**Discussion**

The estimation of breeding values requires stable investment for continuous collection of data. While breeding programmes usually have a secure funding for phenotyping, the funding for genotyping is not yet well established to initiate and regularly update the training population for genomic prediction. In this paper we propose implementing genomic selection by optimizing the investment into phenotyping of milk production traits and genotyping. We show that by reallocating a part of phenotyping resources to genotyping, we can substantially increase genetic gain regardless of the cost and amount of genotyping, and availability of initial training population. We also show that we can increase the genetic gain even further by increasing the investment into genotyping, despite simultaneously decreasing the amount of phenotyping. Similarly, although reduced phenotyping decreased the phenotype accuracy, genomic prediction increased the accuracy for non-phenotyped selection candidates. These results raised four discussion points 1) how optimizing investment in phenotyping and genotyping affects genetic gain with or without an initial training population; 2) how optimizing investment in phenotyping and genotyping affects accuracy with and without an initial training population; 3) limitations and remarks of the study; 4) implications for breeding programmes.

**1 Genetic gain**

**Implementing genomic selection by optimizing the investment in phenotyping and genotyping increased genetic gain compared to the conventional scenario despite reduced phenotyping, even at minimum investment into genotyping.** In this study we assumed that some small populations have access to an international training population, such as InterGenomics for Brown Swiss in Central Europe (Jorjani, 2012). With an initial 10K training population available, all genomic scenarios outperformed the conventional scenario, mainly due to reduced generation interval. This is in agreement with real data and previous simulation studies. Garcia-Ruiz et al. analyzed US Holstein data and showed that the main driver of genetic gain in genomic selection is the reduced generation interval in the sire of bulls and sire of dams paths, that decreased between 25% and 50% compared to the conventional selection. Simulation studies also confirmed that the genomic selection increases genetic gain due to reduced generation interval, despite reduced selection accuracy (Pryce et al., 2010; Obšteter et al., 2019). Van Grevenhof et al. computed the break-even size of the training population required to achieve a response comparable with conventional selection. They showed, that if the generation is not reduced and the number of phenotypes is limited, genomic selection cannot compete with conventional selection. But as the generation decreases, the break-even size decreases rapidly. When generation interval is halved, only ~2000 or ~3500 individuals are needed to achieve the response of selection on traditional BLUP-EBV based on own performance or 10 progeny per sire.

Another major advantage of the genomic scenarios was increased intensity of sire selection. A costly and lengthy procedure of progeny-testing limits the number of tested sires in the conventional schemes. Genomic selection significantly reduced the cost of testing male candidates (Schaeffer, 2006) and thus allowed for increased intensity of male selection. In US Holstein population, genomic selection increased (improved) the selection differential for all traits, even more for the low heritable ones, such as health and fertility traits (Garcia-Ruiz et al:).

Assisting the superiority of genomic scenarios was also the fact, that although reduced phenotyping decreased the phenotype accuracy, it did not affect the selection accuracy in the same way. While the genomic prediction only slightly decreased the selection accuracy for sires, it actually increased the selection accuracy for females. We discuss the reasons for this in more details below.

**We further increased the genetic gain with increasing the investment into genotyping**, mainly due to increased intensity of sire selection. Investing more into genotyping resulted in more male candidates tested, but same number selected, which intensified selection and drove the genetic gain. This can be seen as increasing investment into genotyping did not further reduce generation interval nor increase the accuracy of sire selection candidates (discussed in the next section). A minor drive of increasing genetic gain was also enlarging the update and total size of the training population. This is in agreement with Thomasen et al., 2020, showing that adding more cows yearly to the training population increases genetic gain. In our simulation a larger training population in turn increased selection accuracy of female candidates. The benefit of this was however diminished, since the intensity of selection in females was very low. It is also worth mentioning, that some of the high-genotyping scenarios achieved the observed genetic gain at a lower total cost, since they did not use all the resources for genotyping females. The saved resources could be invested back into phenotyping females for milk production or novelty traits, genotyping more male candidates, or some other breeding action.

