Genetic parameters for mid-infrared spectroscopy–predicted fertility

I. van den Berg,1* P. N. Ho,1 M. Haile-Mariam,1 and J. E. Pryce1,2

Graphical Abstract

Summary

MFERT is the probability of conception to first insemination using mid-infrared spectroscopy of a milk sample collected during lactation as part of routine milk recording, age at calving, days in milk, and milk production. Because MFERT can be estimated during routine milk recording, it can become available on a much larger number of records than the traditional fertility phenotypes. To test the potential of MFERT for genetic selection, we estimated genetic parameters for MFERT. While MFERT had a higher heritability than traditional fertility phenotypes, genetic correlations between MFERT and traditional fertility phenotypes were low to moderate. The strongest genetic correlation was found with calving to first service, with a genetic correlation of −0.61. Therefore, MFERT will only be able to increase the accuracy of a fertility index if there are many more records for MFERT than for traditional fertility phenotypes.

Highlights

• Mid-infrared (MIR)-predicted fertility has greater heritability than traditional fertility phenotypes.
• MIR-predicted fertility has moderate genetic correlations with traditional fertility phenotypes.
• Because MIR-predicted fertility mainly requires a milk sample, it could be available on many more cows than traditional fertility phenotypes.
• MIR-predicted fertility can increase the accuracy of a fertility index when available on a much larger number of cows than traditional fertility phenotypes.
Genetic parameters for mid-infrared spectroscopy–predicted fertility

I. van den Berg,¹*, P. N. Ho,¹ M. Haile-Mariam,¹ and J. E. Pryce¹,²

Abstract: Female fertility is a challenging trait to improve genetically because of its low heritability, its unfavorable genetic correlation with milk yield, and its relatively small number of records. The MFERT trait is the probability of conception to first insemination predicted using mid-infrared (MIR) spectroscopy of a milk sample collected during lactation as part of routine milk recording, age at calving, days in milk, and milk production. As such, MFERT could become available on many more cows than traditional fertility traits. Our objectives were (1) to estimate the heritability of MFERT; (2) to estimate genetic correlations between MFERT, traditional fertility traits, and milk production traits; and (3) to assess the potential of MFERT to be used as an indicator trait for fertility in a selection index. The MFERT trait had a heritability of 0.16, which was higher than that (0.05) obtained for traditional fertility traits. Genetic correlations between MFERT and traditional fertility traits were low to moderate. The weakest and strongest correlations (mean ± standard error) were with pregnancy at the end of the mating season (0.13 ± 0.05) and calving to first service (−0.61 ± 0.03), respectively. Based on our estimates, including MFERT in a fertility index will only substantially increase the accuracy of the index when there are many more records available for MFERT than for the traditional fertility traits. This is likely to be the case because the number of milk samples from commercial machines belonging to milk recording companies in Australia that are capable of generating MIR spectra is growing. Hence, the number of records for MFERT is expected to increase substantially in the near future.

Female fertility is an important trait in dairy cattle. Many studies have shown that it has been subject to declines due to its negative genetic correlation with production traits (Berry et al., 2014; Lucy, 2019). As a response to this, fertility has been a major part of the breeding objective worldwide over the last 20 yr (Miglior et al., 2005; Cole and VanRaden, 2018). Although this has led to favorable genetic trends in fertility (Berry et al., 2014), fertility remains the most common reason for culling in Australia (Workie et al., 2019). The low heritability of fertility (Ma et al., 2019) and unfavorable genetic relationships with other traits under selection limit the response to selection. Furthermore, variation in economic factors and genetic parameters between countries results in differences in the magnitude of response to selection (Pryce et al., 2014). Hence, it is more challenging for some countries to achieve positive genetic trends than for others.

The accuracy of EBV, whether they are predicted based on progeny records or genomic prediction, increases when the number of records for a trait increase. A larger number of progeny records increases the accuracy of the breeding value of a bull, and a larger number of records in the reference population used for genomic prediction increases the accuracy of genomic prediction (Goddard, 2009). This is especially true for traits with low heritability. If the reference population is sufficiently large, it can provide accurate breeding values for all genotyped bulls, including those with no or limited number of progeny records. Although fertility is routinely recorded for a large number of cows in Australia, the number of fertility records is smaller than that of other traits such as milk yield. Furthermore, unlike milk production data, most fertility data are not available when selection decisions are made. Having a fertility record for all cows with production records could help to substantially increase the reference population.

