ABSTRACT

The main objective of this study was to investigate the benefit of accuracy of genomic prediction when combining records for an intermediate physiological phenotype in a training population with records for a traditional phenotype. Fertility was used as a case study, where commencement of luteal activity (C-LA) was the physiological phenotype, whereas the interval from calving to first service and calving interval were the traditional phenotypes. The potential accuracy of across-country genomic prediction and optimal recording strategies of C-LA were also investigated in terms of the number of farms and number of repeated records for C-LA. Predicted accuracy was obtained by estimating population parameters for the traits in a data set of 3,136 Holstein Friesian cows with 8,080 lactations and using a deterministic prediction equation. The effect of genetic correlation, heritability, and reliability of C-LA on the accuracy of genomic prediction were investigated. When the existing training population was 10,000 bulls with reliable estimated breeding value for the traditional trait, predicted accuracy for the physiological trait increased from 0.22 to 0.57 when 15,000 cows with C-LA records were added to the bull training population; but, when the interest was in predicting the traditional trait, we found no benefit from the additional recording. When the genetic correlation was higher between the physiological and traditional traits (0.7 instead of 0.3), accuracy increased less when adding the 15,000 cows with C-LA (from 0.51 to 0.63). In across-country predictions, we observed little to no increase in accuracy of the intermediate physiological phenotype when the training population from Sweden was large, but when accuracy increased the training population was small (200 cows), from 0.19 to 0.31 when 15,000 cows were added from the Netherlands (genetic correlation of 0.5 between countries), and from 0.19 to 0.48 for genetic correlation of 0.9. The predicted accuracy initially increased substantially when recording on the same farm was extended and multiple C-LA records per cow were used in prediction compared with single records; that is, accuracy increased from 0.33 with single records to 0.38 with multiple records (on average 1.6 records per cow) from 2 yr of recording C-LA. But, when the number C-LA per cow increased beyond 2 yr of recording, we noted no substantial benefit in accuracy from multiple records. For example, for 5 yr of recording (on average 2.5 records per cow), accuracy was 0.47; on doubling the recording period to 10 yr (on average 3.1 records per cow), accuracy increased by 0.07 units, whereas when C-LA was recorded for 15 yr (on average 3.3 records per cow) accuracy increased only by 0.05 units. Therefore, for genomic prediction using expensive equipment to record traits for training populations, it is important to optimize the recording strategy. The focus should be on recording more cows rather than continuous recording on the same cows. Key words: multitrait genomic prediction, milk progesterone, dairy cattle, accuracy

INTRODUCTION

Although milk progesterone (P4) levels have been widely accepted as valid indicators of fertility in dairy cows, their application in routine genetic evaluation schemes has been constrained by the high cost associated with collecting sufficient samples per cow. Endocrine fertility traits derived from milk P4 levels have been recommended as indicators for fertility in dairy cows because they are physiological phenotypes that more directly reflect parts of a cow’s reproductive physiology (Bulman and Lamming, 1978; Lamming and Darwash, 1998; Darwash et al., 1999). For example, the interval from calving to first ovulation as determined by P4 levels in milk could be used as a direct indicator of a cow’s ability to return to luteal activity after calving, instead of an indirect indicator such as the interval from calving to first service (CFS). Several studies have revealed that endocrine fertility traits yield higher...
heritability estimates than traditional fertility traits derived from calving and insemination data. Notably, for the interval from calving to commencement of luteal activity (C-LA), heritability estimates of 0.15 to 0.28 have been reported (Darwash et al., 1997; Veerkamp et al., 1998; Petersson et al., 2007; Tenghe et al., 2015). Furthermore, examining the phenotypic correlation of C-LA with traditional fertility traits showed that early reestablishment of cyclic activity in postpartum cows increases the probability of an early insemination after calving, shortens the interval from calving to conception, increases conception rate, and reduces the number of services per conception (Darwash et al., 1997). These observations further suggest that early reestablishment of cyclic activity is an important prerequisite for good fertility.

Until recently, methods to measure milk P4 levels were labor intensive. They often entailed manually taking several milk samples per cow, analyzing, and recording the results. However, inline technology such as the Herd Navigator (DeLaval Int., Tumba, Sweden) now exists to measure P4 levels on individual cows during each milking, making it possible to sample a larger number of animals at a reasonable cost. But because few herds currently have the inline recording system, endocrine fertility traits that are defined from milk P4 levels would be most useful in a genomic selection scheme, where cows from contract herds form the reference population. In comparison, traditional fertility traits are routinely measured on a large scale, are more practical phenotypes for fertility, and are still useful target traits in national genetic evaluations. Therefore, combining (genomic predictions for) endocrine and traditional fertility traits may provide a more accurate prediction of fertility. In a previous study, using real data, the improvement in accuracy achievable from using endocrine fertility traits concurrent with a traditional traits (CFS) in genomic prediction of fertility was quantified (Tenghe et al., 2016a). However, the extent of exploring the effect of number of cows with endocrine traits in the training population in that study was limited by the data size. In the current study, the limitation was overcome by using a deterministic prediction equation to evaluate accuracy of prediction.

The main objective of our study was to investigate the benefit on accuracy of genomic prediction when combining records for an intermediate physiological phenotype in a training population with records for a traditional phenotype. Fertility was used as a case study, where the endocrine trait C-LA was a physiological or biological phenotype, whereas CFS and calving interval (CInt) were traditional phenotypes currently used in selection for fertility breeding values. The main objective was achieved by (1) evaluating the potential accuracy of different scenarios when combining a cow training population measured for the endocrine fertility trait C-LA with a training population of bulls with daughter observations for traditional fertility traits in multitrait genomic prediction of fertility; (2) evaluating the potential accuracy of across-country genomic prediction; (3) investigating recording strategies that optimally use the Herd Navigator for genomic prediction in terms of the number of farms and number of lactations per cow. For our objectives, predicted accuracy was obtained by estimating population parameters for the traits in a data set of Holstein Friesian cows and using a deterministic equation (Wientjes et al., 2016) that uses actual population parameters to predict the accuracy of genomic breeding values when different populations are combined in 1 training population. The effect of different factors such as genetic correlation between populations, heritability, and reliability of the physiological phenotype on the accuracy of genomic prediction were investigated.