Although increasing investment into genotyping increased genetic gain, the increase was not proportional. Instead, **increasing genotyping had a diminishing return relationship with genetic gain,** **which reached a plateau**. Results showed that when phenotyping and genotyping had equal cost, investing more than the resources of six phenotype records into genotyping did not significantly improve the genetic gain. The first reason for this is, that the accuracy of sire selection in genomic scenario did not increase with increasing genotyping, but was high regardless. We discuss the reasons for this in more details below. Secondly, the intensity of male selection was high in all top performing scenarios. This agrees with Reiner-Benaim et al., 2017, showing that genetic gain increases with the number of tested candidates, but with a diminishing return. They showed that with 4 sires selected, the optimal number of tested calved yielding maximum profit is 1721. They also showed, that 99% or 90% of the profit is achieved with 740 or 119 calves tested. However, they assumed that the price of genomic evaluation is $95, which has reduced since then. And lastly, the top performing scenarios had a large percentage of training population update and a large training population. Due to similar reasons, **we achieved a comparable maximum genetic gain regardless the relative price of phenotyping to genotyping**. In general, selecting less than 2% of the tested males and updating the training population with more than 35% of first parity cows resulted in the maximum genetic gain.

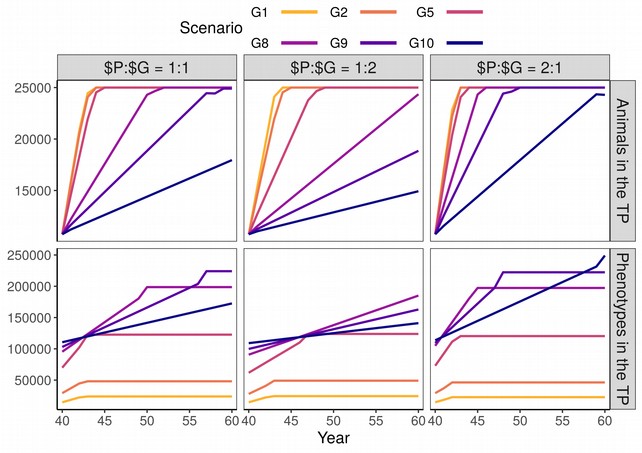
Our results agree with previous studies showing that adding a female to the training population has diminishing return relationship with accuracy and economic genetic gain (Van Grevenhof et al., 2012; Gonzalez-Recio et al., 2014). Gonzales-Recio et al. showed that when the number of females in a training population is small, an additional record has a larger value for the genetic gain than when the female training population is large. Since our scenarios in question all started with a ~10K training population, additional revenue from enlarging the training population was small to begin with, but still decreased with increasing genotyping.

In this study we additionally showed, that while genetic gain does increase with the number of females in training population (although with diminishing return), adding repeated phenotypes does not have the same effect. As the scenarios increased the number of females in the training population, they also decreased the number of (repeated) phenotypes (Figure S1). The top performing scenarios therefore had a training population with the most females but also the least (repeated) phenotypes. However, since we ran single-step genomic prediction, the phenotypes of the non-genotyped animals contributed to the estimation as well.

In this paper we also considered a situation, when small populations **do not have access to a training population and have to initialize one themselves**. **These genomic scenarios still increased the genetic gain between 31% and 134% compared to the conventional scenario**. However, compared to the corresponding scenarios with an initial training population available, these scenarios achieved lower genetic gain. This was mainly due to smaller training population and delay in implementing genomic selection. Increasing the investment into genotyping compensated for starting without a training population due to two reasons. Firstly, investing more into genotyping shortened the delay in implementation of genomic selection (down to one year in high‑genotyping scenarios). And secondly, a smaller initial training population (until reaching 25K) did not proportionally translate into smaller genetic gain, since increasing the number of females in the training population has diminishing return. For example, Gonzales-Recio et al. showed, that for most traits the additional gain from increasing the number of females above 10,000 is negligible.

When implementing genomic selection with a delay, we did not observe any increase in genetic gain above the conventional scenario prior to implementing genomic selection of sires. On the other hand, we also did not observe a decreased genetic gain compared to the conventional scenario prior to the implementation, despite reduced phenotyping. This suggests that breeding programmes can run a conventional breeding programme with reduced phenotyping until they accumulate genotypes to initiate a training population without harming the genetic gain in the accumulation (transition?) period.

