Genomic evaluation for calf wellness traits in Holstein cattle

Zoetis Genetics, Kalamazoo, MI 49007

ABSTRACT

Reducing calf morbidity and mortality is important for attaining financial sustainability and improving animal welfare on commercial dairy operations. Zoetis (Kalamazoo, MI) has developed genomic predictions for calf wellness traits in Holsteins that include calf respiratory disease (RESP; recorded between 0 and 365 d of age), calf scours (DIAR; recorded between 2 and 50 d of age), and calf livability (DEAD; recorded between 2 and 365 d of age). Phenotype and pedigree data were from commercial dairies and provided directly by producers upon obtaining their permission. The number of records ranged from 741,484 for DIAR to 1,926,261 for DEAD. The number of genotyped animals was 325,025. All traits were analyzed using a univariate threshold animal model including fixed effect of year of birth × calving season × region, and random effects of herd × year of birth and animal. A total of 45,425 SNP were used in genomic analyses. Animals genotyped with low-density chips were imputed to the required number of SNP. All analyses were conducted using single-step genomic BLUP implementing the “algorithm for proven and young” (APY) animals designed to accommodate very large numbers of genotypes. Estimated heritabilities were 0.042, 0.045, and 0.060 for RESP, DIAR, and DEAD, respectively. The genomic predicted transmitting abilities ranged between −8.0 and 24.0, −11.5 and 28.5, and −6.5 to 22.8 for RESP, DIAR, and DEAD, respectively. Reliabilities of breeding values were obtained by approximation based on partitioning of a function of reliability into contributions from records, pedigree, and genotypes, where the genotype contribution was approximated using the diagonal value of the genomic relationship matrix. The average reliabilities for the genotyped animals were 41.9, 42.6, and 47.3% for RESP, DIAR, and DEAD, respectively. Estimated genomic predicted transmitting abilities and reliabilities were approximately normally distributed for all analyzed traits. Approximated genetic correlations of calf wellness with Zoetis dairy wellness traits and traits included in the US national genetic evaluation were low to moderate. The results indicate that direct evaluation of calf wellness traits under a genomic threshold model is feasible and offers predictions with average reliabilities comparable to other lowly heritable traits. Genetic selection for calf wellness traits presents a compelling opportunity for dairy producers to help manage herd replacement costs and improve overall profitability. Key words: calf wellness, respiratory disease, scours, calf livability, genomic predictions

INTRODUCTION

Replacement costs are one of the largest financial components on a dairy farm. Costs of raising a calf from birth to first calving have been estimated at $1,200 to over $2,000 (Rossini, 2004). A heifer that experiences a disease at least once has 6% higher rearing costs than a healthy heifer (Mohd Nor et al., 2012). Therefore, keeping calves healthy and minimizing mortality and morbidity are key investments with real future returns that may mean the difference between profit or loss in tight margin years. In a study by Sischo et al. (1990), calf disease costs represented 4% of the total cost a cow incurred during her lifetime. Diarrhea and pneumonia were responsible for 86% of calf disease costs. In spite of improved management and calf-rearing practices, preweaning death loss in dairy calves range from 7.8 to 12%; 53% of those losses are due to digestive problems (scours) and 21% to respiratory diseases (Murray, 2011). Even if the calf survives and recovers from the disease, its performance as a mature cow will be affected. Occurrence of a calfhood disease increases the age at first calving by up to 2 mo, reduces survival through the first and second lactations, and increases culling due to mastitis and other diseases (Rossini, 2004).

Producing calves that are robust and able to thrive under the challenges of modern dairy operations is essential for both the economics of the dairy industry and the welfare of the animals (Gulliksen et al., 2009). In addition to improving herd management practices, increasing calf disease resistance genetically is essential for the success of dairy operations. Using genetic selection as a tool to improve calf wellness in dairy herds has been researched but, so far, not implemented. The only
traits available in dairy genetic evaluations focusing on characteristics of the calf are stillbirth in the United States and calf survival in Canada (Henderson et al., 2011). Several studies based on limited numbers of animals have indicated that calf health and survival are genetically controlled and that the genetic component, although small, can be successfully exploited in breeding programs. Most heritability estimates for calf respiratory diseases were low, ranging from 0.04 in Holstein calves in Ontario (McCorquodale et al., 2013) to 0.05 in Norwegian Red calves (Heringstad et al., 2008) to 0.09 in Holstein calves in New York (Henderson et al., 2011), and up to 0.21 when estimated using a genomic relationship matrix (Neibergs et al., 2014). Similar heritabilities were obtained for scours, bloat, and other digestive diseases in dairy calves (Henderson et al., 2011; McCorquodale et al., 2013). Heritability estimates for calf mortality (or survival) have been lower, ranging from 0.001 to 0.008 in Danish Holsteins (Hansen et al., 2003) and 0.004 for a “heifer livability” trait currently researched by USDA scientists (Norman, 2016).

