

Genomic evaluation for male fertility in Nordic Red dairy cattle

K. Matilainen¹, G.P. Aamand² & E.A. Mäntysaari¹

¹ *Natural Resources Institute Finland (Luke), Myllytie 1, FI-31600 Jokioinen, Finland
kaarina.matilainen@luke.fi (Corresponding Author)*

² *Nordic Cattle Genetic Evaluation, Agro Food Park 15, Skejby, DK-8200 Århus N, Denmark*

Summary

Currently the Nordic Red dairy cattle breeding goal contains female fertility traits, and the bulls are evaluated by their daughters fertility. In 2016 the joint Nordic (Denmark, Finland, Sweden) fertility evaluations were updated to include conception rate (CR). Conception rate is a binary trait that is based on a success or failure of each insemination. This enables the inclusion of service sire information for each insemination. The data for this study included 21 million CR observations for heifers and cows in parities 1, 2 and 3, in which 24 455 bulls were used as service sire. Pedigree consisted of 6 million animals. A base model was female fertility model for which the effects of random service sire, and his age and breed were added as fixed effects. The males' effects were assumed to affect equally in all 4 genetically correlated female CR traits. As expected, low proportion of service sire variance (0.006) was found when REML was applied for pre-corrected data with female effects regarded as known. Further the study revealed that 2/3 of the service sire variance was related to the non-genetic and 1/3 for the genetic effect. For a multi-step genomic evaluation, the predicted service sire effects were deregressed using inverse of the shrinkage factor. Consequent genetic and genomic evaluations applied BLUP and GBLUP for the rescaled service sire effects of 6 148 genotyped bulls. Number of inseminations was used as weight. Genotypes included 46 914 SNPs. Cross-validation was made to assess validation reliability and bias of predicted genomic breeding values (GEBVs) or parental averages (PAs) through regressing the observed rescaled service sire effects of validation bulls on the GEBVs or PAs. Genotyped bulls that had at least 500 inseminations formed a group of 2 753 validation bulls which was further divided randomly to five groups. Validation reliabilities of GEBVs (PAs) were on average 0.26 (0.18) and regression coefficient for GEBVs (PAs) was on average 0.95 (1.03) over five validation groups. In conclusion, genomic evaluation for male fertility was feasible and it was more reliable than pedigree based genetic evaluation. Despite low heritability of male fertility for one insemination, high reliability was obtained for bulls with at least 500 inseminations.

Keywords: male fertility, field data, GBLUP, Nordic Red dairy cattle

Introduction

In dairy cattle the success of insemination depends both on the male and female fertility, but currently the breeding goal in the Nordic Red dairy cattle (RDC) contains mainly the female fertility traits. In 2016, the joint Nordic (Denmark, Finland, Sweden) fertility evaluations were updated to include conception rate (CR). Conception rate is a binary trait with observations on either a success or failure of each insemination. This trait definition enables the inclusion of service sire information to each insemination record.

Female fertility traits have generally low heritability, especially for the binary traits. For CR in Nordic Red dairy cattle heritability of 0.01 and 0.02 are assumed for heifer and cows, respectively. Heritability for male fertility has been reported to be even smaller, probably due to strict semen quality control. For instance, Berry *et al.* (2011) and Kuhn & Hutchison (2008) reported heritabilities of 0.001 and 0.0002 for CR, respectively, and Hyppänen & Juga (1998) and Stålhammar *et al.* (1994) found heritability of 0.001 and 0.000-0.006 for non-return rate. This has led to a situation that bulls are evaluated based on their average phenotypic conception rate or on the predicted conception rate values.

The main aim of this study was to apply two-step genomic evaluation for male fertility in Nordic RDC. For this, at first the joint female and male model was implemented, and after that, the service sire effects were used in the GBLUP to predict genomic enhanced breeding values (GEBVs).

Material and methods

Data

The RDC CR data used for the joint Nordic fertility evaluations in January 2017 was obtained from the Nordic Cattle Genetic Evaluation (NAV, Aarhus, Denmark). Data included 21 million CR observations from 4.6 million heifers and cows in parities 1, 2 and 3. Approximately 5% of the CR observations had unknown service sire and for the rest 24 455 bulls were used. In 3% of the inseminations with known bulls, the bull and the cow were of different breed. Original pedigree was expanded to include also young service sires that had not yet daughters with observation in the female fertility evaluation. The pedigree consisted of 6 million animals from which 6 148 RDC bulls were genotyped using 50k bead chips. After edits, a total of 46 914 SNPs were included in the analyses.

Models

GBLUP for male fertility was performed in two steps: The first step included genetic evaluation for the combined female and male fertility and the second step included genomic evaluation for male fertility solutions.

