Mixed reference population in genomic evaluation for clinical mastitis in Czech Holstein cattle

Ludmila Zavadilová*, Eva Kašná, Zuzana Krupová, Luboš Vostrý

Výzkumný ústav živočišné výroby, v. v. i., Praha – Uhříněves, Česká republika

The accuracy of genomic breeding values (GEBV) for clinical mastitis (CM) of Czech Holstein cattle was analysed. The single-step genomic method and mixed genomic reference populations were employed. Dataset included 92,388 Holstein cows and 160,426 lactations. CM lactation incidence was 19.05%. Cows calved between 2017 and 2022 in 119 herds. A total of 4,969 Holstein sires and 35,814 Holstein females were genotyped. Three genomic matrices were used, two of them encompassing females. The linear animal model with repeatability included fixed effects of herd-year-season and parity-age at calving. The highest average accuracy of GEBV occurred for genotyped cows. SD of GEBV accuracy for bulls was significantly higher than these for cows. The accuracy of genotyped cows with health phenotypes ranged from 0.06 to 0.51. For the genotyped bulls born in 2021 and 2022, the maximal accuracy was 0.37; for genotyped heifers, the maximal accuracy was 0.42. The highest average GEBV accuracy occurred for the reference population with genotyped bulls and genotyped cows with phenotypes. The average accuracy for the young genotyped bulls increased by one percentage point when phenotyped cows were considered in the reference genomic population. The cow's GEBV accuracy benefited from including their genotypes in the prediction. It has been confirmed that due to the expansion of the genomic reference population to include a group of genotyped cows with phenotypes, the individual accuracy of GEBV for CM had increased.

Keywords: dairy cattle, udder health, single-step genomic evaluation, the mixed genomic reference population

1   Introduction

Genomic selection (Meuwissen et al., 2001; Schaeffer, 2006) has been implemented in dairy cattle breeding in many countries. Besides other things, the improved accuracy of breeding values, especially for low heritable traits and, subsequently, the high selection success, were the main reasons. Genomic selection accelerates genetic progress by shortening generation intervals while the accuracy of selection for the young animal is increasing (Schaeffer, 2006; Obšteter et al., 2019). The introduction of genomic selection has its hindrances, for example, the cost of genotyping. The accuracy of the genomic prediction depends on various parameters, including the size of the reference population and its genetic structure (Lee et al., 2017). As Meuwissen et al. (2001) stated, the first step is to genotype a sufficient number of animals with progeny records or phenotypes to create a genomic reference population. Particularly in small livestock populations, it could be a problem, as it is relatively costly to genotype, and a large number of genotyped animals is necessary for successful genomic selection.

Nevertheless, Obšteter et al. (2019) and Jenko et al. (2019) highlight the importance of the increasing number of genotyped females as those represent the genotyped animals with the records. The USA was the first to include females in its reference population (Wiggans et al., 2011). Thomasen et al. (2014, 2020) concluded that genotyping of cows is a quick method to increase accuracies of genomic predictions and upsurge genetic gain in a small population. To select animals as a parent of the next generation, the individual accuracy of genomic breeding value is necessary. For small populations, approximative accuracies can be calculated by the inversion of the left-hand side of a BLUP system of equations (VanRaden 2008). For the genomic evaluation by the single-step method in a large population, the approximative

*Corresponding Author: Ludmila Zavadilová, Výzkumný ústav živočišné výroby, v. v. i., ♀ Přátelství 815, 10400 Praha – Uhříněves, Česká republika

methods are developed by Misztal et al. (2013). Bauer et al. (2015) analysed those methods in the Czech Republic.

Mastitis is the most common disease in dairy cows. The clinical or subclinical form negatively impacts animal welfare and the economic efficiency of dairy farms. The costs associated with treatment, production losses, and reduced animal welfare are high (Jamali et al., 2018; Wolfová et al., 2006), including the increased risk of culling and shortening the production period of a dairy cow. The possibility exists to increase the resistance to mastitis in dairy cattle by selection despite clinical mastitis being a typical low heritable trait (Martin et al., 2018). In most cases, literature results indicate the heritability of mastitis between 0.01 and 0.10. Furthermore, in confirming this suggestion, Heringstadt et al. (2001) proved that the longstanding process of genetic selection to reduce the incidence of clinical mastitis led to a positive genetic trend in Norwegian dairy cattle from the 1970s to the 1990s.

Genomic evaluation for Czech Holstein cattle (Přibyl et al., 2012) started with the single-step genomic method (ss-GBLUP) for production traits. The bull reference population has been increasing gradually, with the oldest bulls born in the 1970s and originating mainly from North America and European countries. Still, most of the increase in the number of genotyped animals is due to the genotyping of females. The Czech Holstein breeder's organisation recently started a long-term project, FitCow, to raise the number of genotyped cows and increase the accuracy of genomic prediction. FitCow is in sync with the project: the national cattle health monitoring system "The Diary of Diseases and Medication" (Kasna et al., 2017). The Diary, as the web application, was made available to the farmers in August 2018 after a one-year trial. It consists of a farmers' online health recording form and a simplified key of diagnoses based on ICAR recommendations. Both projects' goal is to enable the selection for increasing disease resistance by including health traits in the selection index. The outputs of those projects are utilised in genomic evaluation for the health traits of Czech Holstein cattle.