Another obstacle in implementing genetic evaluations for calf health is the lack of routinely recorded data. Dairy farms, including calf-raising facilities, usually record calf health events and treatment using their herd management software, but many factors influencing calf health and survival, such as birth weight, amount of colostrum fed, blood protein level, and so on, are not routinely measured and recorded. However, previous studies have shown that producer-recorded data, albeit imperfect, can be successfully used in genomic evaluations of dairy cow wellness traits (e.g., Zwald et al., 2004; Parker Gaddis et al., 2014; Vukasinovic et al., 2017). The inclusion of genomic data substantially improves reliabilities for these traits (Parker Gaddis et al., 2014; Vukasinovic et al., 2017).

The objectives of this study were to explore the genetic background of calf wellness traits, to develop a system for genomic evaluation for those traits in Holstein dairy calves based on producer-recorded data, and to estimate genetic correlations among calf wellness traits and traits in the national dairy genetic evaluation as well as Zoetis dairy wellness traits (Zoetis, Kalamazoo, MI).

**MATERIALS AND METHODS**

**Phenotype Data**

Phenotype data were obtained directly from dairy producers upon obtaining their permission. As of March 2018, data from 326 herds located in 24 US states was available. The characteristics of the data are shown in Table 1.

Backup files produced using herd management software were delivered to Zoetis, opened, and processed using internally written scripts. Initially, all events up to 1 yr of age were extracted to “format CH files,” files identical in format to the standard USDA format 6 (https://redmine.uscdb.com/projects/cdcb-customer-service/wiki/Format_6). As of March 2018, a total of 3,288,342 events recorded during the first year of life were available. Pedigree information was extracted from the backup files.

Out of the events typically occurring at young age, respiratory problems (RESP), pneumonia (PNEU), scours/diarrhea (DIAR), and mortality (DEAD) were chosen for further analyses. Umbilical hernias or bloat, although important calf health traits, were not considered because too few records were available. Further, respiratory events (RESP and PNEU) were merged into RESP because some herds recorded only “RESP,” some only “PNEU,” and some both.

**Trait Definition.** The incidence of calf health disorder was plotted against the age when a certain calf health event was first recorded. Distributions of RESP/PNEU, DIAR, and DEAD events by age from 0 to 365 d are given in Figures 1, 2, and 3, respectively.

The distribution of respiratory diseases shows 4 distinct peaks: the first, and highest, occurs very shortly after birth, likely due to aspiration pneumonia (Poulsen and McGuirk, 2009); the second peak occurs during the first week of life; the third peak happens around 35 d of age, and the fourth, and largest, occurs between

<table>
<thead>
<tr>
<th>Trait</th>
<th>Acronym</th>
<th>Description</th>
<th>Records</th>
<th>Incidence (%)</th>
<th>No. of herds</th>
<th>Average herd size1</th>
<th>G+P2</th>
<th>Pedigree3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calf respiratory disease</td>
<td>RESP</td>
<td>RESP + PNEU4 0–365 d of age</td>
<td>1,331,626</td>
<td>21.0</td>
<td>188</td>
<td>7,083</td>
<td>67,289</td>
<td>2,310,723</td>
</tr>
<tr>
<td>Calf scours</td>
<td>DIAR</td>
<td>DIAR 2–50 d of age</td>
<td>741,484</td>
<td>26.1</td>
<td>118</td>
<td>6,283</td>
<td>70,275</td>
<td>1,583,960</td>
</tr>
<tr>
<td>Calf livability</td>
<td>DEAD</td>
<td>DEAD 2–365 d of age</td>
<td>1,926,261</td>
<td>4.7</td>
<td>233</td>
<td>8,267</td>
<td>105,012</td>
<td>2,923,038</td>
</tr>
</tbody>
</table>