For the first step a base model was the female fertility model (Tyrisevä *et al.*, 2017) for which the effects related to service sire were added as follows:

$$\mathbf{y} = \mathbf{X}_a \mathbf{b}_a + \mathbf{X}_s \mathbf{b}_s + \mathbf{Q}_p \mathbf{p} + \mathbf{Q}_s \mathbf{s} + \mathbf{Z} \mathbf{a} + \mathbf{e}, \quad (1)$$

where \mathbf{b}_a and \mathbf{b}_s are vectors of fixed effects related to females and service sires, respectively, \mathbf{s} is vector of random service sire effects, \mathbf{p} and \mathbf{a} are vectors of permanent environment and additive genetic effects of females, respectively, and \mathbf{e} is vector of random residuals. \mathbf{X} , \mathbf{Q} and \mathbf{Z} are incidence matrices that relate the appropriate effects to each observation. Random effects were assumed to follow normal distribution with mean zero and variances $\mathbf{I}\sigma_s^2$, $\mathbf{I}\sigma_p^2$ and $\mathbf{A}\sigma_a^2$ for service sire, permanent environment and additive genetic effects, respectively, where \mathbf{I} is identity matrix of size equal to number of effect levels and \mathbf{A} is a numerator relationship matrix. Fixed effects related to service sire were bull's age at insemination (grouped as 1, 2, and at least 3 years old) and breed with identification of pure- or crossbred insemination. Service sire effects were assumed to be the same for each of the four parities.

In the second step, the predicted service sire effects were deregressed using inverse of

the shrinkage factor as follows:

$$\lambda_j = \frac{r}{1 + r(N_j - 1)} \quad (2)$$

where r as proportion of service sire variance and N_j is number of inseminations made by bull j . Finally, the GBLUP was applied for deregressed service sire effects (SCR) of the 6 148 genotyped RDC bulls. The model contained non-genetic and genetic effects for service sires and the number of inseminations was used as weight as follows:

$$\text{SCR} = \mathbf{1}\mu + \mathbf{Z}\mathbf{pe} + \mathbf{Z}\mathbf{ss} + \mathbf{e}^* \quad (3)$$

where μ is the overall mean and random effects \mathbf{pe} , \mathbf{ss} and \mathbf{e}^* are normally distributed: $\mathbf{pe} \sim N(\mathbf{0}, \mathbf{I}\sigma_{pe}^2)$, $\mathbf{ss} \sim N(\mathbf{0}, \mathbf{G}\sigma_g^2)$ and $\mathbf{e}^* \sim N(\mathbf{0}, \mathbf{W}\sigma_e^2)$ with inverse of weights in the diagonal of \mathbf{W} . Genomic relationship matrix \mathbf{G} was made by method 1 in Van Raden (2008) with multiplication of diagonals with 1.001 to prevent problems in inverting the \mathbf{G} . Data allele frequencies were used in constructing \mathbf{G} . For comparison, also genetic evaluation with pedigree based relationship matrix \mathbf{A} instead of genomic relationship matrix \mathbf{G} was applied.

Variance parameters for the service sire effects in the two models were estimated by REML using the pre-corrected data with female effects regarded as known. Low proportion of service sire variance was obtained ($r = 0.006$) which was expected as reported in other studies. This was used for the service sire effect in the first model, and the parameters for the female effects were adopted from the official NAV evaluations. Further study revealed that proportion of the variance of the non-genetic service sire effect was 0.004 and that for the additive genetic effect (h^2) was 0.002. This information was used in the second model.

Validation

Five-fold cross-validation was used to assess validation reliability and bias of GEBVs and parental average EBVs (PAs). Same data and parameters were used for GEBV and PA analyses. Bulls that had at least 500 inseminations but no male progenies formed a group of 2 753 possible validation bulls. Those bulls were divided randomly to five sub-groups with 550 or 551 validation bulls in each, which was 9% of all genotyped bulls, while the rest (4/5 of 2 753) were included in the bull reference. Validation reliability and bias for each of the five sub-groups were obtained by regression of the SCR on the breeding value predictions of the validation bulls. The resulting coefficient of regression was regarded as bias, and reliabilities were obtained as squared correlation between observed and predicted SCRs, divided by the average reliability of the SCRs as follows:

$$r_{\text{val}} = \frac{b}{\sqrt{\text{RE}}} \quad (4)$$

Results are reported as mean validation reliability and bias over the five validation sub-groups.

Results and discussion

The data analyses were done using the MiX99 program package, which uses iterative preconditioned conjugate gradient (PCG) method. Convergence was set to meet when the sum of squares of relative differences between solutions of consecutive iteration rounds was less

than 1e-10. Both the original and joint model took less than 2500 PCG iterations, although solving time increased from 1 hour to 2 hours. The analyses for genotyped sires in the second step converged fast. Genomic evaluation took less than one minute for 141 PCG iterations.

As shown in Figure 1, no clear genetic trend was observed. Genetic trends based on the standardized GEBVs and EBVs followed nicely each other although correlation between GEBVs and EBVs over all 6 148 bulls in the evaluations was 0.88. The number of common sires among the best 50 genotyped bulls was found to be 22.