The objectives of this study were to analyse the individual accuracy of genomic breeding values (GEBV) for clinical mastitis (CM) of Czech Holstein cattle when the mixed reference populations were employed.

2 Material and methods

The edited dataset included 92,388 Holstein cows and 160,426 lactations, with a lactation incidence of clinical mastitis for all lactations at 19.05%. Cows calved between 2017 and 2022 in 119 herds. Only 75% or more Holstein breed cows were included in the edited dataset. The proportions of breed admixture are based on pedigree.

2.1 Phenotypes

Farmers collected CM records and registered them voluntarily in the national cattle health monitoring system called "The Diary of Diseases and Medication" (Kasna et al., 2017). CM was defined as a binary trait with 0 (no case) and 1 (at least 1 case) in a lactation up to 305 days in milk. The records collected from the first to sixth parity were used for analysis. The herds included in the analysis must meet the requirements of regularly recording CM health records and exhibit a minimal CM lactation incidence rate (LIR) of 5% in the recording period. The SAS software package, version 9.4 (SAS, 2012), was used to edit data and calculate correlations and basic statistics.

2.2 Genotypes

A total of 4,969 Holstein sires and 35,814 Holstein cows and heifers were genotyped by the Illumina BovineSNP50 BeadChip V2 (Illumina Inc., San Diego, USA). 54,609 SNP genotypes were available.

Categories of genotyped animals, their numbers and years of birth are listed in Table 1.

Three compositions (GM) of the genomic relationship matrix (G) according to the inclusion of the categories of genotyped animals in the genomic matrix were considered (Table 2 and Figure 1):
- (GM_1) only genotyped bulls,
- (GM_2) genotyped bulls and genotyped cows with a health phenotype,
- (GM_3), genotyped bulls and genotyped cows with and without health phenotype.

Table 1 Genotyped animals, numbers and their years of birth

Category of genotyped Animals	Number	Birth years	Included in analysis
Sires	4,969	1973–2020	GM_1 GM_2 GM_3 GM_1B* GM_2B*
Cows with health phenotype	10,903	2010–2020	GM_2 GM_3
Cows without health phenotype	9,662	2015–2020	GM_3
Cows without health phenotype	10,243	2021–2022	GM_3 GM_1B* GM_2B*

*only in pedigree, without genotype

Table 2   Description of the analyses, number of genotyped animals, pedigrees

Analysis	Number of genotyped animals	Number of animals in pedigree	No. of genotyped bulls born in 2021 and 2022	No. of genotyped females born in 2021 and 2022
GM_1	4,969	200,529	272	–
GM_1B	4,969	220,398	272	10,243*
GM_2	15,872	200,529	272	–
GM_2B	15,872	220,398	272	10,243*
GM_3	35,814	220,398	272	10,243

* only in pedigree, without genotype
GM_1 only genotyped bulls, GM_2 genotyped bulls and genotyped cows with a health phenotype, GM_3, genotyped bulls and genotyped cows with and without health phenotype. The GM_1B and GM_2B analyses complement the respective G matrix layouts; the genotyped females without health phenotypes were added to the pedigree without considering their genotypes for comparison between the analyses
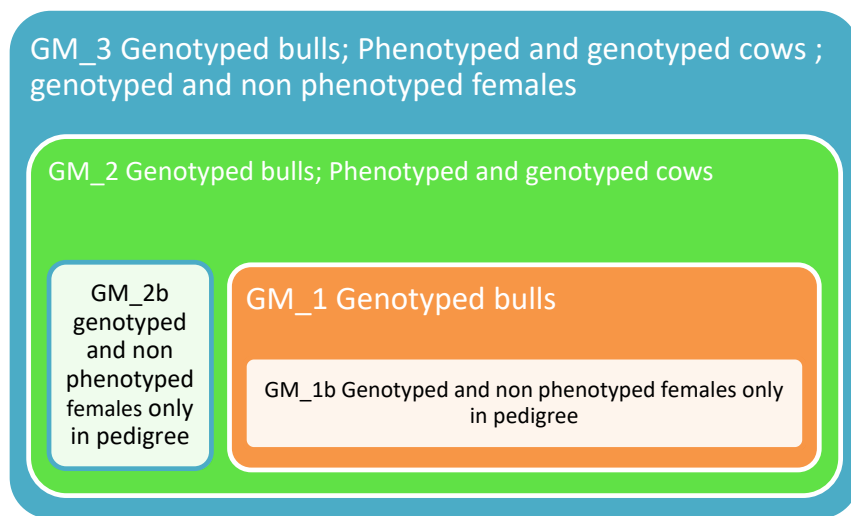

Figure 1   Genomic matrix composition

The GM_1B and GM_2B analyses complement the respective G matrix layouts (Table 2). In GM_1B and GM_2B, the genotyped females without health phenotypes were added to the pedigree without considering their genotypes for comparison between the analyses. Because of their phenotypes, genotyped cows with health phenotypes are included in the GM_1 analysis. Nevertheless, their genotypes were not considered in the GM_1.

The effective no. of SNP was 34,847 (GM_1), 36,250 (GM_2) and 35,144 (GM_3).

2.3 Statistical methods

A mixed linear animal model was used to predict the GEBV for the CM trait.