1Average number of phenotypic records per herd across all years of data.
2Number of animals with a genotype that also had a phenotype for each trait.
3Number of animals in the final pedigree for each trait.
4Pneumonia.
about 60 and 100 d of age, and is most likely caused by moving animals from individual housing to group pens (Hulbert and Moisá, 2016). After that, the incidence of respiratory diseases gradually decreases up to 365 d. Incidence of scours or diarrhea peaks at about 10 d of age and reduces sharply after that. Mortality is high during the first 2 d of life, and again around 7 to 10 d, after which it gradually decreases.

To cover the time periods when the respiratory disease, scours, and death are most likely to happen in young animals and have most economic impact, the traits for further analysis were defined as shown in Table 1.

Calf respiratory disease (RESP) was defined to cover the entire period from 0 to 365 d of age. Calf scours (DIAR) was limited to the range of 2 to 50 d of age, because after that, the incidence of scours was negligible. Calf livability (DEAD) was defined to describe all mortality cases recorded between 2 and 365 d of age (the first 2 d of life were omitted to avoid confounding early deaths with stillbirth).

**Data Editing.** The information from the CH files (health events for calves) was merged with the pedigree information to include all healthy herdmates of the affected animals. “Healthy” animals were defined as animals not recorded for a particular trait. All traits were treated as binary events, having a value of 1 if the animal was recorded as having a disorder (or died) and 0 otherwise.

All animals having the event “SOLD” in the first week of life as the only event were removed from the data. All male calves, as well as animals with events recorded before their birth date, were removed. Records were also removed for all animals (healthy and sick) that did not reach the opportunity period, which was defined as the upper age limit for each trait and calculated as the number of days between the birth and the date when the farm backup file was created. At this point, all non-Holstein calves were also removed.

Finally, for each trait, herd × year groups with fewer than 10 records were removed. Also removed were herd × year groups with the incidence of disorder being <0.5% or >95% to avoid underreporting or matching cases with insufficient number of healthy herdmates. The final numbers of animals and records for each trait with phenotypes and average incidence are in Table 1.

**Genotype Data**

Animals from commercial herds submitted to Zoetis for genomic testing were genotyped with various versions of low-density and medium-density chips. Raw genotypes were edited following the criteria described in
Statistical Model

Each trait was analyzed separately, using the following univariate threshold animal model:

$$\lambda = X\beta + Z_b h + Z_a a + e,$$

where $\lambda$ represents a vector of the animals’ unobserved liabilities to the given disorder; $\beta$ is the vector of fixed effects of year $\times$ season of birth $\times$ region combination (years 1996 to 2017 were considered); 4 birth seasons were defined within each year: Winter (Dec–Feb), Spring (Mar–May), Summer (Jun–Aug), and Fall (Sep–Nov); and 5 geographical regions based on climate were defined as in VanRaden et al. (2004); $h$ is the random herd-year effect, where $h \sim N(0, \sigma^2_h)$, with variance $\sigma^2_h$; $I$ is the identity matrix; $a$ is the random animal effect, with $a \sim N(0, \sigma^2_a)$, where $\sigma^2_a$ is the additive genetic variance and $H$ is the pedigree relationship matrix augmented using genotypes; $X$, $Z_b$, and $Z_a$ are incidence matrices corresponding to the fixed effects in $\beta$ and the random effects of herd $\times$ year and animal, respectively; and $e$ is the residual.

Variance Component Estimation

Variance components for calf wellness traits were estimated using the same model but, for simplicity, without genomic information. Previous studies have shown that differences between values of heritabilities estimated with and without including markers were not relevant (e.g., Parker Gaddis et al., 2014; Gonzalez-Pena et al., 2018). The variance components were estimated using ASReml version 4.1 (Gilmour et al., 2015), using the REML methodology. All binary traits were analyzed using generalized linear models with a binomial distribution and a link function of LOGIT where the residual variance to be 1, the variance components estimated by ASReml using the binomial model with the logit link function were adjusted to the residual variance of 1.00 by dividing each component by 3.29, so that their ratios remained the same. Heritabilities and their standard errors were estimated using the “pin” function in ASReml, and the formula $\text{Var}(g)/[\text{Var}(g + \text{Var}(ly) + \text{Var}(e))]$, where $\text{Var}(g)$ and $\text{Var}(ly)$ represent additive genetic and herd-year variances, respectively, and $\text{Var}(e)$ equals $1/3\pi^2$ or 3.29.