However, in this study we did not implement genomic selection in the female path nor did we assume the use of female reproductive technologies, such as embryo transfer. This would further decrease the generation interval and increase genetic gain of genomic scenarios (Pryce et al., 2010; Garcia-Ruiz et al., 2016). Implementing genomic selection of females would require a minor modification of the scheme used in this paper, i. e. genotyping heifers instead in first-parity cows. Regarding female reproductive technologies, some of the tested scenarios saved some of the available resource and could invest in embryo transfer or some other technology.

**Figure S1:**The number of animals and phenotypes in the training population.

**Accuracy**

Accuracy trends follow the genetic gain trends (vice versa?)

1. Accuracy for male candidates high regardless amount of genotyping and phenotyping

* offspring of elite matings, the accuracy of parent average is very high (fathers genotyped)

1. The accuracy for sires decreases with reduced phenotyping, despite increased genotyping

* small number of sires - their accuracy varied considerably
* The accuracy for sires decreased with reduced phenotyping - this is a consequence of us trying to rank (distinguish between) sires (animals) in the tail of the distribution, where details matter – and every additional phenotype helps to correctly differentiate between sires. However, since this is the accuracy after the selection has been made, it is not of great interest for the breeders.

1. Accuracy for dams increases with increasing genotyping, despite reduced phenotyping and less own records

* higher accuracy than with conventional prediction – genomic prediction better estimates Mendelian sampling; higher/better genetic connectedness
* accuracy increasing with increasing investment in genotyping: growing training population, more females genotyped and estimated a genomic breeding value; and better genetic connectedness
* but increases with diminishing returns

1. Accuracy for female candidates follows the dam’s trend and increases with increasing genotyping

* female candidates are not genotyped – the accuracy increases due to increased accuracy of the dams and consequently increased parent average

5. Without an initial training population

6. Papers comparing test-day vs 305-yield models: high correlation of 0.97 (Shaeffer), minor changes in sire ranking (Swalve) and appropriate for Gir; MEUWISSEN PART\_LACT correlations!!!!

* Accuracy of predicting milk yield from alternative milk recording schemes: correlation between A4 and A8 milk yield!!!!
* Prediction of lactation milk yield using various milk recording methods – non-significant sifferences between A4 and A6 methods!!!
* Estimation of 305-day milk yield from test-day records of Chinese Holstein cattle
* Comparison of test-day models for genetic evaluation of production traits in dairy cattle
* Prediction of Daily Milk Yields from a Limited Number of Test Days Using Test Day Models – we can improve the prediction!!!

7. What would happen at different heritability?

The Gonzales-Recio et al. paper also provides insight into what would happen at a different heritability of the phenotype. They showed that as the heritability increases, the contribution of an additional phenotype to the genetic gain decreases. This means less phenotypes (unique phenotypes / females) are needed to achieve the maximum gain and we would need to invest less into genotyping to achieve the maximum genetic gain. In contrast, at lower heritabilities we would need the invest even more to genotype more phenotyped females.

8. Genotype the phenotyped individuals

Bijma; Although increasing the size of the progeny groups increases the accuracy of the sire’s EBV, the number of sires with which the GS reference population is constructed has a much larger impact on the accuracy of genomic EBV. Thus, when the number of phenotypic records is limiting, it is optimal to genotype the individuals that produce the phenotype, not their parents

Figure 1 shows that the increase in the accuracy of genomic EBV with the number of phenotypic records is strongly non-linear, showing a diminishing-return relationship. As a consequence, increasing the total number of phenotypic records increases accuracy less than proportional. I

- not so bad for milk yield, since a lot of historical data available

**Limitations of the study**

**1.** 25K limit for the training population – cite break-even sizes from Bijma and sizes from Gonzales-Recio; also the genetic gain plateaued

2. Only additive effects, single-trait for all lactations (cite correlation between lactation yields)

**Implications**

1. Genotypes also for parentage testing

2. Phenotypes also for management – how many records do you need to manage a herd

3. Cheaper updates with genomic selection!!!