26 _{0.01}	0.46 _{0.01}	0.50 _{0.02}	0.71 _{0.01}
$r_g = 0.2$	0.27 _{0.01}	0.44 _{0.01}	0.50 _{0.01}	0.71 _{0.01}
$r_g = -0.2$	0.27 _{0.01}	0.48 _{0.01}	0.49 _{0.01}	0.72 _{0.01}

60 to a minimum value of 0.71 for the daughter group size of 240 (Table V).

4. Discussion

4.1. Genetic correlation and heritability

Our main objective was to propose a bivariate model for genetic evaluation of both; mastitis susceptibility and recovery ability in a joint analysis. We put emphasis on the correlations (r_g) between the traits to evaluate whether the model is identifiable. Franzén et al. (2012) studied whether or not the different values of simulated genetic correlations were reproducible from single trait analyses. Their results showed that the correlations between EBVs were very scattered. Though they confirmed the simulated pattern, the positive and zero values of r_g (0.2 and 0.0) were very scattered compared to the simulated values. However, for the simulated negative values of r_g (−0.2) their approach failed to reproduce the negative pattern of correlations. The current results match in pattern to the simulated values, with the exception of the alternative with small daughter group size and low mastitis incidence; disregarding this alternative, the mean estimated correlation deviated at most 0.05 from the simulated value. However, estimates were characterized by a relatively high degree of uncertainty (e.g. the length of the 95% HPD width for $\hat{r}_g$ in scenario one for daughter group size of 60 was larger than one). Nevertheless, we found that for a large proportion of the replicates, the true parameter value was within the 95% HPD interval. To the best of our knowledge the new approach for analyzing SCC data in the current study is a first approach to model mastitis susceptibility and recovery ability in a simultaneous analysis using a bivariate threshold sire model. Therefore, even though we are not able to further compare the results with previous studies, by reproducing the simulated correlations, we have demonstrated that the new approach is valid and capable to estimate genetic correlations between mastitis susceptibility and recovery ability, in a joint analysis. In situations where there is little information about recovery (e.g. due to low mastitis incidence) the genetic correlation can only be estimated with moderate precision.

The low estimated permanent environmental correlation ($\hat{r}_c$) between the traits is not unexpected as the effect of cow permanent environment on mastitis susceptibility and recovery ability was assumed uncorrelated during the data simulation. But, the expected value for $\hat{r}_c$ from the analysis is not zero, because the cow permanent environmental effect absorbs three quarters of the genetic variance in a sire model. The average of

$\hat{r}_c = -0.023$ (for $r_g = 0.0$) and $\hat{r}_c = 0.020$ (for $r_g = 0.2$) are, therefore, approximately consistent to the expected value of 0.019. For $r_g = -0.2$, the mean $\hat{r}_c = -0.047$ shows slight deviation from the expected value of −0.019. For the small daughter group size combined with low mastitis incidence deviations from expected value are higher. As observed for the genetic correlations, the lower precision, especially observed for the daughter group size of 60 in scenario one could be due to little information about recovery because of low mastitis incidence.

The heritability estimates for mastitis susceptibility inferred from the weekly SCC data were approximately in agreement with the simulated (within-herd) heritability of 0.039. In a study with the same simulation design, different models for genetic evaluation of mastitis (linear, threshold, and survival analysis) were compared and a range of heritability estimates were reported (Carlén et al., 2006). For survival analysis of time to first mastitis they reported a value of 0.037, which is very close to the simulated value of 0.039. This supports our estimates, because their survival analysis model is conceptually equivalent to analyzing state transitions with a probit model with repeated observations as in the current study. One difference, and advantage, is that our model allows for repeated mastitis cases, whereas their survival analysis did not.

Estimated mean of heritabilities for recovery ability were slightly upward biased and higher than for mastitis susceptibility. However, the posterior modes of heritabilities for both mastitis susceptibility and recovery ability were closer to the true value. It is clearly shown that posterior mode predicts better than posterior mean. A simulation study by Lipschutz-Powell et al. (2012) showed failure of conventional models (such as sire models) to capture the genetic variation in susceptibility even if it is present in a disease data, our new bivariate sire model did capture all the variations and resulted in slightly overestimated heritabilities. The over estimation can be due to the nature of the trait and the low number of observations (Table I) for the DH dataset. Recovery ability can only be observed if there has been mastitis, but with low incidence level the number of observations is low for the DH dataset.

Sire models with the MCMC method and Gibbs sampling (the model used in the current study) are pertinent in the genetic analysis of disease traits or binary data like mastitis (Kadarmideen et al., 2001). Such threshold models have been favored over the restricted maximum likelihood (REML) applied to generalized linear mixed models (GLMM) in several studies (e.g. Sørensen et al., 2009; Franzén et al., 2012; Gernand et al., 2012). This is because in threshold models where the trait of interest is defined on an underlying susceptibility scale,

there are theoretical advantages as they are more appropriate to depict the physiological background of a disease data compared to linear models and the associated model assumptions such as normality (Yin et al., 2014).