Model equation:

$$y_{ijklm} = \text{parity\_age}_i + \text{herd}_j + \text{year\_season}_k + pe_l + a_m + e_{ijklm}$$

where y_{ijklm} – the clinical mastitis (CM) as all-or-none trait in a lactation; parity_age$_i$ – the effect of the parity and age at calving i (15 levels); herd$_j$ – the effect of the herd j (119 levels); year_season$_k$ – the effect of the calving year (2017–2021) and season (January–March; April–June; July–September; December) k (19 levels); pe_l – the permanent environmental effect of the cow l; a_m – the additive genetic effect of the animal m; e_{ijklm} – the residual effect

Pedigree (Table 2) included 200,529 animals (GM_1; GM_2): 8,614 bulls and 191,915 cows; or 220,398 animals (GM_1b; GM_2b; GM_3): 8,614 bulls and 211,779 cows.

The single-step genomic method (SSGBLUP) used for genomic prediction enables to obtain GEBV for each animal included in the analysis (Misztal et al., 2009). It included the genomic information following the next equation, where genetic relationships from the genomic data (G matrix) and pedigree data (A matrix) were

summarised/merged into an H matrix according to the following equation:

$$H = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{G} \end{bmatrix} = \mathbf{A} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G} - \mathbf{A}_{22} \end{bmatrix}$$

where $\mathbf{H}$ – a modified relationship matrix incorporating genomic information; $\mathbf{A}_{22}$ – the numerator relationship matrix for genotyped animals; $\mathbf{G}$ – a genomic relationship matrix

A genomic relationship matrix $\mathbf{G}$ was created and scaled using the procedure described by VanRaden (2008) that weights markers by reciprocals of their expected variance pi:

$$G = ZDZ'$$

where: D – diagonal with

$$D = \frac{1}{\left(m_{jj} \left[2p_i(1-p_i) \right] \right)}$$

as in Amin et al. (2007) and Leuttenger et al. (2003). To estimate genomic relationship and inbreeding, VanRaden (2008) introduced $\mathbf{M}$, it is the matrix that specifies which marker alleles each individual inherited. Dimensions of $\mathbf{M}$ are the number of individuals (n) by the number of loci (m). Equations can include marker information using $n \times n$ matrix $\mathbf{MM'}$ or $m \times m$ matrix $\mathbf{M'M}$. If elements of $\mathbf{M}$ are set to -1, 0, and 1 for the homozygote, heterozygote, and other homozygote, respectively, diagonals of $\mathbf{MM'}$ count the number of homozygous loci for each individual, and off-diagonals measure the number of alleles shared by relatives. In contrast, diagonals of $\mathbf{M'M}$ count the number of homozygous individuals for each locus, and off-diagonals measure the number of times alleles at different loci were inherited by the same individual. Let the frequency of the second allele at locus i be p_i, and let $\mathbf{P}$ contain allele frequencies expressed as a difference from 0.5 and multiplied by 2, such that column i of $\mathbf{P}$ is $2(p_i - 0.5)$. Subtracting $\mathbf{P}$ from $\mathbf{M}$ gives incidence matrix $\mathbf{Z}$, which sets the mean values of the allele effects to 0.

The genomic data were edited both by animal and by loci. Animals with genotype call rates <95% were excluded from further evaluation. SNPs were discarded if more than 5% of SNP calls were missing.

The BLUPF90 family programs by Misztal et al. (2018) were employed to predict the genomic breeding values and calculate their accuracy. We run the programme Renumf90 for preparing the renumbered datasets. For the prediction of GEBV, the programme Blupf90, and for calculating the accuracy of GEBV, the programme ACCGS was used.

The accuracy of GEBV by program ACCGS is calculated using the second approximative procedure by Misztal et al. (2013). This method involved the diagonal elements of inverses of the genomic relationship matrix and the pedigree relationship matrix for genotyped animals, and the accuracies are corrected for inflation.

3 Results and discussion

3.1 Genomic breeding values for clinical mastitis

The principal advantage of ss-GBLUP is that it works directly with performance records and whole sets of animals included in the pedigree, simultaneously utilising genomic information of the animals. Genotyped animals' estimates are primarily influenced by using genomic information in breeding value prediction (Lee et al., 2017). Non-genotyped animals' estimates are influenced less according to the degree of kinship to genotyped animals. The changes due to the employment of genomic information appear in the predicted breeding values and the accuracies of breeding values. The changes in GEBV and its accuracy depend on the genetic structure of the genomic relationship matrix and the size of the reference population (VanRaden, 2008).

Table 3 presents the average breeding values of particular groups of animals for different analyses according to the composition of the G matrix. The average values were positive, ranging from 0.18 (young females without phenotypes) to 0.29 (females with phenotypes). The differences in the average GEBV are more noteworthy between the particular groups of genotyped animals than between the analyses. The lowest mean values occurred for bulls and females without phenotypes and the highest for cows with phenotypes (from 0.18 to 0.21). Both young animal groups showed lower mean values compared to older animals, especially cows (from 0.26 to 0.28). The analyses within three groups of genotyped animals, the bulls, young bulls and cows with phenotypes, yielded very similar means and SD: the bulls (0.019–0.023), young bulls (0.18–0.021), cows with phenotypes (0.025 to 0.026). Analysis GM_3 showed the highest SD within all groups. GM_3 means were the same as those for GM_1 except for young bulls born in 2021 and 2022. GM_2 analysis that included the genotyped bulls and the phenotyped and genotyped cows exhibited the highest means and average SD compared to GM_1 and GM_3, except for young bulls.