**MATERIALS AND METHODS**

To obtain potential accuracy of predicting genomic breeding values using the deterministic equation, 3 steps were involved. In the first step, the available phenotypes and genotypes were used to estimate population parameters. The phenotypes were used to estimate genetic parameters (heritability and genetic correlations) in a mixed model, whereas the genotypes were used to estimate the effective number of chromosome segments (Me). In the second step, the population parameters were used in the deterministic equation to predict potential accuracy for different scenarios of genomic selection. In the last step, recording strategies for endocrine fertility traits were investigated.

**Phenotypes**

The phenotype data consisted of inline P4 records from 15 commercial herds in the Netherlands. In total, phenotypic data were available for 8,080 lactations on 3,136 Holstein Friesian cows. Milk sampling, measuring and recording of P4 level was performed with the Herd Navigator (DeLaval Int.) on all farms. Sampling frequency for P4 measurement was based on a biological model (Frieggens et al., 2008), but on average was undertaken every 2 d. The endocrine fertility trait investigated was C-LA, which was defined as the number of days between day of calving and first day on which milk P4 level was elevated ≥5 ng/mL. The traditional fertility traits were CFS and CInt.
Genotypes

The genotypes were used to calculate the effective number of chromosome segments (Me) across predicted and training populations; Me is one of the parameters used in the deterministic equation to predict accuracy. Genotypes were available from a previous study (Tenghe et al., 2016b), where 1,907 cows from the Netherlands were genotyped with the GeneSeek Genomic Profiler Bovine HD (GeneSeek, Lincoln, NE), containing 76,883 SNP (80k), and an additional 6 cows were genotyped with a custom Illumina 6k array, whereas 1,946 cows from Ireland, Sweden, and the United Kingdom were genotyped with the Illumina BovineSNP50 v1 BeadChip (Illumina Inc., San Diego, CA), which contained 54,001 SNP (50k). The 50k and 80k SNP arrays had 25,815 SNP in common, the rest were imputed from one data set to the other and vice versa using BEAGLE 3.3.2. (Browning and Browning, 2009), resulting in a total of 102,062 SNP. The quality criteria applied before imputation were minimum call rate of 95% per animal and 97% per SNP, and minimum GenCall score of 0.3. After imputation, SNP were retained for analysis if they fulfilled all quality criteria, namely (1) minor allele frequency >0.01 and (2) no extreme deviation from Hardy-Weinberg equilibrium (P < 0.01). Following all edits, the final data contained 85,485 SNP for 3,739 cows. Of the 3,739 cows with genotypes, 1,623 cows were used for edits, the final data contained 85,485 SNP for 3,739 cows. The details on how Me was calculated are presented below.

Estimating Variance Components

Variance components were estimated with mixed linear models that use the restricted maximum likelihood method. The model used was

\[ y_{ijkl} = \mu + p_j + h_{ys_k} + b_1 \times ca_1(p_j) + b_2 \times ca_2(p_j) + a_i + pe_i + e_{ijkl}, \]

where \( y_{ijkl} \) was the phenotype of the analyzed trait for individual \( i \); \( \mu \) was the overall mean; \( p_j \) was the fixed effect of parity \( j \) (cows in parity 4 or above were grouped to a common class); \( h_{ys_k} \) was the fixed effect of herd-year-season combination \( k \), with calving season defined as winter (December–February), spring (March–May), summer (June–August), autumn (September–November); \( b_1 \times ca_1(p_j) \) was the fixed regression on age at calving \( l \) (in months), nested within parity \( p_j \) with regression coefficient \( b_1 \); \( b_2 \times ca_2(p_j) \) was the fixed regression on age at calving \( l \), fitted as a quadratic covariate, nested within parity \( p_j \) with regression coefficient \( b_2 \); \( a_i \) was the random genetic effect of cow \( i \sim N(0, \sigma_a^2) \), where \( A \) was the additive genetic relationship matrix and \( \sigma_a^2 \) was the additive genetic variance; \( pe_i \) was the random permanent environmental effect of cow \( i \) to account for repeated measures within cow, \( \sim N(0, \sigma_{pe}^2) \), where \( I \) was an identity matrix including all animals (but animals without information on repeated records are set at zero) and \( \sigma_{pe}^2 \) was the permanent environment variance; and \( e_{ijkl} \) was the random error term, \( \sim N(0, \sigma_e^2) \), where \( \sigma_e^2 \) was the residual variance. ASReml 4.1 (Gilmour et al., 2014) was used for estimating the variance components.

Deterministic Prediction of Accuracy

We used the deterministic equation of Wientjes et al. (2016), which uses population parameters as input, to predict the accuracy of genomic prediction when different populations are combined into 1 training population. The populations might be populations from different lines or environments, or populations measured for different traits. The prediction equation is

\[ r_{GEBV} = \frac{r_{LD_{AC}}r_{G_{AC}}}{\sqrt{h_A^2 Me_{AC}}} \left( \frac{h_B^2}{Me_{BC}} + \frac{1}{N_A} \right) + \frac{r_{LD_{BC}}r_{G_{BC}}}{\sqrt{h_B^2 Me_{BC}}} \frac{h_B^2}{Me_{AC} Me_{BC}} \]