4.2. Accuracies of breeding values

The model generated more accurate EBVs for the higher mastitis frequency, where the proportion of false classified cases is expected to be lower (Franzén et al., 2012). Likewise, the increase in accuracy with the increase in the number of daughters (from 60 to 240) is expected as accuracy is highly dependent on the amount of performance information of an animal and information from its relatives. Especially in dairy cattle the large number of progeny evaluated is very useful to achieve a high level of accuracies. The accuracy (correlations between EBV and TBV) for the daughter group size of 60 was higher than the values reported by Franzén et al. (2012), despite that the simulation design was exactly the same. This could be traced back to a programming mistake in the conversion of health states to state transitions that affected the results in Franzén et al. (2012) negatively (Fikse and Franzén, pers. comm.), which was corrected in the current study. In an earlier analysis of single trait threshold sire model using the same simulation design, the accuracy of breeding values ranged from 0.53 to 0.60 (Carlén et al., 2006) for the daughter group of 60 for the lower mastitis frequency per lactation. The corresponding accuracy in the current study (ranging from 0.56 to 0.57) was consistent with that study.

The accuracies of breeding values for recovery ability compared to mastitis susceptibility were low in all cases. Accuracies of breeding values for recovery ability that ranged from 0.24–0.48 were reported by Franzén et al. (2012). The corresponding accuracy in the current study (ranging from 0.26 to 0.48) was a bit higher. We believe the improvement in accuracy is due to the correction of the previously mentioned programming mistake in Franzén et al. (2012). The other reason for the lower accuracies for recovery ability (compared to mastitis susceptibility) is the much reduced information in the DH data; because not all cows did get mastitis, which means small number of transitions from D to H state. The number of records for mastitis susceptibility (HD) was over one million per replicate, whereas it was around 10,000 and 35,000 for recovery ability (DH) for low and high mastitis frequency, respectively (Table I). Nevertheless, we have achieved much higher accuracies (as high as 0.72) by increasing the daughter group size to 240.

In general, as expected the accuracy of breeding values was highly dependent on daughter group sizes and the frequency of mastitis, like the findings by Franzén et al. (2012) where higher mastitis frequency combined with more daughters per sire resulted in higher accuracies. The different simulated values of genetic correlations ($r_g = 0.0$, $r_g = 0.2$, and $r_g = -0.2$) between mastitis susceptibility and recovery abilities had effect in neither the current nor in the previous study (Franzén et al., 2012). Accuracies of breeding values were not affected even with higher genetic correlations (e.g. $r_g = -0.5$ (unreported results)). The observed low heritabilities are one reason for this.

The method was tested on simulation data. In real life the data are expected to have different quality and quantity. Applying the method and investigating the analysis with real-life data will be the next step. One aspect is to expand models (5) and (6) with additional systematic effects, for example to allow for difference in mastitis susceptibility and recovery ability over the course of lactation. In addition, classification of cows as healthy or mastitic based on SCC data is critical, and algorithms like those proposed by Sørensen et al. (2016) offer potential for this purpose. It is also important to investigate the performance of the model to utilize direct information on CM and recovery instead of only SCC which is indirect measure of mastitis.

In conclusion, we have developed a new approach with a bivariate model for joint estimation of breeding values for susceptibility and recovery from mastitis on the basis of state transitions according to changes in SCCs. The developed bivariate model could be used to estimate possible genetic correlations between mastitis susceptibility and recovery. The high accuracies of breeding values (0.89 in scenario two) demonstrate the importance and potential of SCC in the genetic evaluation of mastitis susceptibility. Effect of level of mastitis incidence on accuracy of breeding value for both mastitis susceptibility and recovery ability was as important as the number of daughters per sire. The more daughters per sire combined with more mastitis incidence, the more accurate the breeding values were observed. The new approach with bivariate model gives an option for joint analysis of mastitis susceptibility and recovery ability. The model identifiability was shown by reproducing the simulated genetic correlations from the model analysis.

Acknowledgements

We are grateful to Yvette de Haas from the Animal Breeding and Genomics Centre of Wageningen UR Livestock Research for her input and comments on earlier

version of this manuscript. We thank Emma Carlén, Animal geneticist at Växa Sverige for her insightful comments.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

BGW benefited from a joint grant from the European Commission and Swedish University of Agricultural Sciences (SLU), within the framework of the Erasmus-Mundus joint doctorate EGS-ABG.