When the genotyped females without the health phenotype were considered in the analysis but only in pedigree, without their genotypes (GM_1B and GM_2B) included in the genomic matrix, the resulting means were the same when using G matrixes GM_1 and

GM_1B. At the same time, SD in GM_2B decreased in comparison with GM_2. The means and SD of GM_1B and GM_2B were very similar for the groups of genotyped females without the health phenotypes. For GM3, when genotypes of those females are considered, the mean of GEBV decreased, and SD increased.

As Gao et al. (2015) stated, the variance of GEBV was inflated by including genotyped females in genomic prediction, producing more bias than predictions with proven bulls only. The conclusion is that the standard deviation of the breeding values will increase not simply by adding animals to the pedigree but by adding their genotypes. Of course, the GEBV value is also affected. In what direction changes of GEBV appear depends on how the genomic matrix specifies the relationship of animals, especially to the phenotyped animals.

Our analysis's primary aim was to determine if the use of the mixed genomic reference population (including genotyped cows and bulls) instead of only the bulls' genomic reference population influenced the accuracy of individual genomic breeding values, especially for clinical mastitis. Following the literature (Pryce et al., 2012; Obšteter et al., 2019), we hypothesised that it is advantageous in the genomic prediction to use genomic information on cows, not only bulls, because the GEBV accuracy will improve due to the increase in the reference population size.

The results on accuracy are presented in Table 4. We found the highest average accuracy of GEBV for the genotyped cows, not for the genotyped bulls, when it comes to analyses GM_2, GM_2b and GM_3 where the cow's genotype is considered. We assume that the known phenotype plays a vital role in the final GEBV accuracy for cows. On the other hand, the SD of GEBV accuracy for

bulls is significantly higher than for cows, indicating that the range of GEBV accuracy in bulls is much higher than this for cows. Maximal accuracy values for genotyped bulls were 0.91, and the correlation between accuracy and the number of bulls' daughters lay around 56%. It shows that the GEBV accuracies of bulls depend partially on the number of daughters of the bull.

The accuracy of genotyped cows with health phenotypes ranged from 0.06 to 0.51. These cows showed higher GEBV accuracy, probably because they were phenotyped. For the genotyped bulls born in 2021 and 2022, maximal accuracy amounts to 0.37, minimal to 0.08, while for genotyped cows born in 2021 and 2022, maximal accuracy amounts to 0.42, minimal to 0.01. We can conclude that the cows benefited from including their genotypes in the prediction.

Comparing the GM analyses (Table 4), the highest average GEBV accuracy occurred for GM_2 or GM_2B; for the young genotyped bulls born in 2021 and 2022 and the cows with phenotypes born in 2010 and 2020. On the contrary, for the group of genotyped bulls, the decrease occurred in GM_2. Generally, the average accuracy decreased in GM_3, where all genotyped animals were included, for all groups of genotyped animals.

Figures 1 and 2 show the average accuracy per birth year (from 2017 to 2022) for genotyped bulls (Figure 2) and genotyped cows (Figure 3), respectively. It turns out that the increase in accuracy in GM_2 is also observable in genotyped bulls if results are presented per birth year. Following Table 4, the lowest accuracy value for genotyped bulls occurred for GM_3. In Figure 3, the average accuracy for genotyped cows for GM_1 represent the accuracy without cows' genotypes. Therefore GM_1 average is much lower than those of

Table 3 The means of genomic breeding values for the particular groups of animals

	Groups of genotyped animals				
	sires	bulls born in 2021 and 2022	cows with phenotypes born 2010–2020	females without phenotypes born 2015–2020	genotyped females born in 2021 and 2022
Number	4,969	272	10,903	9,699	10,243
Analysis	mean ±SD				
GM_1	0.020 ±0.039	0.021 ±0.030	0.026* ±0.032		
GM_1B	0.019 ±0.038	0.021 ±0.031	0.025* ±0.031	0.021* ±0.031	0.023* ±0.023
GM_2	0.023 ±0.044	0.020 ±0.039	0.029 ±0.046		
GM_2B	0.023 ±0.040	0.021 ±0.029	0.028 ±0.041	0.021* ±0.029	0.024* ±0.026
GM_3	0.020 ±0.056	0.018 ±0.055	0.026 ±0.059	0.018 ±0.055	0.021 ±0.057

* only in pedigree, without genotype
GM_1 only genotyped bulls, GM_2 genotyped bulls and genotyped cows with a health phenotype, GM_3, genotyped bulls and genotyped cows with and without a health phenotype. The GM_1B and GM_2B analyses complement the respective G matrix layouts; the genotyped females without health phenotypes were added to the pedigree without considering their genotypes for comparison between the analyses

