J. Dairy Sci. 97:1–11
http://dx.doi.org/10.3168/jds.2013-7830

Random regression test-day model for clinical mastitis: Genetic parameters, model comparison, and correlations with indicator traits

E. Gernand* and S. König†[1]
*Thuringian State Institute of Agriculture, 07743 Jena, Germany
†Department of Animal Breeding, University of Kassel, 37213 Witzenhausen, Germany

ABSTRACT

The objective was to study genetic (co)variance components for binary clinical mastitis (CM), test-day protein yield, and udder health indicator traits [test-day somatic cell score (SCS) and type traits of the udder composite] in the course of lactation with random regression models (RRM). The study used a data set from selected 15 large-scale contract herds including 26,651 Holstein cows. Test-day production and CM data were recorded from 2007 to 2012 and comprised parities 1 to 3. A longitudinal CM data structure was generated by assigning CM records to adjacent official test dates. Bivariate threshold-linear RRM were applied to estimate genetic (co)variance components between longitudinal binary CM (0 = healthy; 1 = diseased) and longitudinal Gaussian distributed protein yield and SCS test-day data. Heritabilities for liability to CM (heritability ~0.15 from 0 to 305 d after calving) were slightly higher than for SCS for corresponding days in milk (DIM) in the course of lactation. Daily genetic correlations between CM and SCS were moderate to high (genetic correlation ~0.70), but substantially decreased at the very end of lactation. Genetic correlations between CM at different test days were close to 1 for adjacent test days, but were close to zero for test days far apart. Daily genetic correlations between CM and protein yield were low to moderate. For identical DIM (e.g., DIM 20, 160, and 300), genetic correlations were −0.03, 0.11, and 0.18, respectively, and disproved pronounced genetic antagonisms between udder health and productivity. Correlations between estimated breeding values (EBV) for CM from the RRM and official EBV for linear type traits of the udder composite, including EBV from 74 influential sires (sires with >60 daughters), were −0.31 for front teat placement, −0.01 for rear teat placement, −0.31 for fore udder attachment, −0.32 for udder depth, and −0.08 for teat length. Estimated breeding values for CM from the RRM were compared with EBV from a multiple-trait model and with EBV from a repeatability model. For test days covering an identical time span and on a lactation level, correlations between EBV from RRM, multiple-trait model, and repeatability model were close to 1. Most relevant results suggest the routine application of threshold RRM to binary CM to (1) allow selection of genetically superior sires for distinct stages of lactation and (2) achieve higher selection response in CM compared with selection strategies based on indicator type traits or based on the indicator-trait SCS.

Key words: clinical mastitis, random regression model, indicator trait

INTRODUCTION

Test-day models have been the standard methodological tool used in official genetic evaluations for Gaussian distributed longitudinal production traits and SCS for almost 2 decades. The first relevant scientific studies for the estimation of genetic parameters in the course of lactation were accomplished by Meyer et al. (1989) and by Ptak and Schaeffer (1993) for milk yield, and by Reents et al. (1995) for SCS. Strategies for analyzing longitudinal test-day data were summarized and compared by Swalve (2000). Strategies included the comparison of repeatability models with multiple-trait models, and the comparisons of test-day models using fixed regressions, covariance functions, and random regressions. Random regression models (**RRM**) present the genetic trajectory of the longitudinal trait in the course of a time-dependent or an environmental continuous descriptor, or both (e.g., Brügemann et al., 2011). Breeding values for each test day allow the use of different selection criteria (e.g., the selection of animals characterized by desired measurements for lactation persistency; Jamrozik et al., 1997). In most RRM applications to longitudinal data, regressions were modeled based on Legendre polynomials. Baldi et al. (2010) used RRM with Legendre polynomials to fit all random components, including direct genetic, maternal genetic,

---

Received December 13, 2013.
Accepted February 24, 2014.
[1] Corresponding author: sven.koenig@uni-kassel.de

animal permanent environmental, and maternal permanent environmental effects.