As part of routine milk recording, mid-infrared (MIR) spectroscopy is used to quantify the composition of milk, including fat, protein and lactose concentration. Recently, Ho et al. (2019) developed an equation to predict the probability of conception of dairy cows to first insemination (MFERT), using MIR data, age at calving, DIM, and milk production. MFERT was able to phenotypically classify cows into high or low fertile groups. The inference here is that by training the MFERT model that contrasts the most fertile cows (i.e., cows that conceived at first insemination) with the least fertile (i.e., cows with only one insemination within the mating season without conception), clearer biological signals might be observed and thus higher modeling accuracy could be obtained (Ho et al., 2019; Ho and Pryce, 2020).

Currently, the breeding objective of the fertility Australian Breeding Value (ABV) includes 6-week-in-calf rate derived from calving interval (CI), lactation length (LL), pregnancy at the end of the mating season (PREG), first-service nonreturn rate, and interval from calving to first service (CFS) using a 5-trait model (Haile-Mariam et al., 2013). We hypothesized that MFERT may be useful to be included as an additional trait for predicting the EBV for fertility. However, for this to be successful for genetic improvement of fertility, MFERT should be heritable, significantly genetically correlated with other fertility traits already in the breeding objective, and be available in a large quantity. Currently, the genetic parameters of MFERT are unknown.

The negative genetic correlation between milk yield (MY) and fertility also means that MY could be used as an indicator trait.
to predict fertility (Biffani et al., 2005; Harris et al., 2005). Milk yield has a higher heritability than fertility and is available on a larger number of cows than fertility traits. However, including MY for predicting fertility may unintentionally mean selecting for lower MY; as such (although of interest to include in this study), it is unlikely to be considered as a selection criterion for national evaluations of fertility. As MFERT is derived from a milk sample, understanding genetic correlations with milk production traits is also important. Our objectives were (1) to estimate the heritability of MFERT; (2) to estimate genetic correlations among MFERT, traditional fertility traits, and MY; and (3) to compare the potential of MFERT to be used as indicator trait for fertility in a selection index with MY.

Table 1 summarizes the data used for our analyses. Phenotypes for traditional fertility traits (TFERT) including CI, LL, CFS, and PREG, as well as test-day MY, test-day MY, test-day protein yield (FY), test-day fat percentage (FY), and test-day protein percentage (P%) were provided by DataGene (Bundoora, Victoria, Australia), for 618,856, 598,732, 357,110, 302,069, 642,617, 641,340, 641,339, 641,341 and 641,399 Holstein cows with records for CI, LL, CFS, PREG, MY, FY, P%, and P%, respectively, with one record per cow. Depending on the trait, between 3,157 (MFERT) and 39,177 (MY) cows with phenotypes were used. The MIR spectra used in our analyses were derived using the model described in detail by Ho et al. (2019) and Ho and Pryce (2020). Data used for our study were collected between 2016 and 2018 (inclusive) from 29 commercial Australian dairy herds. The MIR spectra were obtained from the analysis of milk samples at Hico Pty Ltd. (Maffra, Victoria, Australia), TasHerd Pty Ltd. (Hadsdpen, Tasmania, Australia), or DairyExpress (Armidale, New South Wales, Australia) using NexGen Series FTS Combi machines (Bentley Instruments). Other data such as age at calving, DIM at herd testing, and milk production were obtained from DataGene. The MIR spectra used in our analyses were recorded between 0 and 265 DIM, with 97% of MFERT records being within 150 DIM.

To independently obtain the value of MFERT for cows in each herd, we applied a procedure that is similar to external validation. To achieve this, we predicted the probability of conception of cows (ranging from 0 to 1) in each herd in turn by applying a model that was trained using data from the other 28 herds. Hence, the model was completely independent from the herd being predicted. In this study, we only used predicted probabilities of Holstein cows. Where a cow had multiple records within a lactation, only the first record was selected. This resulted in 4,124 cows with MFERT records.

Genetic parameters were estimated using GIBBS2F90 (Misztal et al., 2014). A combined pedigree (constructed using 3 generations of pedigree) and genomic relationship matrix (H) was constructed following Aguilar et al. (2010), using genotypes of 576,431 variants on the HD chip. A univariate animal model fitting fixed effects (herd-year-season, age within parity, and month of calving), alongside H and a vector of random residuals, was used to estimate heritabilities. Bivariate model with the same effects was used to estimate genetic correlations between traits.