Before using ASReml software, we attempted to implement the threshold model in the program “thrgibbs90”, but it showed convergence problems, and some parameters were out of range when different options were tried for the model (data not shown), probably because of a large number of levels of the fixed effect of year $\times$ season $\times$ region (Miszal, 2005).

Genomic Evaluation

Genomic evaluation was performed using the programs from the BLUPF90 family (Miszal et al., 2002). A univariate threshold animal model based on single-step genomic BLUP methodology (ssGBLUP) was applied to all traits. In ssGBLUP, the inverse of the traditional pedigree relationship matrix, $A^{-1}$, is replaced by the inverse of the $H$ matrix that combines pedigree and genomic relationships (Legarra et al., 2009; Aguilar et al., 2010):

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ \tau G^{-1} - \omega A^{-1} \end{bmatrix},$$

where $A^{-1}$ is an inverse of the pedigree relationship matrix, $G^{-1}$ is an inverse of the genomic relationship matrix (VanRaden, 2008), $A_{22}^{-1}$ is an inverse of the pedigree relationship matrix for genotyped animals only; and $\tau$ and $\omega$ are scaling factors to condition the genomic relationship matrix to be compatible with the pedigree information that were both set to 1.0 (Miszal et al., 2010, 2013). To accommodate the large number of genotypes, the algorithm for proven and young animals (APY) was applied (Miszal et al., 2014a). The APY algorithm generates $G^{-1}$ using genomic recursion based on a subset of animals (“proven” or “core” animals). Only a relationship matrix for animals defined as “core” needs to be inverted; elements of $G^{-1}$ for all other animals (“young” or “non-core”) are calculated linearly by recursion, thus reducing computational requirements. Computational details of APY are described in Fragomeni et al. (2015) and Masuda et al. (2016). The program “cblup90iod2” (version 3.21; Miszal et al., 2014b) was used to obtain genomic breeding values by iteration on data using preconditioned conjugate gradient. The core consisted of 25,000 randomly selected animals. The core size was determined through
eigenvalue decomposition of the genomic relationship matrix and the numbers of largest eigenvalues explaining 99% of the variation were selected (Pocrnic et al., 2016) as implemented in the program “preGSF90” (version 1.10; Aguilar et al., 2014; Daniela Lourenco, Department of Animal and Dairy Science, University of Georgia, Athens, personal communication, 2017). The parameters determining the relative contribution of the genomic variance versus the residual polygenic variance (also known as α and β parameters) were set to 0.95 and 0.05, respectively. The preconditioned conjugate gradient method was run until the convergence criteria reached the threshold of $1 \times 10^{-12}$. Inbreeding was considered when constructing a pedigree relationship matrix with a depth of 20 generations (or the maximum number of generations available for each animal). The reliabilities of EBV were obtained with the program “accf90GS” (version 2.42), which approximates reliabilities using contributions from genotypes, phenotypes, and pedigree (Tsuruta et al., 2016). To reduce computational requirements, the contribution from genotypes is replaced by the value of the diagonal of the genomic relationship matrix pertinent to animal $i$, $g_{ii}$. The following formula was used to approximate reliabilities for genotyped animals (Daniela Lourenco, personal communication, 2016):

$$d^g_i = \alpha \left( \frac{\bar{REL}}{g_{ii}} + \frac{REL - Rel_{PA}}{1} \right),$$

where $d^g_i$ is the reliability of (genomic) EBV for animal $i$, $\alpha$ is the ratio of error variance to additive genetic variance, $g_{ii}$ is the diagonal element of the genomic relationship matrix pertinent to animal $i$, $\bar{REL}$ is the average base change in genomic reliability, and $Rel - Rel_{PA}$ is the contribution from phenotypes of genotyped animals. Thus, reliabilities were adjusted for the number of genotyped animals having phenotypes. All calculations were performed on a computer running Linux (x86_64) RedHat release 7.3 with Intel Xeon E5–4620 CPU (2.6 GHz) processors with 1 TB memory and 64 computing cores.