where \( r_{GEBV} \) is the accuracy of genomic estimated breeding values; \( A \) and \( B \) are the populations that are combined into one training population; \( C \) represents the animals in the predicted population for which only genotypes are available; \( Me \) is the effective number of chromosome segments across predicted and training populations; \( h^2 \) is the reliability of the breeding value used (i.e., the trait heritability for a single record on a cow), a reliability for progeny tested bulls (a value of 0.8 was assumed), and the reliability when cows have...
Three main scenarios were examined for accuracy of genomic estimated breeding values (GEBV). The first 2 scenarios (Figure 1) represent multitrait genomic prediction scenarios, where we considered prediction for 2 extreme scenarios; that is, the target trait was either a physiological or traditional fertility phenotype. In these scenarios, a training population consisting of records for a traditional fertility trait (population A) was combined with a cow training population (population B) with own records for an endocrine fertility trait. Population A was either a bull training set of 10,000 bulls with EBV based on daughter information (on average 60 daughters per bull) for a traditional trait, or a cow training set of 10,000 random cows with own phenotypes for a traditional trait. The combined training set (population A and B) was then used to predict GEBV for animals from population C. The animals from population C could be young genotyped bulls or cows with genotypes only and no phenotypes. In scenarios 1 and 2, the prediction in population C is for GEBV of a physiological and traditional fertility phenotype, respectively. For these scenarios, the parameters affecting predicted accuracy that were investigated included heritability, genetic correlation between the endocrine and traditional traits, and number of cows with endocrine traits added in training population B.

Scenario 3 represents multipopulation prediction where countries collaborate to setup a training population. For example, the training set of cows with endocrine fertility traits in Sweden (population A) is increased by adding a training set of cows with endocrine traits from the Netherlands (population B) to predict GEBV for animals in Sweden (population C). The effect of heritability, genetic correlation between countries, and number of cows in training population A and training population B on predicted accuracy were investigated in scenario 3.

Population Parameters

For scenarios 1 and 2, the population parameters for the deterministic equation were computed using the data from the Netherlands. The heritability and genetic correlation estimates were computed using the variance estimates from the mixed model described above. The $M_e$ was computed as in Goddard et al., (2011), using the formula $M_e = \frac{1}{\text{Var}[G_{ij} - E(G_{ij})]}$, where $G_{ij}$ is the genomic relationship matrix and $E(G_{ij})$ is the expected value for the genomic relationships between all individual $i$ and $j$, with variance taken over all pairwise relationships between individuals $i$ and $j$. For scenarios where countries collaborate to setup a training population, $M_e$ was computed as in Wientjes et al. (2016), using the genotypes from both countries, with the formula

$$M_{e A,B} = \frac{1}{\text{Var}[G_{\text{pop. } A, \text{pop. } B_j} - E\left(G_{\text{pop. } A, \text{pop. } B_j}\right)]},$$

where $G_{\text{pop. } A, \text{pop. } B_j}$ contains the genomic relationships, and $E\left(G_{\text{pop. } A, \text{pop. } B_j}\right)$ contains the expected values for the genomic relationships between all individuals $i$ from population $A$ and individuals $j$ from population $B$, and the variance is taken over all pairwise relationships.

The heritability of C-LA was obtained from the mixed model described above using the data from the Netherlands and was assumed to be the same for both countries. The $rLD$ in the prediction equation represents the square root of the proportion of genetic variance captured by SNP in the training population. This parameter can only be estimated based on empirical data by comparing the predicted and empirical accuracy. Because the aim of our study was to evaluate predicted accuracy only, $rLD$ estimates from literature were used. The $rLD$ for Holstein has been estimated in the range of 0.76 to 0.85 (Daetwyler, 2009; Erbe et al., 2013; Wientjes et al., 2015); therefore, the value of $rLD$ was set to 0.8 in all scenarios. The input parameters used for all scenarios are in Table 1.

Recording Strategies for Endocrine Fertility Traits

To investigate different strategies that optimally use the investment in the Herd Navigator for genomic prediction, we evaluated the effect of the number of farms and number of cows on predicted accuracy for scenario 1. We also varied the number of lactations for endocrine traits, because simply varying the number of records for genomic prediction is not sufficient given that cows
can be recorded for several lactations on each farm. That is, the number of records in the training population can be increased by increasing the number of cows measured or increasing the number of lactations measured for each cow. We used a strategy with different numbers of farms (15, 30, and 50) equipped with the Herd Navigator and different lengths of recording periods, namely 1, 2, 3, 5, 10, and 15 yr of recording C-LA. Each farm was assumed to have 100 lactating cows at the start of recording, with a replacement and culling rate of 25% per year, where the oldest animals were replaced and each cow would have 1 calf per year. With multiple records per cow, the heritability estimate for C-LA in the prediction equation defined above was replaced with the reliability calculated as in Van Vleck., (1993) using the formula reliability = \( \frac{mh^2}{(m-1)t+1} \), where \( m \) = number of records per cow, \( h^2 \) = heritability of the trait, and \( t \) = repeatability of the trait. The \( h^2 \) (0.13) and \( t \) (0.26) of C-LA obtained from the mixed model described above were used. The number of cows with C-LA records for each recording period and corresponding reliability estimates were used in the prediction equation to predict accuracy of GEBV. The details of the computed reliabilities, number of cows, and records for each recording period are in Table 2.

**Figure 1.** Flowchart of multitrait scenarios investigated, where a cow training population (population B) with own records for an intermediate physiological phenotype (an endocrine fertility trait) is combined with training population A, consisting of records for a traditional fertility phenotype. The endocrine trait is commencement of luteal activity (C-LA), and calving interval (CInt) is the traditional trait. Population A is either a bull training set with EBV based on daughter information (on average 60 daughters per bull) for the traditional trait, or a cow training set where cows have own records for the traditional trait. The combined training set is then used to predict genomic breeding values for animals from population C. The target trait for prediction in population C is a physiological fertility phenotype in scenario 1, and a traditional fertility phenotype in scenario 2; Nbull in population A is the number of bulls with EBV; Ncows represents the number of cows with their own records. GEBV = genomic EBV.
RESULTS

Genetic Parameters

The estimated heritability and repeatability for each trait are in Table 3. Heritability estimates were 0.06 for CInt, 0.11 for CFS, and 0.13 for C-LA. Repeatability estimates were low to moderate, ranging from 0.09 to 0.26, with the highest value observed for C-LA. The estimated genetic correlations (±SE) were 0.58 (±0.12) between C-LA and CFS, and 0.31 (±0.21) between C-LA and CInt.