To assess the potential advantage of including MFERT in a fertility index, we compared the accuracy of the EBV of a bull for several indices. The basic index contained CI, LL, CFS, and PREG and is referred to as FERTI, where \( n_I \) is the number of progeny records for CI, LL, CFS, and PREG, with \( n_I = 0, 10, 30, \) or 60 for each of the TFERT traits. The accuracy of the FERT index was compared with indices with \( n_I \) records for FERT and 60 for MFERT (FERTI+60, MY (FERTI+60, MY), or MFERT and MY (FERTI+60, MY). The accuracy of each index was calculated using software developed by J. van der Werf (https://jyanderw.une.edu.au/software.htm), following Mrode (2005):

\[
\rho_{HI} = \frac{\sigma_I}{\sigma_H},
\]

where \( \sigma_I \) is the standard deviation of the index and \( \sigma_H \) the standard deviation of the breeding objective. In our indices, the breeding objective was always CI. All other traits were used as selection criteria; that is, as correlated traits with economic weights of 0. Hence, the standard deviation of the breeding objective equaled the standard deviation of CI.

The heritability of MFERT was substantially larger than that of the fertility traits currently included in the Australian selection indices (0.16 vs. 0.05; respectively; Table 2). Heritabilities for TFERT were in line with the report of Ma et al. (2019). As a consequence of the number of records available, the standard error of

| Acronym | Trait | nP | nG | Mean | SD |
|---------|-------|----|----|------|----|
| CI      | Calving interval (d) | 618,856 | 35,014 | 410 | 61 |
| LL      | Lactation length (d) | 598,732 | 34,050 | 360 | 59 |
| CFS     | Calving to first service (d) | 357,110 | 34,899 | 93.7 | 46.30 |
| PREG    | Pregnancy (0,1) | 302,069 | 30,633 | 0.9 | 0.36 |
| MFERT   | Probability of conception to first insemination predicted using MIR spectroscopy (0, 1) | 4,124 | 3,157 | 0.8 | 0.12 |
| MY      | Test-day milk yield (L) | 642,617 | 39,177 | 27.6 | 8.07 |
| FY      | Test-day fat yield (kg) | 641,340 | 39,107 | 1.0 | 0.34 |
| PY      | Test-day protein yield (kg) | 641,399 | 39,111 | 0.9 | 0.25 |
| P%      | Test-day fat percentage (%) | 641,340 | 39,107 | 0.9 | 0.34 |
| P%      | Test-day protein percentage (%) | 641,399 | 39,111 | 3.1 | 0.32 |
the heritability of MFERT was larger than that of TFERT. Standard errors were 0.03 for MFERT and ≤0.01 for TFERT. Despite the relatively large standard error, the heritability of MFERT was more significant than that of TFERT. This means that genetic improvement of MFERT may be easier to achieve than that of TFERT. MFERT might have a higher heritability than the other fertility traits because some of the factors that contribute to the residual variance of TFERT. For example, culling and recording errors may affect MFERT less than TFERT. Furthermore, the trait MFERT predicts the probability of conception and may be a more precise and therefore more heritable phenotype than TFERT.

MFERT had significant genetic correlations with all other traits. The weakest correlation was found with PREG (0.13) and the strongest correlation with CFS (−0.61; Table 2). Based on the trait definitions, we expected MFERT to be closer to PREG than CI, LL, and CFS: MFERT predicts the probability of conception, whereas PREG measures the outcome of conception. However, genetic correlations were stronger with CI, LL, and CFS than with PREG. This may be because, like MFERT, CI, LL, and CFS are all continuous traits, whereas PREG is a binary trait. Because the strongest correlation was detected between MFERT and TFERT, adapting MFERT to predict CFS rather than PREG may further increase the genetic correlation between MFERT and CFS. Furthermore, MFERT is better at predicting high and low fertile cows than moderately fertile cows (Ho et al., 2019; Ho and Pryce, 2020). If we assume a normal distribution for fertility, moderately fertile cows would make up a substantial proportion of our data. This may be earlier than most TFERT. While having MFERT records obtained routinely through herd testing would result in a much larger number of records for MFERT than for TFERT, the current situation in Australia is the reverse: we have many more records for TFERT than for MFERT. Therefore, the benefit of using MFERT for genetic selection is limited at the moment. However, MFERT is a promising tool for phenotypic prediction and management (Ho et al., 2019; Ho and Pryce, 2020). MIR spectra are being collected and used for prediction of a wide range of other traits (e.g., Ho et al., 2021). Consequently, the number of records for MFERT