**Expression of EBV**

For each trait, the solutions from the “chbulp90iod2” program (raw EBV) were transformed into probabilities of exceeding the threshold value. The threshold represents the transition value between the 2 stages of the categorical variable (“healthy” and “sick”). Threshold values for all traits were calculated from the current data. For each animal solution, we calculated the probability that a standard normal variable with the mean equal to this solution and the variance of 1 exceeds the threshold. The probabilities were then multiplied by 100 (to represent percent) and divided by 2 (to obtain PTA). Animals with larger PTA values have higher relative disease risk compared with herdmates with smaller PTA values. Similar to Zoetis dairy wellness traits (Vukasinovic et al., 2017), the resulting PTA values were expressed as the deviation of the average PTA of all animals born in 2010 with a phenotypic record for that trait.

---

**Table 2.** Estimated variance components for calf wellness traits, heritabilities, and their standard errors, from the ASReml analysis with binomial distribution and a LOGIT link function.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Var(g)</th>
<th>Var(hy)</th>
<th>Var(e)</th>
<th>$h^2$</th>
<th>SE ($h^2$)</th>
<th>Var(g) adj</th>
<th>Var(hy) adj</th>
<th>Var(e) adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>RESP</td>
<td>0.251</td>
<td>2.436</td>
<td>3.29</td>
<td>0.042</td>
<td>0.002</td>
<td>0.076</td>
<td>0.740</td>
<td>1.00</td>
</tr>
<tr>
<td>DIAR</td>
<td>0.305</td>
<td>3.133</td>
<td>3.29</td>
<td>0.045</td>
<td>0.002</td>
<td>0.093</td>
<td>0.952</td>
<td>1.00</td>
</tr>
<tr>
<td>DEAD</td>
<td>0.308</td>
<td>1.540</td>
<td>3.29</td>
<td>0.060</td>
<td>0.003</td>
<td>0.094</td>
<td>0.468</td>
<td>1.00</td>
</tr>
</tbody>
</table>

$1^{st}$ Var(g) = additive genetic variance; Var(hy) = herd-year variance; Var(e) = residual variance; SE ($h^2$) = SE of the heritability; Var(g) adj, Var(hy) adj, Var(e) adj = variance components adjusted to Var(e) = 1 to fit the requirements of the threshold model in BLUPF90 programs. Adjusted values show the actual values used in BLUPF90 programs for genetic evaluation.

$2^{nd}$ RESP = calf respiratory disease; DIAR = calf scours; DEAD = calf livability.

---

**Table 3.** Mean, SD, minima, and maxima of estimated genomic PTA for calf wellness traits for genotyped animals with values of the diagonal of the genomic relationship matrix ($G_{diag}$) within ±3 SD from the mean ($n = 321,532$).

<table>
<thead>
<tr>
<th>Trait</th>
<th>Mean</th>
<th>SD</th>
<th>Minima</th>
<th>Maxima</th>
</tr>
</thead>
<tbody>
<tr>
<td>RESP</td>
<td>2.715</td>
<td>3.398</td>
<td>−8.00</td>
<td>24.007</td>
</tr>
<tr>
<td>DIAR</td>
<td>1.229</td>
<td>3.720</td>
<td>−11.510</td>
<td>28.515</td>
</tr>
<tr>
<td>DEAD</td>
<td>0.499</td>
<td>2.226</td>
<td>−6.473</td>
<td>22.797</td>
</tr>
</tbody>
</table>

$1^{st}$ RESP = calf respiratory disease; DIAR = calf scours; DEAD = calf livability.

---

**Table 4.** Mean, SD, minima, and maxima of genomic reliabilities for calf wellness traits for genotyped animals with values of the diagonal of the genomic relationship matrix ($G_{diag}$) within ±3 SD from the mean ($n = 321,532$).