Multitrait Genomic Prediction to Improve GEBV of Physiological Fertility Traits

First we evaluated the accuracy of genomic prediction for fertility when the target was GEBV for the physiological traits. We investigated three scenarios: Multitrait prediction of genomic EBV for physiological phenotypes, Multitrait prediction of genomic EBV for traditional phenotypes, and Across-country prediction of genomic EBV for physiological phenotypes.

Table 1. Input parameters of scenarios investigated for accuracy of genomic prediction when combining training populations and considering an endocrine or classical breeding goal for fertility in dairy cattle

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trait training population ((A))</td>
<td>Traditional (CInt or CFS)</td>
<td>Traditional (CInt or CFS)</td>
<td>Endocrine (C-LA)</td>
</tr>
<tr>
<td>Trait training population ((B))</td>
<td>Endocrine (C-LA)</td>
<td>Endocrine (C-LA)</td>
<td>Endocrine (C-LA)</td>
</tr>
<tr>
<td>Trait predicted population ((C))</td>
<td>Endocrine (C-LA)</td>
<td>Traditional (CInt, CFS)</td>
<td>Endocrine (C-LA)</td>
</tr>
<tr>
<td>(N) population (A)</td>
<td>10,000 bulls or cows</td>
<td>10,000 bulls or cows</td>
<td>200–10,000 cows</td>
</tr>
<tr>
<td>(N) population (B)</td>
<td>0–15,000 cows</td>
<td>0–15,000 cows</td>
<td>0–15,000 cows</td>
</tr>
<tr>
<td>Range (h^2_A)</td>
<td>0.8 (bulls) or 0.05 (cows)</td>
<td>0.8 (bulls) or 0.05 (cows)</td>
<td>0.13</td>
</tr>
<tr>
<td>Range (h^2_B)</td>
<td>0.05–0.2</td>
<td>0.05–0.2</td>
<td>0.13</td>
</tr>
<tr>
<td>Range (r_{G_{A,B}})</td>
<td>0.1–0.7</td>
<td>0.1–0.7</td>
<td>0.5–0.9</td>
</tr>
<tr>
<td>Range (r_{G_{A,C}})</td>
<td>0.99</td>
<td>0.1–0.7</td>
<td>0.5–0.9</td>
</tr>
<tr>
<td>Range (r_{G_{B,C}})</td>
<td>0.99</td>
<td>0.99</td>
<td>0.8</td>
</tr>
<tr>
<td>(r_{LD_{A,C}})</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>(r_{LD_{B,C}})</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>(Me_{A,C})</td>
<td>1,566</td>
<td>1,566</td>
<td>416</td>
</tr>
<tr>
<td>(Me_{B,C})</td>
<td>1,566</td>
<td>1,566</td>
<td>2,683</td>
</tr>
</tbody>
</table>

1\(A\) and \(B\) are the different populations that are combined into 1 training population; \(C\) is the predicted population; \(N\) is the number of individuals in each training population; \(h^2\) is the heritability in each training population; \(r_G\) is the genetic correlation between the populations; \(r_{LD}\) is the genetic variance in the predicted population captured by SNP in the training population; \(Me\) is the effective number of chromosome segments across predicted and training populations.

2\(C-LA = \) commencement of luteal activity, CInt = calving interval, CFS = interval from calving to first service.

Table 2. Input parameters used to investigate optimal recording strategies for endocrine fertility traits for use in genomic prediction

<table>
<thead>
<tr>
<th>No. of farms</th>
<th>No. of recording years</th>
<th>No. of unique cows per farm</th>
<th>No. of records per farm</th>
<th>Total cows on all farms</th>
<th>Total records on all farms</th>
<th>No. of records per cow</th>
<th>Reliability C-LA</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>1</td>
<td>100</td>
<td>100</td>
<td>1,500</td>
<td>1,500</td>
<td>1.0</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>125</td>
<td>200</td>
<td>1,875</td>
<td>3,000</td>
<td>1.6</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>150</td>
<td>300</td>
<td>2,250</td>
<td>4,500</td>
<td>2.0</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>200</td>
<td>500</td>
<td>3,000</td>
<td>7,500</td>
<td>2.5</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>325</td>
<td>1,000</td>
<td>4,875</td>
<td>15,000</td>
<td>3.1</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>450</td>
<td>1,500</td>
<td>6,750</td>
<td>22,500</td>
<td>3.3</td>
<td>0.27</td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td>100</td>
<td>100</td>
<td>3,000</td>
<td>3,000</td>
<td>10</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>125</td>
<td>200</td>
<td>3,750</td>
<td>6,000</td>
<td>1.6</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>150</td>
<td>300</td>
<td>4,500</td>
<td>9,000</td>
<td>2.0</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>200</td>
<td>500</td>
<td>6,000</td>
<td>15,000</td>
<td>2.5</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>325</td>
<td>1,000</td>
<td>9,750</td>
<td>30,000</td>
<td>3.1</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>450</td>
<td>1,500</td>
<td>13,500</td>
<td>45,000</td>
<td>3.3</td>
<td>0.27</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>100</td>
<td>100</td>
<td>5,000</td>
<td>10,000</td>
<td>1.0</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>125</td>
<td>200</td>
<td>6,250</td>
<td>10,000</td>
<td>1.6</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>150</td>
<td>300</td>
<td>7,500</td>
<td>15,000</td>
<td>2.0</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>200</td>
<td>500</td>
<td>10,000</td>
<td>25,000</td>
<td>2.5</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>325</td>
<td>1,000</td>
<td>16,250</td>
<td>50,000</td>
<td>3.1</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>450</td>
<td>1,500</td>
<td>22,500</td>
<td>75,000</td>
<td>3.3</td>
<td>0.27</td>
</tr>
</tbody>
</table>