| Trait | $h^2 ± SE$ | $r_{g,}\text{MFERT} ± SE$ | $r_{g,}\text{MY} ± SE$ | $r_{g,}\text{P%} ± SE$ | $r_{g,}\text{P%} ± SE$ |
|-------|------------|------------------------|----------------------|------------------|------------------|
| CI    | 0.05 ± 0.01 | −0.32 ± 0.04           | 0.44 ± 0.02          | −0.03 ± 0.03     | −0.23 ± 0.02     |
| LL    | 0.05 ± 0.00 | −0.36 ± 0.03           | 0.45 ± 0.03          | −0.08 ± 0.02     | −0.28 ± 0.02     |
| CFS   | 0.05 ± 0.01 | −0.61 ± 0.03           | −0.08 ± 0.04         | 0.05 ± 0.04      | 0.06 ± 0.04      |
| PREG  | 0.05 ± 0.01 | 0.13 ± 0.05            | −0.32 ± 0.05         | 0.50 ± 0.05      | 0.46 ± 0.04      |
| MFERT | 0.16 ± 0.03 | —                      | −0.32 ± 0.05         | 0.50 ± 0.05      | 0.46 ± 0.04      |
| MY    | 0.11 ± 0.01 | −0.32 ± 0.05           | —                    | 0.52 ± 0.01      | −0.53 ± 0.01     |
| FY    | 0.08 ± 0.00 | 0.24 ± 0.06            | 0.37 ± 0.02          | 0.59 ± 0.01      | 0.07 ± 0.02      |
| PY    | 0.07 ± 0.00 | −0.14 ± 0.04           | 0.82 ± 0.00          | −0.26 ± 0.02     | 0.04 ± 0.02      |
| F%    | 0.14 ± 0.01 | 0.50 ± 0.05            | −0.52 ± 0.01         | —                | 0.53 ± 0.01      |
| P%    | 0.16 ± 0.00 | 0.46 ± 0.04            | −0.53 ± 0.01         | 0.53 ± 0.01      | —                |

1CI = calving interval; LL = lactation length; CFS = calving to first service; PREG = pregnancy; MFERT = mid-infrared spectroscopy predicted fertility; MY = test-day milk yield; FY = test-day fat yield; PY = test-day protein yield; F% = test-day fat percentage; and P% = test-day protein percentage.
is expected to increase substantially in the near future as MIR data become routinely collected and stored from milk recording organizations across Australia. When more data for MFERT become available, we recommend re-estimating the genetic parameters on a larger data set and subsequently reassessing the potential benefit of including MFERT in a fertility index. In addition to more animals with records, having more records per cow, both within lactation and across lactations, may also help improve the accuracy of the MFERT prediction equations, the estimation of genetic parameters, and genomic prediction.

References

Aguilar, I., J. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. J. Dairy Sci. 93:743–752. https://doi.org/10.3168/jds.2009-2730.

Berry, D. P., E. Wall, and J. E. Pryce. 2014. Genetics and genomics of reproductive performance in dairy and beef cattle. Animal 8:105–121.

Biffani, S., M. Marusi, F. Biscarini, and F. Canavesi. 2005. Developing a genetic evaluation for fertility using angularity and milk yield as correlated traits. Interbull Bull. 33:63.

Cole, J. B., and P. M. VanRaden. 2018. Symposium review: Possibilities in an age of genomics: The future of selection indices. J. Dairy Sci. 101:3686–3701. https://doi.org/10.3168/jds.2017-13335.

Douglas, M. L., L. C. Maret, K. L. Macmillan, J. M. Morton, M. C. Hannah, A. D. Fisher, and M. J. Auldist. 2016. Associations of high and low milk protein concentrations with energy allocation, milk production, and concentrations of blood plasma metabolites and hormones in Holstein-Friesian cows. J. Dairy Sci. 99:10057–10066. https://doi.org/10.3168/jds.2016-11388.

Goddard, M. 2009. Genomic selection: Prediction of accuracy and maximisation of long term response. Genetica 136:245–257. https://doi.org/10.1007/s10709-008-9308-0.

Haile-Mariam, M., P. J. Bowman, and J. E. Pryce. 2013. Genetic analyses of fertility and predictor traits in Holstein herds with low and high mean calving intervals and in Jersey herds. J. Dairy Sci. 96:655–667. https://doi.org/10.3168/jds.2012-5671.

Harris, B. L., J. E. Pryce, Z. Z. Xu, and W. A. Montgomery. 2005. Fertility breeding values in New Zealand, the next generation. Interbull Bull. 33:47.