<table>
<thead>
<tr>
<th>Trait</th>
<th>Mean</th>
<th>SD</th>
<th>Minima</th>
<th>Maxima</th>
</tr>
</thead>
<tbody>
<tr>
<td>RESP</td>
<td>0.419</td>
<td>0.061</td>
<td>0.189</td>
<td>0.99</td>
</tr>
<tr>
<td>DIAR</td>
<td>0.426</td>
<td>0.061</td>
<td>0.203</td>
<td>0.99</td>
</tr>
<tr>
<td>DEAD</td>
<td>0.473</td>
<td>0.056</td>
<td>0.246</td>
<td>0.99</td>
</tr>
</tbody>
</table>

$1^{st}$ RESP = calf respiratory disease; DIAR = calf scours; DEAD = calf livability.
Correlations of Calf Wellness Genomic PTA with Those for Other Economically Relevant Traits

To check associations among the calf wellness traits and with the genomic (g)PTA for other traits in the Zoetis and the national genetic evaluations, genetic correlations were approximated using method by Calo et al. (1973), as described in Parker Gaddis et al. (2014):

\[
\hat{r}_{12} = \sqrt{\left(\sum_{i=1}^{n} REL_{1i}\right)\left(\sum_{i=1}^{n} REL_{2i}\right)} \times \hat{r}_{12},
\]

where \(\hat{r}_{12}\) is the approximated genetic correlation between traits 1 and 2, \(REL_{1i}\) and \(REL_{2i}\) are reliabilities of (g)PTA for trait 1 and 2, respectively, for animal \(i\), and \(\hat{r}_{12}\) is the Pearson (product-moment) correlation between (g)PTA for traits 1 and 2. The standard error of the approximate genetic correlation was calculated following Sokal and Rohlf (1995), as described in Parker Gaddis et al. (2014). Only (g)PTA of animals having values of the diagonal of the genomic relationship matrix within mean ± 3 standard deviations (SD) were used in this calculation.

RESULTS AND DISCUSSION

Number of Records and Incidence of Calf Health Disorders

Table 1 shows the number of records used in genetic evaluation and the incidence (percent of affected animals in the data) for each trait. The numbers of records varied depending on the trait. The variation was caused by differences in herds recording various traits and the implementation of the “opportunity period” for each trait. Although most herds recorded death events, not all herds recorded respiratory events and scour with the same precision.

The incidence of RESP for the entire period of 0 to 365 d of age was 21%. The highest incidence of all
traits in this analysis was found for DIAR, at 26.1%. The incidence of deaths between 2 and 365 d of age was 4.7%. For all traits, the incidence found in our data was lower than that reported in the literature, which might be associated with the fact that the producers who provided the data were above average in herd management and disease control.

### Variance Components

Table 2 shows estimates for genetic, herd × year, and residual variance, heritabilities, and their standard errors, and the adjusted values of variance components used in BLUPF90 programs (cblup90iod2 and accf90GS).

The heritabilities for the calf wellness traits, estimated on the liability scale, were 0.042, 0.045, and 0.060 for RESP, DIAR, and DEAD, respectively. The heritability estimates for RESP and DIAR were comparable with values published in the literature (e.g., Henderson et al., 2011; McCorquodale et al., 2013).

The estimated heritability of 0.060 for DEAD was higher than the heritability of the comparable trait “heifer livability” (defined as a survival of an animal between 5 d and 18 mo of age), which was estimated by the Council on Dairy Cattle Breeding (CDCB, Bowie, MD) at 0.004 (Norman, 2016). The reason for higher heritabilities in our analyses relative to CDCB results could not be associated with a single factor, considering that the data used in this study may differ in geographical distribution, data size, trait definition, and model used.

### Genomic Evaluation

The number of rounds needed for convergence varied by trait, ranging from 16 for DEAD to 43 for DIAR. It is in the interest of computation time to reach the convergence in the fewest possible rounds; however, fewer than 10 iterations is not recommended because of the poorer estimation of the threshold (Shogo Tsuruta, Department of Animal and Dairy Science, University of Georgia, Athens, personal communication). It appears that the number of iteration rounds was inversely proportional to the number of records (fewer records = more iteration rounds) but also the incidence of the disorder (lower incidence needs more iterations). We expect the convergence for RESP and DIAR to improve in the future because the number of phenotypic records will increase.

### gPTA and Reliabilities

The results regarding gPTA and reliabilities are shown only for animals with values of the diagonal of the genomic relationship matrix (G_diag) within ±3 SD from the mean (1.006 ± 0.12). Because the G matrix was scaled using average allele frequencies (VanRaden, 2008), animals having genotypes very different from the overall population’s genomic profile (because they originated from a different population or were highly inbred) had extreme values of G_diag (Simeone et al., 2012). As a result of applying the approximation formula, these animals had extremely low reliabilities. From the initial 325,025 animals with genotypes, 3,493 had values of G_diag beyond the mean ± 3 SD.

