Proc. Assoc. Advmt. Anim. Breed. Genet. 20:25-28

EFFECT OF DAUGHTER MISIDENTIFICATION ON DAIRY SIRE EVALUATION

A.M. Winkelman

LIC, Private Bag 3016, Hamilton, New Zealand

SUMMARY

The impact of daughter misidentification on dairy sire breeding value (BV) estimation was investigated by comparing sire progeny group means of DNA-verified cows to sire progeny group means of cows that had paternity determined through mating records. The daughters' BVs were adjusted for the dam contribution prior to the calculation of the means. BVs for milk volume, fat yield, protein yield, somatic cell score (SCS) and liveweight, and a five-trait index (breeding worth (BW)) containing these traits, were analyzed. Comparisons were done within sire breed (Friesian, Jersey and Friesian Jersey (FJ) Cross).

Estimates of progeny group means of the production traits of DNA-verified daughters were, on average, higher than those of daughters for which paternity had been assigned via mating records. The estimates ranged from 4.2 to 10.7 litres for milk volume, 0.16 to 0.30 kg for fat yield and 0.16 to 0.28 kg for protein yield. The differences between the progeny group means was close to zero for SCS, while the differences ranged from -0.076 to 0.18 kg for liveweight. The differences between the progeny group means of the five-trait BW were less than 2 BW points.

The magnitudes of the effect tended to increase with increased genetic merit. Higher genetic merit sires are likely to have greater bias than lower genetic merit sires. There was, however, considerable sire-to-sire variation in the difference between the progeny group means.

INTRODUCTION

Internationally, estimates of the percentage of cows that are misidentified to sire range from 5% (Ron *et al.*, 1996) to 23% (Gelderman *et al.* 1986). LIC proves their young sires in progeny test herds, referred to as Sire Proving Scheme (SPS) herds, prior to widespread use. In SPS herds, 95% of the cows are mated to young bulls and 5% are mated to proven bulls. Results from DNA paternity verification found that the rate of misidentification in SPS herds was 4.7%, 6.6% and 5.5% in seasons 2005, 2006 and 2007, respectively (assessed using 3602, 4427, 5120 sire-daughters tests, respectively, in the seasons) (Ric Sherlock personal communication). The expectation is that the percentage of misidentification is lower in SPS herds than in the non-SPS herds. Hence, the proportion of misidentified progeny is expected to increase from first to subsequent proofs.

A number of approaches have been used to assess the effect of misidentification of sires on genetic evaluation. Van Vleck (1970) used a deterministic model of the sire-daughter inheritance path to assess the effect of sire misidentification on genetic evaluation and estimates of genetic trends. He found that misidentification resulted in biased genetic evaluations and estimates of genetic trends. The bias increased with an increased proportion of misidentified daughters. Geldermann *et al.* (1986) also used a deterministic model (again considering only paternal pedigree errors) to show that a misidentification rate of 15% decreased accuracy of genetic evaluation, decreased estimates of heritabilities and reduced genetic gain. Estimates of the drop in genetic gain ranged from 8.7% to 16.9% for heritabilities of 0.5 and 0.2, respectively. Losses of similar magnitudes were found using stochastic simulation (Harder *et al.* 2005) and by Banos *et al.* (2001). Misidentification is expected to shrink the scale of the estimated breeding values (BVs). This is because the progeny that were incorrectly assigned to superior sires would more likely be the progeny of sires with a lower genetic merit than the top-end bulls. Hence, the top sires' genetic evaluations would be biased downwards. Similarly, progeny incorrectly assigned to sires of low

genetic merit, would more likely be the progeny of sires with a higher genetic merit than the lower-end bulls, thereby biasing the genetic evaluation of these sires upwards.

DNA verification of paternity involves comparing the DNA markers of an animal to those of its putative sire. LIC first offered a SNP-based DNA sire verification service to customers in the mid 1990s. Later, the service was provided by GeneMark. The test is based on approximately 100 SNPs, where the sire was deemed correct if the concordance between him and his progeny was at least 99%. The question arose as to whether data on cows for which paternity had been DNA-verified could be used to assess the impact of misidentification on the genetic evaluation of sires. The purpose of this study was to determine whether sire genetic evaluations based on cows that had paternity assigned via DNA verification differed to evaluations for which paternity was determined using mating records alone.

MATERIALS AND METHODS

The impact of misidentification on sire evaluation was assessed by comparing the sire contribution to daughters' BVs where paternity had been assigned via DNA verification to those that had been assigned using mating records. A sire's contribution to his daughter's BV can be partially determined by removing the dam's contribution to the BV. This approach does not remove the daughter's own Mendelian sampling (MS) contribution to the BV. However, if the MS is assumed to have a progeny group mean of zero, then averaging the sire's contribution across all his daughters within each of his progeny groups (i.e. DNA-verified and otherwise) should be a means of determining the impact of misidentified daughters on the sire's proof. No difference in the progeny group means would indicate that sire evaluation is not affected by misidentification of progeny. If the mean of the progeny genetic evaluation for the DNA-verified group is higher than that of the group that had paternity determined via mating records, then there is evidence that the misidentification is biasing the sire genetic evaluations downwards.