Table 4   Average accuracy of genomic breeding values for the groups of animals

	Groups of genotyped animals				
	sires	bulls born in 2021 and 2022	cows with phenotypes born 2010–2020	females without phenotypes born 2015–2020	genotyped females born in 2021 and 2022
Number	4,969	272	10,903	9,699	10,243
Analysis	mean ±SD				
GM_1	0.29 ±0.127	0.22 ±0.037	0.16* ±0.066		
GM_1B	0.29 ±0.129	0.23 ±0.040	0.17* ±0.065	0.14* ±0.062	0.12* ±0.059
GM_2	0.28 ±0.133	0.23 ±0.038	0.36 ±0.052		
GM_2B	0.29 ±0.139	0.24 ±0.041	0.32 ±0.051	0.14* ±0.062	0.12* ±0.059
GM_3	0.25 ±0.146	0.21 ±0.044	0.29 ±0.056	0.28 ±0.054	0.27 ±0.050

* only in pedigree, without genotype
GM_1 only genotyped bulls, GM_2 genotyped bulls and genotyped cows with a health phenotype, and GM_3 genotyped bulls and genotyped cows with and without a health phenotype. The GM_1B and GM_2B analyses complement the respective G matrix layouts; the genotyped females without health phenotypes were added to the pedigree without considering their genotypes for comparison between the analyses

GM_2 and GM_3. The lowest average accuracy for GM_3 in Figures corresponds with Table 4.

If the average accuracy is presented for all bulls (Figure 4) and all cows (Figure 5), the trends found for all bulls (Figure 4) are the same as for genotyped bulls (Figure 2). Still, the averages are lower for all bulls than for the genotyped bulls, and the maximum accuracy occurred in 2018. In the calculation of average accuracy, the nongenotyped bulls with lower accuracy represent the higher ratio of bulls. Figure 5 shows an interesting picture where trends in accuracy for all cows and heifers are captured. The accuracy for GM_2 and GM_3 (0.19–0.24) is lower than for the genotyped females in Figure 3 (0.27–0.33). GM_1 is again the lowest of all layouts. In

Figure 4, GM_2 and GM_3 report the same values until 2019. The reason for these lower accuracy averages in the year 2022 for GM_2 compared with GM_3 is the higher number of genotyped females in GM_3. That increase in the number of genotyped females caused the rise of the GM_3 average accuracy in 2022, as it is represented in Figure 5.

Figures 3 and 4 show that the inclusion of the genotyped animal in genomic evaluation influences the accuracy of GEBV for non-genotyped animals only slightly. The positive effect on the individual GEBV accuracy mainly manifests for the genotyped animals.
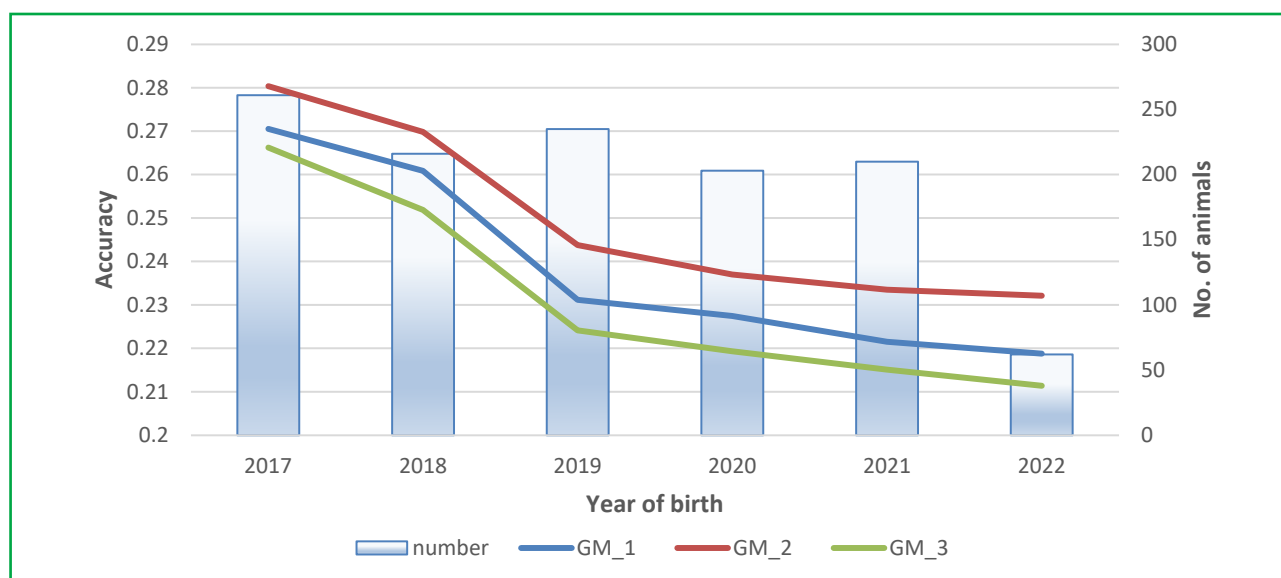