Random regression models also were applied to longitudinal binary data structures by including threshold methodology. Random regression threshold sire models were fitted for the genetic analyses of binary conception rate by DIM (Tsuruta et al., 2009), and by temperature × humidity indices (Brügemann et al., 2013). A recent focus on the implementation of health data-recording systems enables substantial data sets for health traits as a prerequisite for RRM applications also to binary health data. Among health traits, priority is on genetic improvements for clinical mastitis (**CM**). This is due to the obvious effect of CM on involuntary dairy cow disposals, and due to the detrimental effect of CM on dairy farming profitability and economy (e.g., Sadeghi-Sefidmazgi et al., 2011).

The direct inclusion of CM into an overall breeding goal, the improvement of breeding goals, and the definition of breeding strategies requires knowledge of genetic correlations with production traits, and with currently used indicator traits for udder health. The most important indicator trait for CM is SCS, and estimates for genetic correlations ($\mathbf{r_g}$) between CM and SCS from lactation models were $r_g \approx 0.70$ (e.g., Emanuelson et al., 1988). Indicator traits for CM additionally include linear type traits of the category "udder." Among linear type traits, best predictors were udder depth, udder attachment, and udder balance (Rupp and Boichard, 1999). The direct inclusion of health traits into an overall breeding goal raises the question about the necessity of indicator trait recording. König et al. (2013) applied selection index methodology to phenotypic and genomic information sources. In their study, selection response in the targeted health trait decreased with a stepwise increase of health indicator traits in the breeding goal. In the study by de Haas et al. (2002), an indirect selection strategy based on milk yield and SCS was ineffective in reducing simultaneously incidences of the whole spectrum of specific major pathogens for CM.

Following those aspects for reducing CM incidence by genetic selection, the objective of the present study was 4-fold: to (1) to estimate genetic (co)variance components of CM in the course of lactation, (2) estimate genetic correlations between CM and milk yield in the course of lactation, (3) estimate correlations between breeding values for CM and breeding values for indicator type traits in the course of lactation, and (4) compare EBV for CM from the threshold RRM with EBV for CM from more robust repeatability models (**RM**) and multiple-trait models (**MTM**).

MATERIALS AND METHODS

Data

Holstein cows used in this study were the same as used in the study by Gernand et al. (2013) for the genetic analysis of claw disorders in the course of lactation. Also for CM, data recording in a "contract herd system," including 15 large-scale herds in the state of Thuringia, was based on the diagnosis key by Feucker and Staufenbiel (2003). This diagnosis key is implemented in electronically on-farm recording systems. The contract herd system implies a strong collaboration between the breeding organization and dairy cattle farms, and advantages of such a system in terms of health data quality and quantity was outlined in several previous studies (e.g., Gernand et al., 2012, 2013). Data recording spanned a period of 5 yr from 2007 to 2012, and included 26,651 Holstein cows from parities 1 to 3. Longitudinal CM data were generated as described in detail by Gernand et al. (2013) and illustrated in Figure 1. The first and major feature of this strategy is to create time windows between consecutive official test days. Official test days are in monthly intervals and, consequently, the length of time intervals is approximately 30 d. In the second step, we assigned CM cases to exactly the date of the adjacent official test day. Hence, diseased cows (1 or more CM cases within a given interval) received a score of 1 and healthy cows received a score of 0 at the official test date. Based on this methodology, a substantial data set of repeated binary measurements was created (Table 1). The number of test days per herd ranged from 11,967 to 63,194. Furthermore, for the application of MTM, we divided the lactation into 4 periods (Figure 1): period I comprised DIM 1 to 60, period II DIM 61 to 180, period III DIM 181 to 270, and period IV DIM 271 to 360.

The genetic structure of the contract herd cows was characterized as follows: the 26,651 cows were from 1,258 different sires, which implies, on average, 21 daughters per sire and 74 sires with more than 60 daughters for all traits. National official EBV from these 74 sires were available from the genetic evaluation of August 2012. In the present study, we used official EBV for the overall index, the conformation index, the longevity index, the SCS index (**RZS**), the production index, the milking speed index (**RZD**), and the udder type trait index as well as relative breeding values for linear type traits of the udder composite, including rear udder height, central ligament, front teat placement (**FTP**), rear teat position (**RTP**), fore udder attachment, udder depth, and teat length (**TL**). Relative official national breeding values are standardized to a mean of 100 and a