\(C-LA = \) commencement of luteal activity. Reliability computed as \(mh^2/(m − 1)t + 1\), \(m = \) number of records per cow, \(h^2 = \) heritability of the trait, and \(t = \) repeatability of the trait.
ological fertility phenotypes in scenario 1, where C-LA is the physiological trait; training population A was either 10,000 bulls with EBV or 10,000 cows with own records for a traditional trait (CInt or CFS), whereas training population B was cows with own records for C-LA. The training populations were combined in a multitrait model and used to predict C-LA for animals from population C. Predicted accuracy as a function of number of cows from population B added to training population A are shown in Figure 2, assuming a heritability of 0.05 to 0.2 for C-LA and genetic correlations of 0.1 to 0.7 between C-LA and the traditional trait. When training population A consists of bull EBV (Figure 2a) and heritability of C-LA was 0.1 with a genetic correlation of 0.3, accuracy increased from 0.22 when only population A was used, to 0.57 when 15,000 cows were added from population B; with genetic correlation of 0.7, accuracy increased from 0.51 to 0.63. With a higher heritability for C-LA of 0.2, the accuracy increased more (from 0.22 to 0.65) when 15,000 cows were added from population B for genetic correlation of 0.3, and from 0.51 to 0.68 with genetic correlation of 0.7. When training population A consisted of cows with own records of CInt or CFS instead of bull EBV (Figure 2b), and the heritability of C-LA was 0.1 with a genetic correlation of 0.3, the accuracy increased from 0.12 when only population A was used to 0.56 when 15,000 cows are added from population B; with a genetic correlation of 0.7, the accuracy increased from 0.28 to 0.57. A similar trend of increased accuracy was observed with heritability of 0.2 for C-LA (Figure 2b). In general, the accuracy was higher when the genetic correlation between C-LA and the traditional trait was higher, and the accuracy increased when an increasing number of cows with C-LA records was added to the training population, especially when the genetic correlation was small.

Multitrait Genomic Prediction to Improve GEBV of Traditional Fertility Traits

In the second scenario, accuracy of prediction was examined when selection for fertility was based on GEBV for traditional phenotypes (CInt or CFS). Training population A consisted of 10,000 bulls with EBV or 10,000 cows with own records for CInt or CFS, and training population B was cows with own records for C-LA. The training populations were combined and used to predict CFS or CInt for animals from population C. Predicted accuracy as a function of number of cows added to training population B is shown in Figure 3. When training population A consisted of bull EBV, adding information on cows for C-LA did not increase accuracy, irrespective of the genetic correlation between traits, or heritability of C-LA (Figure 3a). However, when the 10,000 animals in training population A were cows with own records for the traditional trait, some improvement was noted in accuracy when adding cows with C-LA records from population B (Figure 3b). That is, when heritability of C-LA was 0.1 with a genetic correlation of 0.3, the accuracy increased from 0.39 when only population A was used to 0.41 when 15,000 cows were added from population B, and from 0.39 to 0.50 with a correlation of 0.7. For heritability of 0.2 with a genetic correlation of 0.3, the accuracy increased from 0.39 when only population A was used to 0.42 when adding population B, and from 0.39 to 0.53 with a genetic correlation of 0.7.

Across-Country Genomic Prediction

The across-country predictions show that when the training population with the physiological phenotype from Sweden is small, a substantial increase in accuracy can be obtained by adding animals from the Netherlands. For example, the predicted accuracy with 200 cows from Sweden was 0.19 and increased to 0.31 when 15,000 cows were added from the Netherlands, for a genetic correlation of 0.5 between countries; and for a genetic correlation of 0.9, the accuracy increased from 0.19 to 0.48 (Figure 4a). When the genetic correlation between countries was high, fewer animals from the Netherlands were needed to obtain substantial accuracy. For example, with 200 animals from Sweden and a correlation of 0.5, 15,000 animals from the Netherlands were needed to achieve an accuracy of 0.31, but when

<table>
<thead>
<tr>
<th>Trait</th>
<th>Mean</th>
<th>SD</th>
<th>Number of lactations</th>
<th>Number of cows</th>
<th>h²</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-LA, d</td>
<td>38.96</td>
<td>18.98</td>
<td>5,043</td>
<td>2,748</td>
<td>0.13 (0.03)</td>
<td>0.26 (0.02)</td>
</tr>
<tr>
<td>CInt, d</td>
<td>409.76</td>
<td>66.54</td>
<td>4,030</td>
<td>1,916</td>
<td>0.06 (0.03)</td>
<td>0.09 (0.02)</td>
</tr>
<tr>
<td>CFS, d</td>
<td>88.94</td>
<td>33.93</td>
<td>6,655</td>
<td>2,898</td>
<td>0.11 (0.02)</td>
<td>0.11 (0.02)</td>
</tr>
</tbody>
</table>