Figure 2   Average accuracy of genomic breeding values for genotyped bulls are presented, GM_1 only genotyped bulls, GM_2 genotyped bulls and genotyped cows with a health phenotype, GM_3, genotyped bulls and genotyped cows with and without health phenotype
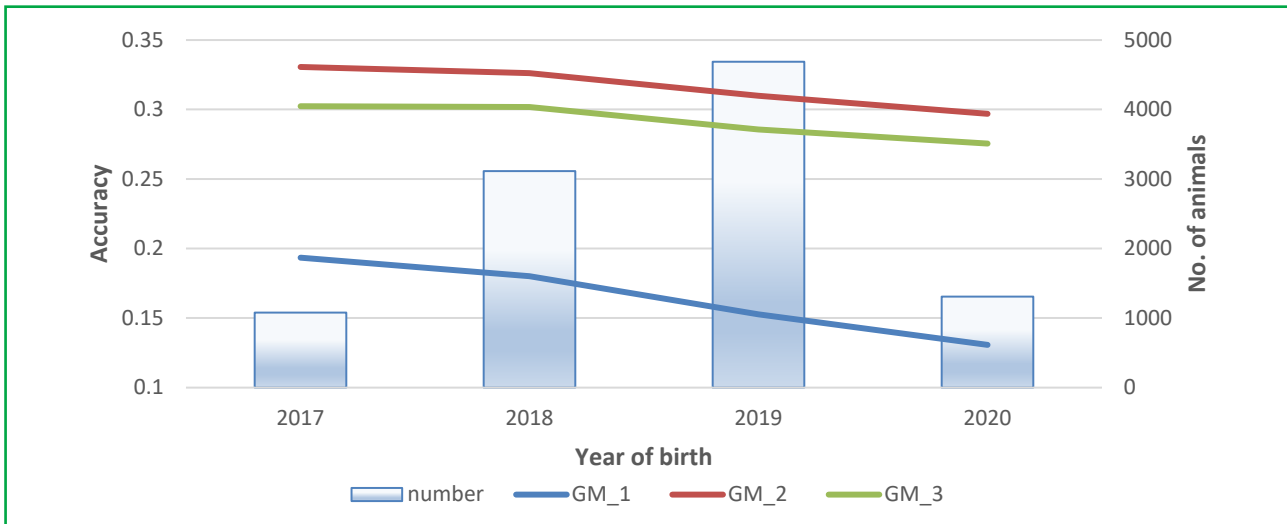
Figure 3 Average accuracy of genomic breeding values for genotyped cows, genotypes of cows involved only in GM_2 and GM_3
GM_1 only genotyped bulls, GM_2 genotyped bulls and genotyped cows with a health phenotype, GM_3, genotyped bulls and genotyped cows with and without health phenotype
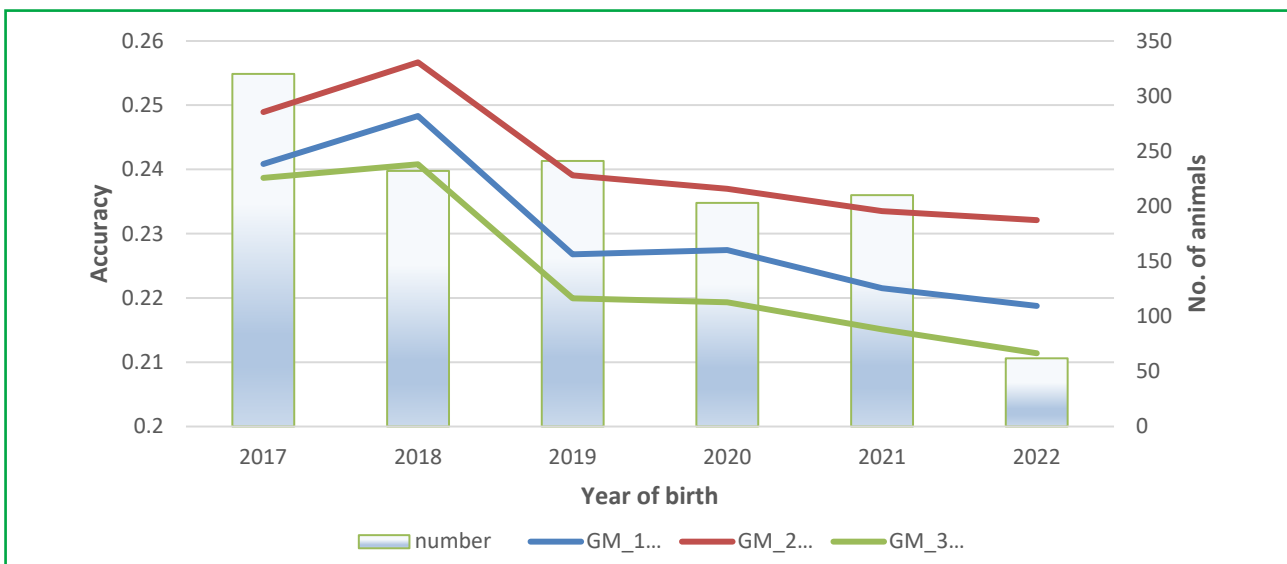

Figure 4 Average accuracy of genomic breeding values for all bulls
GM_1 only genotyped bulls, GM_2 genotyped bulls and genotyped cows with a health phenotype, GM_3, genotyped bulls and genotyped cows with and without health phenotype