References

- Carlén, E., Emanuelson, U. & Strandberg, E. (2006) Genetic evaluation of mastitis in dairy cattle using linear models, threshold models, and survival analysis: A simulation study. *Journal of Dairy Science*, 89, 4049–4057.
- Carlén, E., Schneider, M. D. & Strandberg, E. (2005) Comparison between linear models and survival analysis for genetic evaluation of clinical mastitis in dairy cattle. *Journal of Dairy Science*, 88, 797–803.
- Carlén, E., Strandberg, E. & Roth, A. (2004) Genetic parameters for clinical mastitis, somatic cell score, and production in the first three lactations of Swedish holstein cows. *Journal of Dairy Science*, 87, 3062–3070.
- De Haas, Y., Barkema, H. W. & Veerkamp, R. F. (2002) The effect of pathogen-specific clinical mastitis on the lactation curve for somatic cell count. *Journal of Dairy Science*, 85, 1314–1323.
- Emanuelson, U., Danell, B. & Philipsson, J. (1988) Genetic parameters for clinical mastitis, somatic cell counts, and milk production estimated by multiple-trait restricted maximum likelihood. *Journal of Dairy Science*, 71, 467–476.
- Franzén, J., Thorburn, D., Urioste, J. I. & Strandberg, E. (2012) Genetic evaluation of mastitis liability and recovery through longitudinal analysis of transition probabilities. *Genetics Selection Evolution*, 44(1), 10.
- Geary, U., Lopez-Villalobos, N., Begley, N., McCoy, F., O'Brien, B., O'Grady, L. & Shalloo, L. (2012) Estimating the effect of mastitis on the profitability of Irish dairy farms. *Journal of Dairy Science*, 95, 3662–3673.
- Gernand, E. & König, S. (2014) Random regression test-day model for clinical mastitis: Genetic parameters, model comparison, and correlations with indicator traits. *Journal of Dairy Science*, 97, 3953–3963.
- Gernand, E., Rehbein, P., Von Borstel, U.U. & König, S. (2012) Incidences of and genetic parameters for mastitis, claw disorders, and common health traits recorded in dairy cattle contract herds. *Journal of Dairy Science*, 95, 2144–2156.
- Gianola, D. & Foulley, J. (1983) Sire evaluation for ordered categorical data with a threshold model. *Genetics Selection Evolution*, 15, 201–224.
- Halasa, T., Huijps, K., Osteras, O. & Hogeveen, H. (2007) Economic effects of bovine mastitis and mastitis management: A review. *Veterinary Quarterly*, 29, 18–31.
- Kadarmideen, H. N., Rekaya, R. & Gianola, D. (2001) Genetic parameters for clinical mastitis in Holstein-Friesians in the United Kingdom: A Bayesian analysis. *Animal Science*, 73, 229–240.
- Koivula, M., Mantysaari, E. A., Negussie, E. & Serenius, T. (2005) Genetic and phenotypic relationships among milk yield and somatic cell count before and after clinical mastitis. *Journal of Dairy Science*, 88, 827–833.
- Lipschutz-Powell, D., Woolliams, J. A., Bijma, P. & Doeschl-Wilson, A.B. (2012) Indirect genetic effects and the spread of infectious disease: Are we capturing the full heritable variation underlying disease prevalence? *PloS One*, 7(2012), e39551.
- Madsen, P. & Jensen, J. (2013) A user's guide to DMU. A package for analysing multivariate mixed models. Version 6, release 5.2. Available at: http://dmu.agrsci.dk/DMU/Doc/Current/dmuv6_guide.5.2.pdf
- R Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: <https://www.r-project.org/>
- Schukken, Y. H., Lam, T. & Barkema, H. W. (1997) Biological basis for selection on udder health traits. Proceedings of the International workshop on genetic improvement of functional traits in cattle health. Uppsala, Sweden, June 8–10, *Bulletin Interbull*, 15, 27–32.
- Smith, B. J. (2007) BOA: An R package for MCMC output convergence assessment and posterior inference. *Journal of Statistical Software*, 21, 1–37.
- Sørensen, L. P., Bjerring, M. & Løvendahl, P. (2016) Monitoring individual cow udder health in automated milking systems using online somatic cell counts. *Journal of Dairy Science*, 99, 608–620.
- Sørensen, L. P., Mark, T., Madsen, P. & Lund, M. S. (2009) Genetic correlations between pathogen-specific mastitis and somatic cell count in Danish Holsteins. *Journal of Dairy Science*, 92, 3457–3471.
- Uhler, C. (2009) Mastitis in dairy production: Estimation of sensitivity, specificity and disease prevalence in the absence of a gold standard. *Journal of Agricultural Biological and Environmental Statistics*, 14, 79–98.
- Yin, T., Bapst, B., Von Borstel, U. U., Simianer, H. & König, S. (2014) Genetic analyses of binary longitudinal health data in small low input dairy cattle herds using generalized linear mixed models. *Livestock Science*, 162, 31–41.