1C-LA = commencement of luteal activity, CInt = calving interval, CFS = interval from calving to first service records per cow, and t = repeatability of the trait.
Figure 2. Predicted accuracies for genomic (G)EBV of commencement of luteal activity (C-LA), which is a physiological phenotype, when adding different number of cows (population B) with own C-LA records to a training population (population A) consisting of 10,000 individuals (bulls or cows) with records for a traditional fertility trait for multitrait genomic prediction. The traditional traits considered were calving interval (CInt) and interval from calving to first service (CFS). When population A was a bull training population with EBV based on daughter information, heritability or reliability of the traditional trait was 0.8 (a); when population A was a cow training population with own records, heritability of the traditional trait was 0.05 (b). The heritability for the endocrine trait in population B varied from 0.05 to 0.2, and genetic correlation between the endocrine and classical trait \( r_{G,A,B} \) varied from 0.1 to 0.7. The effective number of chromosomes \( Me \) was 1,566, and the proportion of variance captured by SNP \( r_{LD} \) was 0.8. The accuracy was based on the formula of Wientjes et al. (2016).
Figure 2 (Continued). Predicted accuracies for genomic (G)EBV of commencement of luteal activity (C-LA), which is a physiological phenotype, when adding different number of cows (population B) with own C-LA records to a training population (population A) consisting of 10,000 individuals (bulls or cows) with records for a traditional fertility trait for multitrait genomic prediction. The traditional traits considered were calving interval (CInt) and interval from calving to first service (CFS). When population A was a bull training population with EBV based on daughter information, heritability or reliability of the traditional trait was 0.8 (a); when population A was a cow training population with own records, heritability of the traditional trait was 0.05 (b). The heritability for the endocrine trait in population B varied from 0.05 to 0.2, and genetic correlation between the endocrine and classical trait (rgAB) varied from 0.1 to 0.7. The effective number of chromosomes (Me) was 1,566, and the proportion of variance captured by SNP (rLD) was 0.8. The accuracy was based on the formula of Wientjes et al. (2016).
Figure 3. Predicted accuracies for genomic (G)EBV of calving interval (CInt) or interval from calving to first service (CFS), when adding different number of cows (population B) with own measures for the endocrine trait commencement of luteal activity (C-LA) to a training population (population A) consisting of 10,000 individuals (bulls or cows) with traditional fertility traits for multitrait genomic prediction. When population A was a bull training population with EBV based on daughter information, heritability or reliability of the traditional trait was 0.8 (a); when population A was a cow training population with own records, heritability of the traditional trait was 0.05 (b). The heritability for the endocrine trait in population B varied from 0.05 to 0.2, and the genetic correlation between the endocrine and traditional trait (rgAB) varied from 0.1 to 0.7. The effective number of chromosomes (Me) was 1,566, and the proportion of variance captured by SNP (rLD) was 0.8. The accuracy was based on the formula of Wientjes et al. (2016).
Figure 3 (Continued). Predicted accuracies for genomic (G)EBV of calving interval (CInt) or interval from calving to first service (CFS), when adding different number of cows (population B) with own measures for the endocrine trait commencement of luteal activity (C-LA) to a training population (population A) consisting of 10,000 individuals (bulls or cows) with traditional fertility traits for multitrait genomic prediction. When population A was a bull training population with EBV based on daughter information, heritability or reliability of the traditional trait was 0.8 (a); when population A was a cow training population with own records, heritability of the traditional trait was 0.05 (b). The heritability for the endocrine trait in population B varied from 0.05 to 0.2, and the genetic correlation between the endocrine and traditional trait (rgAB) varied from 0.1 to 0.7. The effective number of chromosomes (Me) was 1,566, and the proportion of variance captured by SNP (rLD) was 0.8. The accuracy was based on the formula of Wientjes et al. (2016).
the correlation was 0.9, only 3,000 animals from the Netherlands were needed to achieve the same accuracy (Figure 4a). Similarly, with 2,000 animals from Sweden, 15,000 animals from the Netherlands were needed to achieve an accuracy of 0.52 when the correlation was 0.5, but only 3,000 animals were needed to attain the same accuracy when the correlation was 0.9 (Figure 4c).

However, when the training population from Sweden was large, for example 10,000 animals, little to no increase in accuracy was observed when adding animals from the Netherlands. That is, for 10,000 animals from Sweden, the accuracy was 0.69 with and without adding animals from the Netherlands, for a genetic correlation of 0.5; for genetic correlation of 0.9, accuracy increased from 0.69 to 0.70 (Figure 4d).

**Recording Strategies for Endocrine Fertility Traits**

Figure 5 presents predicted accuracy of GEBV for C-LA as a function of recording strategies for C-LA in terms of the different number of farms on which cows...
were recorded and number of years during which C-LA was recorded. The recording strategies influence the number of cows recorded and the number of lactations for each cow added from training population B. The reliability of C-LA for the different recording strategies are in Table 2, they were 0.13 for 1 yr of recording, 0.18 (for 2 yr), 0.21 (for 3 yr), 0.23 (for 5 yr), 0.26 (for 10 yr), and 0.27 (for 15 yr). The number of lactations (records) per cow, number of cows, and total number of records for each recording strategy are in Table 2. The predicted accuracy increased substantially when multiple C-LA records were used in prediction compared with single records; the accuracy increased from 0.33 with single records to 0.38 with multiple records (on average 1.6 records per cow) for 2 yr of recording (Figure 5a). However, when C-LA was recorded for longer periods, no substantial benefit in accuracy from multiple periods was observed. For example, for 5 yr of recording (on average 2.5 records per cow), the accuracy was 0.47; on doubling the recording period to 10 yr (on average 3.1 records per cow), the accuracy increased by 0.07 units (Figure 5a), but when C-LA was recorded for 15 yr (on average 3.3 records per cow) it increased only by 0.05 units. The extra lactations for each cow after the second lactation did not add much benefit to the accuracy of prediction. The results also indicate that, with repeatability of 0.26 for C-LA, the number of cows with C-LA is more important than the number of records for genomic predictions. For example, 15,000 records obtained from 3 different recording strategies yielded different accuracies due to different number of cows in the training population. That is, for 15,000 records obtained from 15 farms on 4,875 cows in 10 yr the accuracy of predicting C-LA was 0.54, when 15,000 records were obtained from 30 farms on 6,000 cows in 5 yr accuracy was 0.56, and the accuracy was 0.57 for the 15,000 records obtained from 50 farms on 7,500 cows in 3 yr (Figure 5a). A similar trend was observed when a low estimate for repeatability (0.01) was assumed for C-LA (results not shown). Therefore, it seems more important to have more cows than more lactations per cow in the training population.