Tables 3 and 4 show means, standard deviations, minima, and maxima for gPTA and reliabilities, respectively, for calf wellness traits for all genotyped animals in the analysis with values of G_diag within the mean ± 3 SD. The average values, SD, and extreme values of gPTA for calf wellness traits were comparable to those for wellness traits in dairy cows (Vukasinovic et al., 2017). The distribution of gPTA for RESP, DIAR, and DEAD for all genotyped animals is given in Figure 4. The average reliabilities of calf wellness traits were
lower than those reported for Zoetis dairy wellness traits, reflecting their lower estimated heritabilities and the smaller amount of information currently available. The average reliabilities were 42% for RESP, 43% for DIAR, and 47% for DEAD. All traits showed consistent SD of about 6 percentage points, with extremes reaching 19 and 99%. Figure 5 shows the distribution of reliabilities for RESP, DIAR, and DEAD. Most animals had reliabilities between about 40 and 45%, with only 175 animals having reliabilities of 90% or higher. All of the high-reliability animals were genotyped bulls with more than 300 phenotyped daughters each.

Tables 5 and 6 show means, SD, minima, and maxima for gPTA and reliabilities, respectively, for calf wellness traits for young genotyped Holstein animals without phenotype or progeny. The means, SD, and extremes showed very similar values as the entire set of genotyped animals, simply because young genotyped animals made up the majority of the genotyped animals for all traits. The average reliabilities of gPTA for calf wellness traits for young genotyped animals exceeded 40% for all traits, with the maximum reaching 62% for well-connected animals; that is, fully pedigreed animals having many ancestors and relatives with phenotypes and genotypes.

Tables 7 and 8 show means, SD, minima, and maxima for gPTA and reliabilities, respectively, for calf wellness traits for genotyped bulls having at least 100 phenotyped progeny.

### Genetic Correlations

**Genetic Correlations Among Calf Wellness Traits.** Table 9 shows correlations (product-moment correlations and approximated genetic correlations) among calf wellness traits. The approximated genetic correlations among calf wellness traits were all positive in sign and moderate in magnitude, ranging from 0.46 between DIAR and DEAD to 0.70 between RESP and DEAD. This indicated that animals that suffer from one disease are more likely to get another disease or die. A lower genetic correlation (0.29) was described in German Holstein in animals up to 2 mo of age between DIAR and RESP (Mahmoud et al., 2017); DIAR is strongly associated with RESP because of the impact that both have on the weight and subsequent recovery of the animal under treatment (Lundborg et al., 2003; McCorquodale et al., 2013).

**Genetic Correlations with Zoetis Dairy Wellness Traits.** Table 10 shows approximated genetic correlations (SE in parentheses) of calf wellness traits with Zoetis (Kalamazoo, MI) dairy wellness traits. The approximated genetic correlations between calf wellness traits and Zoetis dairy wellness traits were all positive in sign and moderate in magnitude, ranging from 0.46 between DIAR and DEAD to 0.70 between RESP and DEAD. Most correlations with DIAR and DEAD were positive, which was expected. The correlations with RESP were mostly negative, but low in magnitude.
magnitude. Positive correlations with DIAR suggested that animals susceptible to calfhood diseases may show increased incidence of disorders later in production. The highest positive correlations were observed between DEAD and MAST and between DIAR and KETO ($r \sim 0.24$). The correlations with mortality traits were estimated solely through genetic relationships, because animals with positive mortality phenotypes did not live long enough to have dairy wellness trait phenotypes.

Similar low correlations were found between calf diseases and disease occurrences of first-lactation German Holstein cows (Mahmoud et al., 2017). In agreement with that report, our results imply that calf diseases are poor predictors of cows’ health. Therefore, it is valuable to have both calf and cow health trait genomic predictions for selection.