When cows were included in the genomic matrix, Nguyen et al. (2016) found some improvement in genomic prediction accuracy, expressed as a correlation between predictions for validation sires, for Holsteins but not for Jersey. They explain the low increase in validation to the low number of genotyped cows. Compared to the analysis we present, the number of cows in the genomic reference population was lower than the number of bulls. The analysis was focused on the production traits while we analysed the health traits. It is not a rule that there will be an increase in reliability with the inclusion of cow's genotypes, as Cooper et al., 2015 found in US Holstein when they added the genotyped cows to the reference population. The reason probably was a large number of genotyped sires in the genomic reference population. Dehnavi et al. (2018) found an increase in the accuracy of genomic prediction of about 0.166 for low heritable traits due to adding cow genotypic and phenotypic information to the bulls' reference population. A slightly higher increase occurred for low heritable traits than for production traits. Jenko et al. (2017) verified that adding the cows to the bulls' genomic reference population increased the accuracy of genomic prediction for production traits and calving interval in Guernsey cattle. The increases in the correlation between GBLUP and BLUP approaches were by 0.060 ±0.015 for milk, 0.036 ±0.019
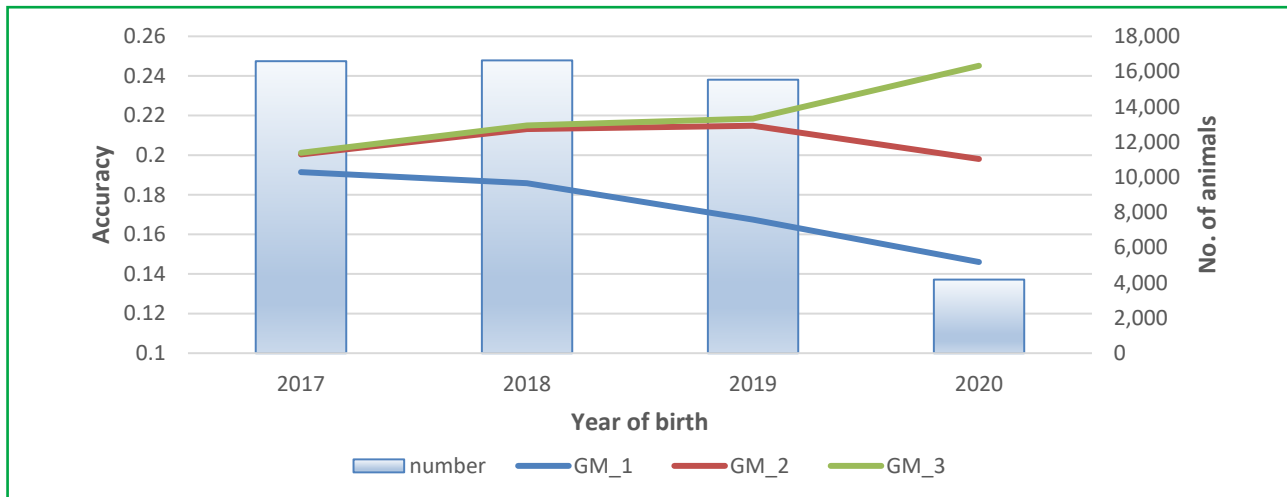
Figure 5 Average accuracy of genomic breeding values for all cows, genotypes of cows involved in GM_2 and GM_3
GM_1 only genotyped bulls, GM_2 genotyped bulls and genotyped cows with a health phenotype, GM_3, genotyped bulls and genotyped cows with and without health phenotype

for fat, 0.033 ±0.015 for protein, and 0.024 ±0.024 for calving interval. For mastitis, Gao et al. (2015) found that the reliability of genomic prediction expressed as squared Pearson correlation coefficient between GEBV and deregressed proof divided by the average reliability of the deregressed proof of the validation bulls achieved a gain of 5.1 percentage points when all genotyped cows were included into the genomic matrix. Pryce et al. (2012) point out that genotyped females may be included in the reference population of genotyped animals but only cautiously considering preferential treatment of cows.

Nonetheless, the inclusion of females in the genomic reference population is beneficial; e.g., a rise of 8% in bulls' reliabilities has been found due to adding 10,000 genotyped cows to the reference population (Pryce et al., 2012). Buch et al. (2012) showed the positive impact of the inclusion of cows in the reference genomic population on the accuracy of GEBV, especially on functional traits in conditions of small-scale phenotyping, which corresponds to conditions in the current population of Czech Holstein cattle and selection for health traits. For Czech Holstein cattle, where the extent of disease monitoring is limited to about a third of the population, we can, according to Buch et al. (2012), expect to increase the reliability of the estimate until the number of phenotyped population surges.

We can justify why the accuracy of genomic breeding values decreased after the inclusion of non-phenotyped genomic females into the prediction. These females were daughters or relatives of the phenotyped cows, but they added no extra information. The only information they provided was about their genotype and, thus, relatedness to other animals in the estimate. The results presented lead to the conclusion that there is no need to include those genotyped cows, which will never again have a phenotype or be candidates for selection, in the calculation of breeding values.

4   Conclusions

It has been confirmed that due to the expansion of the genomic reference population to include a group of genotyped cows with phenotypes, the individual reliability of GEBV for CM has increased. Using genotyped cows with phenotypes is one way of successful genomic selection for clinical mastitis in the population of Czech Holstein cattle. With average accuracy for young bulls and heifers of 0.21 to 0.28, we can conclude that genomic selection is a promising approach to accelerate genetic gains for clinical mastitis resistance. Further increase in the accuracies can be expected, especially for young bulls, after the extension of the monitoring period of health traits and obtaining of a higher pool of historical data, an increasing number of monitored herds and dairy cows, and possibly conversion of the presently applied single-trait linear animal model to a multi-trait model using the information on somatic cells and exterior.