**Discussion**

We used fertility as a case study to evaluate the benefit on accuracy of genomic prediction when combining records of a physiological phenotype with records of a traditional phenotype to predict genomic breeding values. The results showed that when interest is in selection to improve traditional fertility phenotypes (e.g., CFS or CInt), and there is an existing large training population with bull EBV for these traditional phenotypes, no benefit in accuracy of prediction exists when adding physiological phenotypes (e.g., C-LA) from cows to the training population. However, when the target GEBV is based on physiological phenotypes, accuracy is substantially improved when the physiological phenotypes are added to the training population. Also investigated was the potential accuracy of across-country genomic prediction, where training populations from Sweden and the Netherlands were combined to predict animals in Sweden. The findings showed that, when the training population from Sweden was small (e.g., 200 animals), substantial increase in accuracy could be obtained by adding animals from the Netherlands; but when the training population from Sweden was large (e.g., 10,000 animals), little to no increase in accuracy was observed when adding animals from the Netherlands. We also investigated recording strategies that optimally use the Herd Navigator for genomic prediction in terms of the number of farms and repeated records for C-LA. The results showed that for genomic prediction of fertility using C-LA, it is more important to have more animals with C-LA records than more C-LA records per animal.

**Accuracy**

The general increase in accuracy from multitrait genomic prediction is in accordance with previous studies that have shown that when the genetic correlation between traits is high, multitrait prediction improves accuracy; this has been shown in stochastic (Calus and Veerkamp, 2011; Jia and Jannink, 2012) and deterministic (Calus et al., 2013) simulations as well as in real data (Calus et al., 2013; Cooper et al., 2015; Tenghe et al., 2016a). Using a subset of the data used in our study and cross validation, Tenghe et al., (2016a) showed that the accuracy of genomic prediction for fertility was improved substantially when endocrine and traditional fertility traits were used for prediction in a multitrait model. A substantial increase in accuracy was observed in scenario 1, where the target was to improve GEBV for C-LA, and cows with own C-LA records were added to a large training population of bulls with EBV for a traditional trait for prediction. In scenario 1, C-LA was not measured in the bull training population. As a result, information from the cows improved the accuracy of prediction for C-LA substantially. However, when interest was in selection for the traditional phenotype (scenario 2), adding 15,000 cows with C-LA records to the bull training population did not improve accuracy of GEBV because the bull training population was measured for the traditional phenotype. This confirms that it is important to measure the target trait in the training population to benefit from the information in
This was true for C-LA, but could equally apply to similar situations where a predictor trait is used to predict the target trait. For example, SCC can be used to predict breeding values for mastitis incidence (Philipsson et al., 1995) and birth weight can be used to predict calving ease or perinatal mortality (Johanson and Berger, 2003). However, it is beneficial to have actual measures of mastitis or perinatal mortal-

Figure 5. Predicted accuracies for commencement of luteal activity (C-LA), as a function of number of cows, and number of records for C-LA, when combining C-LA and a traditional fertility phenotype in multitrait genomic prediction. The traditional phenotype was calving interval (CInt). Training population A is combined with different number of cows (with different number of C-LA records per cow) in population B for genomic prediction. Heritability estimates were 0.13 for C-LA, and 0.06 for CInt. The genetic correlation was 0.31 between C-LA and CInt. The effective number of chromosomes ($Me$) was 1,566, and the proportion of variance captured by SNP ($rLD$) was 0.8. The accuracy was based on the formula of Wientjes et al. (2016).
ity incidence on a set of cows. That is, it is beneficial to have actual measurements on the target trait if a correlated trait is used to predict it.

To assist in the prediction of GEBV of practical fertility traits such as CInt, which are important to farmers, a physiological predictor phenotype such as C-LA should predict a significant proportion of the target trait and have a higher heritability. The estimated genetic correlation of C-LA with CInt was low, and little to no increase in accuracy of prediction was observed when CInt was the target trait (Figure 3). However, the expectation of physiological phenotypes is higher precision because they are closer to the biological function, as reflected in the higher heritability estimate of C-LA compared with CInt. The downside is that the physiological phenotypes explain less of the trait that is of importance to farmers. Hence, the use of physiological phenotypes will depend on the optimum of the heritability and genetic correlation with the target trait.

The accuracy of breeding values depends on the sources of information included in each phenotypic record. The accuracy of a single phenotypic measurement of an animal itself is equal to the square root of the heritability ($\sqrt{h^2}$) and the reliability is equal to $h^2$; however, this is not true when multiple measurements per animal are used. In practice, more than one measurement for C-LA will be available for cows with multiple lactations. Therefore, the predicted accuracy was also evaluated accounting for repeated records, by substituting the heritability of C-LA in the prediction equation with the reliability. When considering C-LA as the target trait, the predicted accuracy increased substantially when more records were available either due to measuring on more farms (i.e., more animals), or due to measuring for a longer period on the farms (i.e., repeated records). However, measuring C-LA on more farms gave more of an increase in accuracy than measuring for longer periods on a few farms. Little to no increase in accuracy was observed when considering CFS or CInt as target trait and accounting for repeat records, especially when training population A consisted of bulls with EBV for CFS or CInt.

The trend of increased accuracy from combining 2 populations from different countries in 1 training population was expected due to the increase in size of the training population. When the training population from Sweden was small, the accuracy was improved when adding animals from the Netherlands. Previous studies that investigated the accuracy of multipopulation genomic prediction by combining Holstein populations from different countries found a similar trend (Lund et al., 2011; Haile-Mariam et al., 2015; de Haas et al., 2015). A similar trend has also been observed in other studies involving different breeds [e.g., Jersey populations (Wiggans et al., 2015; Haile-Mariam et al., 2015) and Brown Swiss populations (Zumbach et al., 2010; Jorjani, 2012)]. Little to no improvement in accuracy was observed when the training population from Sweden was large (e.g., 10,000) because the accuracy obtained with 10,000 animals was already high.