Instead of the top bulls, more interesting would have been to find out the worst sires as soon as possible. Unfortunately, prediction of some worst sires was not possible with the current model. One reason for the poor male fertility can be deleterious haplotypes, for example a deletion on chromosome 12 that is lethal in homozygous embryo (Kadri *et al.*, 2014). For example, in Figure 2, the sire with worst SCR has been tested to be a carrier, but this wasn't possible to predict based on the relationships.

Overall validation of bulls succeeded well. Average reliability of the SCRs was 0.28 among the validation bulls with at least 500 inseminations. Validation reliabilities of GEBVs ranged between 0.22 and 0.31 for the five validation sub-groups, and the average was 0.26. Validation reliabilities of EBVs were lower for all five sub-groups. They ranged between 0.10 and 0.23 and the average was 0.18. Both the GEBVs and EBVs were non-biased: regression coefficients for GEBV and EBVs were on average 0.95 and 1.03, respectively.

In conclusion, the joint analysis of male and female fertility is advantageous because of its ability to include all relevant information. Further genomic evaluation for male fertility was feasible and it was more reliable than corresponding genetic evaluation. Good reliability results were obtained for bulls with at least 500 inseminations despite the low heritability of service sire effect for one insemination. GEBV for animals with lethal haplotypes were impossible to predict however. Further study is required to find out how this could be taken into account in the final model.

List of References

- Berry, D.P., R.D. Evans & S. Mc Parland, 2011. Evaluation of bull fertility in dairy and beef cattle using cow field data. *Theriogenology* 75: 172-181.
- Hyppänen, K. & J. Juga, 1998. Environmental and genetic effects on the 60-day nonreturn rate in Finnish AI bulls. *Interbull Bulletin* 18: 91-98.
- Kadri, N. K., G. Sahana, C. Charlier, T. Iso-Touru, B. Guldbbrandtsen, L. Karim, U. S. Nielsen, F. Panitz, G. P. Aamand, N. Schulman, M. Georges, J. Vilkki, M. S. Lund & T. Druet, 2014. A 660-Kb deletion with antagonistic effects on fertility and milk production segregates at high frequency in Nordic Red Cattle: Additional evidence for the common occurrence of balancing selection in livestock. *PLoS Genet* 10: e1004049.
- Kuhn, M.T. & J.L. Hutchison, 2008. Prediction of dairy bull fertility from field data: Use of multiple services and identification and utilization of factors affecting bull fertility. *J. Dairy Sci.* 91: 2481-2492.
- Stålhammar E.-M., L. Janson & J. Philipsson, 1994. Genetic studies on fertility in AI bulls. II. Environmental and genetic effects on non-return rates of young bulls. *Anim. Reprod. Sci.* 34: 193-207.
- Tyrisevä, A.-M., K. Muuttoranta, J. Pösö, U.S. Nielsen, J.-Å. Eriksson, G.P. Aamand, E.A. Mäntysaari & M.H. Lidauer, 2017. Evaluation of conception rate in Nordic dairy cattle. *Interbull Bulletin*.
- VanRaden, P. M., 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:

4414-4423.

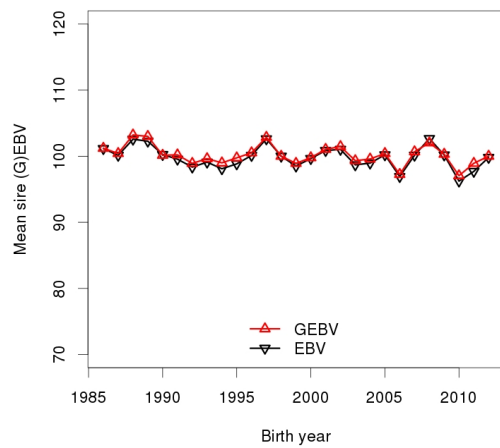


Figure 1. Genetic trend based on standardized genomic enhanced breeding values (GEBV) and estimated breeding values (EBV) of bull fertility based on two-step GBLUP and BLUP, respectively.

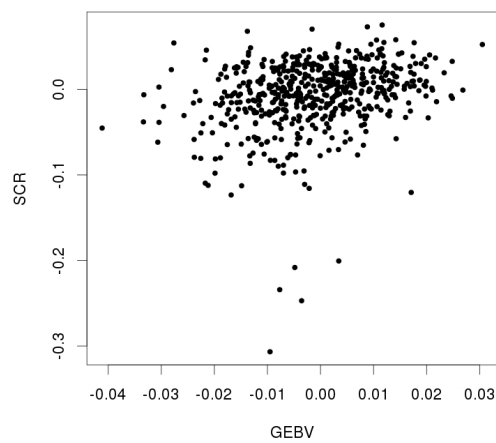


Figure 2. Deregressed service sire effects (SCR) and genomic enhanced breeding values (GEBV) based on the reduced data for a group of validation bulls.