**Genetic Correlations with Traits in the US National Genetic Evaluation.** Table 11 shows approximated genetic correlations among calf wellness traits and the traits in the US national genetic evaluation provided by CDCB. The trait DIAR was negatively associated with net merit (NM$\$), milk and protein yield, productive life, daughter pregnancy rate (DPR), cow conception rate (CCR), type (PTAT), and livability, and positively with SCS and daughter calving ease (DCE) and daughter stillbirth (DSB). These correlations make sense biologically, indicating that animals suffering from scours early in life end up being less profitable and less fertile, having more calving difficulties, a shorter productive life, and a higher risk of dying. Similarly, DEAD was negatively correlated with NM$\$, DPR, productive life (PL), CCR, heifer conception rate (HCR), and livability, and positively correlated with SCS and DCE and DSB; however, correlations with yield traits and type were close to zero. Multiple studies have found that animals suffering from calfhood diseases end up being less productive and having shorter herd life even if they recover (e.g., Rossini, 2004; Dunn et al., 2018). Surprisingly, correlations of RESP with most CDCB traits were in the opposite directions. Similar correlations between bovine respiratory disease (BRD) in calves and reproduction traits were found by Henderson et al. (2011) in a population of Holstein calves from New York. In a review study targeting influence of BRD on animals’ survival and performance in the Netherlands, Van Der Fels-Klerx et al. (2002) concluded that “although several associations were found, the data on the majority of the potential productivity effects and risk factors of BRD are ambiguous or incomplete.” The favorable correlations in this study between RESP and the traits in the national genetic evaluation may be simply linked to the nature of our data, suggesting that herds with higher production levels had more complete recording of RESP events.

<table>
<thead>
<tr>
<th>Trait</th>
<th>$\text{RESP}$</th>
<th>$\text{DIAR}$</th>
<th>$\text{DEAD}$</th>
<th>$\text{Milk}$</th>
<th>$\text{Prot}$</th>
<th>$\text{SCE}$</th>
<th>$\text{SCE}$</th>
<th>$\text{SCE}$</th>
<th>$\text{SCE}$</th>
<th>$\text{SCE}$</th>
<th>$\text{SCE}$</th>
<th>$\text{SCE}$</th>
<th>$\text{SCE}$</th>
<th>$\text{SCE}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{RESP}$</td>
<td>0.110</td>
<td>-0.427</td>
<td>-0.190</td>
<td>-0.180</td>
<td>-0.209</td>
<td>-0.247</td>
<td>-0.247</td>
<td>-0.247</td>
<td>-0.247</td>
<td>-0.247</td>
<td>-0.247</td>
<td>-0.247</td>
<td>-0.247</td>
<td>-0.247</td>
</tr>
<tr>
<td>$\text{DIAR}$</td>
<td>-0.049</td>
<td>0.027</td>
<td>0.007</td>
<td>-0.055</td>
<td>0.025</td>
<td>0.007</td>
<td>0.007</td>
<td>0.007</td>
<td>0.007</td>
<td>0.007</td>
<td>0.007</td>
<td>0.007</td>
<td>0.007</td>
<td>0.007</td>
</tr>
<tr>
<td>$\text{DEAD}$</td>
<td>0.048</td>
<td>-0.236</td>
<td>-0.387</td>
<td>-0.328</td>
<td>-0.378</td>
<td>-0.236</td>
<td>-0.236</td>
<td>-0.236</td>
<td>-0.236</td>
<td>-0.236</td>
<td>-0.236</td>
<td>-0.236</td>
<td>-0.236</td>
<td>-0.236</td>
</tr>
</tbody>
</table>
CONCLUSIONS

Genomic evaluation for calf wellness traits based on producer-recorded data was developed and successfully run with 325,025 genotypes. The estimated heritabilities of calf wellness traits were comparable with those reported in the literature. Large numbers of records result in high reliabilities even for traits with low heritabilities, showing that selection of animals for calf wellness traits can be feasible and potentially contribute to the genetic progress. A validation study may be conducted in the future to evaluate the accuracy and bias of ssGBLUP. Genetic correlations among calf wellness traits were all positive, likely implying that they are controlled by the overall immune system of the calf. Genetic correlations with Zoetis dairy wellness traits were low but mostly positive, indicating that animals suffering from early disorders may show poorer health later in life. Genetic correlations between calf wellness and routinely evaluated traits provided by the Council on Dairy Cattle Breeding were relatively low and in the expected direction for most traits. Having calf health disorders, especially scours, early in life negatively affects future profitability of animals, which is another reason to invest into genetic improvement of calf health.

REFERENCES