**Implications for Practical Dairy Breeding**

In genomic selection, the requirement of collecting many phenotypes on progeny is minimized as opposed to conventional selection. Therefore, genomic selection provides a new opportunity to reconsider genetic improvement using new traits, such as endocrine fertility traits, and to start selecting for them. For genomic selection with new traits, the phenotypes need to be collected only for the animals in the training population; hence, collection of the phenotypes may be conducted on contract farms where collecting expensive or difficult to measure data is feasible. The training population for new traits will be smaller compared with the routinely recorded traits and may result in lower accuracy in genomic prediction. The expected limited size of the training population for new traits stresses the need for strategies that optimally record and use all available information.

One objective of our study was therefore to investigate recording strategies that optimally use the Herd Navigator for genomic prediction. This was done by evaluating the effect of number of farms and recording period for C-LA on accuracy when combining endocrine and traditional fertility traits for prediction. We showed that, in terms of accuracy, it is more beneficial to record more cows for C-LA than more C-LA records per cow. This is important, as it raises an important question of whether it is beneficial to invest in more contract farms and record C-LA for a shorter period or invest in fewer contract farms and record for longer period to attain a target training population size. However, this benefit of more cows over more records per cow is clearly dependent on the extent of repeatability of the trait and the cost of measurement. When the repeatability is lower than found here, and additional cows are more costly than repeated measures per cow, it might be economically optimal to record cows of all ages over multiple years for training rather than having to visit more herds. When the repeatability is higher, it will be more beneficial to record more cows than more lactations per cow. The optimal recording strategy will depend partly on the target trait and cost of measuring the trait. In our case, if we consider that each farm requires 1 Herd Navigator for collecting P4 records and
the target trait is C-LA, then it will be beneficial for the accuracy of estimating genomic breeding values, to invest in more contract farms. For example, there is more benefit investing in 50 contract farms to collect 15,000 records on 7,500 cows in 3 yr than investing in 15 farms to collect the 15,000 records on 4,875 cows in 10 yr. Such recording strategies could equally be applied to traits that are not routinely measured because they are difficult or expensive to measure (e.g., feed intake and methane emission), and for which obtaining large number of phenotypes for the training population in genomic prediction is problematic.

The challenge for dairy breeding companies is to start thinking of ways to optimize phenotyping strategies for these expensive novel traits. Some dairy companies have currently set up partner herds for recording specific phenotypes. A next step for dairy breeding companies would be to optimize the phenotyping structure by establishing phenotyping labs where animals are phenotyped in batches for 1 lactation and dispatched. This will make space for the next batch of cows and leads to phenotypes recorded on as many animals as possible. Based on our results, we argue that this is an optimal way of collecting as many phenotypes as possible on expensive novel traits to set up a training population for genomic prediction. Another advantage of having phenotyping labs is that animals to be measured can be selected in terms of their relatedness to important bulls (or offspring of the sires of the young bulls) so that there is always a close relationship between the cow training population and the candidate bulls, as this will increase the accuracy of genomic prediction. Therefore, it will be more interesting to phenotype heifers (i.e., first parity cows) for expensive traits than later-parity cows.

Investing in phenotyping labs with the Herd Navigator will optimize the recording strategy for endocrine fertility traits in dairy cattle, but this investment will not be worthwhile if the only traits recorded are fertility traits. On the other hand, if investments for recording other expensive traits (e.g., feed intake and methane emission) are included, this might become a realistic option for the future.

The substantial increase in accuracy due to increase in size of the training population by combining the Dutch and Swedish populations in 1 training population denotes a benefit in collaboration between countries, enabling an increased size of the training population for expensive or difficult to measure traits. This collaboration is especially beneficial when the genetic correlation between countries is high. In our study, when the genetic correlation was high, fewer number of animals were needed to attain the same level of accuracy as for when the correlation was lower.

CONCLUSIONS

Genomic selection and new phenotyping technologies, such as the Herd Navigator (DeLaval Intl.), for measuring P4 level in milk present an opportunity to include selection for intermediate physiological phenotypes such as endocrine fertility traits in breeding programs. This study showed that for genomic prediction of fertility, when the target trait is traditional phenotypes (e.g., CFS or CInt) and there is already an existing large training population with bull EBV for CFS or CInt, there is no benefit in accuracy on adding a physiological phenotype from cows to the training population for prediction. However, when the target trait is a physiological phenotype, such as C-LA, accuracy is substantially improved when the physiological phenotype is added. The study also showed that optimizing the recording strategy for C-LA can maximize the accuracy of prediction; that is, for genomic prediction of fertility using C-LA, focus should be on recording more cows rather than continue recording on the same cows. In general, for novel traits, optimal utilization of novel phenotyping technologies may require different phenotyping strategies compared with current practice.

ACKNOWLEDGMENTS

The authors acknowledge Lattec I/S (Hillerød, Denmark) for providing the in-line progesterone data, and the Cooperative Cattle Improvement Organization CRV BV (Arnhem, the Netherlands) for providing the pedigree data for this study. A. M. M. Tenghe benefited from a joint grant from the European Commission and the European Union under the grant agreement no. 311776, within the framework of Erasmus-Mundus joint doctorate “EGS-ABG” (Paris, France). This study was performed as part of the Prolific project, which is financially supported by the European Commission under the Seventh Research Framework Programme, Grant Agreement no. 311776.

REFERENCES


Cooper, T. A., G. R. Wiggins, and P. M. VanRaden. 2015. Short communication: Analysis of genomic predictor population for Holstein