Ho, P. N., V. Bonfatti, T. D. W. Luke, and J. E. Pryce. 2019. Classifying the fertility of dairy cows using milk mid-infrared spectroscopy. J. Dairy Sci. 102:10460–10470. https://doi.org/10.3168/jds.2019-16412.

Ho, P. N., T. D. W. Luke, and J. E. Pryce. 2021. Validation of milk mid-infrared spectroscopy for predicting the metabolic status of lactating dairy cows in Australia. J. Dairy Sci. 104:4467–4477. https://doi.org/10.3168/jds.2020-19603.

Ho, P. N., and J. E. Pryce. 2020. Predicting the likelihood of conception to first insemination of dairy cows using milk mid-infrared spectroscopy. J. Dairy Sci. 103:11535–11544. https://doi.org/10.3168/jds.2020-18589.

Lucy, M. C. 2019. Symposium review: Selection for fertility in the modern dairy cow—Current status and future direction for genetic selection. J. Dairy Sci. 102:3706–3721. https://doi.org/10.3168/jds.2018-15544.

Ma, L., J. B. Cole, Y. Da, and P. M. VanRaden. 2019. Symposium review: Genetics, genome-wide association study, and genetic improvement of dairy fertility traits. J. Dairy Sci. 102:3735–3743. https://doi.org/10.3168/jds.2018-15269.

Miglior, F., B. L. Muir, and B. J. Van Doormaal. 2005. Selection indices in Holstein cattle of various countries. J. Dairy Sci. 88:1255–1263. https://doi.org/10.3168/jds.S0022-0302(05)72792-2.

Misztal, I., S. Tsuruta, D. Lourenco, I. Aguilar, A. Legarra, and Z. Vitezica. 2014. Manual for BLUPF90 family of programs. http://ncc.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all2.pdf.

Morton, J. M., J. E. Pryce, and M. Haile-Mariam. 2018. Components of the covariances between reproductive performance traits and milk protein concentration and milk yield in dairy cows. J. Dairy Sci. 101:5227–5239. https://doi.org/10.3168/jds.2017-13268.

Mrode, R. A. 2005. Linear Models for the Prediction of Animal Breeding Values. CABlI Publishing.

Pryce, J. E., R. Woolaston, D. P. Berry, E. Wall, M. Winters, R. Butler, and M. Shaffer. 2014. World trends in dairy cow fertility. Page 17–22 in Proc. 10th World Congress of Genetics Applied to Livestock Production.
van den Berg, I., P. N. Ho, T. D. W. Luke, M. Haile-Mariam, S. Bolormaa, and J. E. Pryce. 2021. The use of milk mid-infrared spectroscopy to improve genomic prediction accuracy of serum biomarkers. J. Dairy Sci. 104:2008–2017. https://doi.org/10.3168/jds.2020-19468.

Workie, Z. W., J. P. Gibson, and J. H. J. van der Werf. 2019. Age at culling and reasons of culling in Australian dairy cows. Pages 143–146 in Proceedings of the 23rd Conference of the Association for the Advancement of Animal Breeding and Genetics (AAABG), Armidale, New South Wales, Australia, AAABG.

Notes

I. van den Berg (https://orcid.org/0000-0002-9292-8636)
P. N. Ho (https://orcid.org/0000-0001-9481-9612)
M. Haile-Mariam (https://orcid.org/0000-0001-5476-7475)
J. E. Pryce (https://orcid.org/0000-0002-1397-1282)

This study was undertaken as part of the DairyBio program, which is jointly funded by Dairy Australia (Melbourne, Australia) and Agriculture Victoria (Melbourne, Australia) and The Gardiner Foundation (Melbourne, Australia).

The authors thank DataGene (Bundoora, Australia) for supplying the data used in this study and the farmers who participated in this project. Hico Pty Ltd. (Maffra, Victoria, Australia) and TasHerd Pty Ltd. (Hadspen, Australia) are gratefully acknowledged for collecting milk samples and providing MIR spectral data. The authors extend their gratitude to the Agriculture Victoria Research molecular genetics team (Bundoora, Australia), particularly Coralie Reich, Brett Mason, and Amanda Chamberlain, for genotyping. Special thanks to Bolormaa Sunduimijid and Iona MacLeod (Agriculture Victoria Research, Bundoora, Australia) for imputation of genotypes. We are also grateful to Ignacy Misztal, Daniela Lourenco, and Shogo Tsuruta from the University of Georgia (Athens, GA) for technical assistance with the analyses.

The authors have not stated any conflicts of interest.