Acknowledgements

References

Amin, N. et al. (2007). A Genomic Background Based Method for Association Analysis in Related Individuals. *Plos One*, 2, e1274. doi: https://doi.org/10.1371/journal.pone.0001274

Bauer, J. et al. (2015): Contribution of domestic and Interbull records to reliabilities of single-step genomic breeding values in dairy cattle. *Czech J. Anim. Sci.*, 60, 263–267. https://doi.org/10.17221/8240-CJAS

Buch, L. H. et al. (2012). The value of cows in reference populations for genomic selection of new functional traits. *Animal*, 6(6), 880–886. https://doi.org/10.1017/S1751731111002205

Dehnavi, E. et al. (2018). The effect of using cow genomic information on accuracy and bias of genomic breeding values in a simulated Holstein dairy cattle population. *Journal of Dairy Science*, 101(6), 5166–5176. https://doi.org/10.3168/jds.2017-12999

Gao, H. et al. (2015). Including different groups of genotyped females for genomic prediction in a Nordic Jersey population. *Journal of Dairy Science*, 98(12), 9051–9059. https://doi.org/10.3168/jds.2015-9947

Heringstad, B. et al. (2001). Responses to selection against clinical mastitis in the Norwegian Cattle population. *Acta Agriculturae Scandinavica Section A-Animal Science*, 51,(2), 155–160. https://doi.org/10.1080/090647001750193503

Jamali, H. et al. (2018). Invited review: Incidence, risk factors, and effects of clinical mastitis recurrence in dairy cows. *Journal of Dairy Science*, 101(6), 4729–4746. https://doi.org/10.3168/jds.2017-13730

Jenko, J. et al. (2017). Cow genotyping strategies for genomic selection in a small dairy cattle population. *Journal of Dairy Science*, 100(1), 439–452. https://doi.org/10.3168/jds.2016-11479

Kasna, E. et al. (2017) National dairy cattle health recording in the Czech Republic. *Proceedings of EAAP – 68th Annual Meeting*, Tallinn 2017, Wageningen Academic Publishers (pp. 251).

Lee, S.H. et al. (2017). Estimation of genomic prediction accuracy from reference populations with varying degrees of relationship. *Plos One*, 12(12), e0189775. doi: https://doi.org/10.1371/journal.pone.0189775

Leutenegger, A.-L. et al. (2003). Estimation of the Inbreeding Coefficient through Use of Genomic Data. *The American Journal of Human*, 73(3), 516–523. https://doi.org/10.1086/378207

Meuwissen, T. H. E. et al. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157(4), 1819–1829. https://doi.org/10.1093/genetics/157.4.1819

Misztal, I. et al. (2009). Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. *Journal of Dairy Science*, 92(9), 4648–4655. https://doi.org/10.3168/jds.2009-2064

Misztal, I. et al. (2018). *Manual for BLUPF90 family programs*. University of Georgia, Athens, USA (142 p.).

Misztal, I., Tsuruta, S., Aguilar, I., Leggara, A., Van Raden, P.M., & Lawlor, T.J. (2013). Methods to approximate reliabilities in single-step genomic evaluation. *Journal of Dairy Science*, 96, 647–654. https://doi.org/10.3168/jds.2012-5656

Nguyen, T.T.T. et al. (2016). Genomic selection for tolerance to heat stress in Australian dairy cattle. *Journal of Dairy Science*, 99(9), 2849–2862. https://doi.org/10.3168/jds.2015-9685

Obšteter, J. et al. (2019). Efficient use of genomic information for sustainable genetic improvement in small cattle populations. *Journal of Dairy Science*, 102(11), 9971–9982. https://doi.org/10.3168/jds.2019-16853

Přibyl, J. et al. (2012). Single-step prediction of genomic breeding value in a small dairy cattle population with strong import of foreign genes. *Czech Journal of Animal Science*, 57(4), 151–159. https://doi.org/10.17221/5890-CJAS

Pryce, J. et al. (2012). Genotyping dairy females can improve the reliability of genomic selection for young bulls and heifers and provide farmers with new management tools. *38th ICAR Conf.*, Cork, Ireland. ICAR, Rome, Italy (pp. 28).

Schaeffer, L. R. (2006). Strategy for applying genome-wide selection in dairy cattle. *Journal of Animal Breeding and Genetics*, 123(4), 218–223. https://doi.org/10.1111/j.1439-0388.2006.00595.x

Thomasen, J. R. et al. (2014). Adding cows to the reference population makes a small dairy population competitive. *Journal of Dairy Science*, 97(9), 5822–5832

Thomasen, J.R. et al. (2020). Genotyping more cows increases genetic gain and reduces rate of true inbreeding in a dairy cattle breeding scheme using female reproductive technologies. *Journal of Dairy Science*, 103(1), 597–606. https://doi.org/10.3168/jds.2019-16974

Vanraden, P.M. (2008). Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science*, 91(11), 4414–4423. https://doi.org/10.3168/jds.2007-0980

Wolfová, M. et al. (2006). Incidence and economics of clinical mastitis in five Holstein herds in the Czech Republic. *Preventive veterinary medicine*, 77(1–2), 48–64. https://doi.org/10.1016/j.prevetmed.2006.06.002