A total of 680,491 cows DNA-verified to sire were extracted from the national database. Of these, 392,677 had herd test records. These cows were the daughters of 4853 sires. All daughters of these sires were extracted from the national database. A total of 11,892,687 daughters had herd test information. Progeny of Friesian, Jersey and Friesian-Jersey (FJ) cross sires were retained for analysis. Edits were done to ensure that sires had at least five daughters in each progeny group (i.e. paternity assigned via DNA-verification or mating records). BVs for milk volume, fat yield, protein yield, somatic cell score (SCS) and live weight (hereafter referred to as milk, fat, protein, SCS and liveweight) were obtained for the daughters as well as their sires and dams. The BVs did not have Interbull or genomic information incorporated. The final data set included 3452 sires (1847 Friesians, 1159 Jerseys and 446 FJ crosses) with a total of 320,663 DNA-verified daughters and 8,618,574 daughters that had paternity assigned via mating records.

Equation [1] shows the components of a daughter's BV. Equation [2] shows the calculation of the daughter BV adjusted for the dam contribution (BVs_{adj}).

$$daughter\ BV = \frac{1}{2}\ sire\ BV + \frac{1}{2}\ dam\ BV + MS \quad [1]$$

where, MS is the Mendelian sampling; E(MS) = 0.

$$BV_{adj} = daughter\ BV - \frac{1}{2}\ dam\ BV = \frac{1}{2}\ sire\ BV + MS \quad [2]$$

The BVs_{adj} were calculated for all daughters for all 5 traits. Additionally, the adjusted five-trait Breeding Worth index (BW_{adj}) was calculated using the BVs_{adj} and the economic weights of published by NZ's Animal Evaluation Unit (AEU) in February, 2013 (Anonymous, 2013). The BVs_{adj} and BW_{adj}, were averaged over each sire and progeny group. Hence, every sire had two progeny means for each trait – one in which paternity was determined via DNA verification (DNA) and the other in which paternity was determined using mating records (REC)).

The effect of progeny group (DNA versus REC) was estimated using the linear regression of progeny mean on the progeny group and sire BV for each trait. A test of whether the magnitude of the estimate of progeny group was affected by the magnitude of the sire BV was done by including the interaction between the progeny group and sire BV.

RESULTS AND DISCUSSION

Table 1 shows the estimates of the progeny group effect within sire breed for milk, fat, protein, SCS, liveweight and BW. The model was parameterized so that the results are relative to the REC group. Hence, estimates greater than zero indicate that the mean of the DNA progeny group was higher than the mean of the REC group. The estimates were greater than zero for milk, fat and protein and close to zero for SCS. The liveweight mean was greater than zero for the Friesians and the FJ crosses but negative for the Jerseys. The estimates for milk follow the expected trend of being highest for the Friesians, lowest for the Jerseys and intermediate for the FJ crosses. The results for the FJ crosses are not intermediate between the Friesians and Jerseys for fat and protein. Nevertheless, estimates greater than zero are an indication that misidentification to sire biases the sires' BVs downwards.

Table 1. Estimates of effect of progeny group for milk, fat, protein, SCS, liveweight and BW[1]

Sire Breed	N	Milk (l)	Fat (kg)	Protein (kg)	SCS (log(cells/ml))	Liveweight (kg)	BW ($)
Friesian	1847	10.681***	0.271***	0.276***	0.00	0.175***	1.635***
Jersey	1159	4.214***	0.307***	0.191***	-0.001	-0.076	1.965***
FJ Cross	446	6.25***	0.164***	0.164***	0.001	0.142*	0.899***

[1]* = P<0.05, **=P<0.01, ***=P<0.001

Table 2 contains the estimates of the interaction of progeny group and sire BV/BW. The model is parameterized so that the values show the difference between the slopes of the sire BV in the DNA and REC groups. The estimates were small but positive indicating that the difference between the progeny group increases with increasing sire BV. The estimates were not significant for BW.

Table 2. Estimates of the interaction between progeny group and sire BV or BW[1]

Sire Breed	Milk (l)	Fat (kg)	Protein (kg)	SCS (log(counts/ml))	Liveweight (kg)	BW($)
Friesian	0.0125***	0.0070**	0.0045*	0.0079	0.0038***	0.0007
Jersey	0.0127***	0.0071*	0.0032	0.0070*	0.0084*	0.0035
FJ cross	0.0178***	0.0138**	0.0120**	0.0163**	0.0161***	0.0037

[1]* = P<0.05, **=P<0.01, ***=P<0.001

While the estimates of progeny group differences were positive, there was considerable variation in the difference between the DNA and REC means within sire. They ranged from ±200 litres for milk, ±10 kilograms for fat, ±8 kilograms for protein, ±0.4 to 0.4 units for SCS and ±4 kilograms for liveweight. Negative differences occurred for all breeds in both high- and low-BV sires. Such differences could arise from the fact that some sires had very few daughters in the DNA group and thousands in the REC group. The mean MS deviation of a small progeny group could differ markedly from zero. Additionally, there is likely sire-to-sire variation in the

percentage of daughters in the REC group that are misidentified. Nevertheless, the differences in the DNA and REC progeny groups suggest that misidentification does bias the sires' estimated BV.

Harris *et al.* (2007) found an annual per-cow genetic trend in NZ following the introduction of the BW was 2.5 kilograms of milk solids. The weighted (over sire breed) averages of the progeny group effect in Table 1 for fat and protein are 0.269 and 0.233 kilograms, respectively. Summing these gives a total of 0.502 kilograms of milk solids. This value is 20% of the annual genetic gain. The expectation is that top-end sires would be underevaluated for fat and protein by more than 0.5 kilograms. The underevaluation and reduced ability to identify extreme sires would have a negative effect on genetic gain.

The question remains as to what proportion of misidentified daughters could result in a sire's proof being underestimated by the amounts found in this study. Johnson (2010) used a simulation study to determine the effect of sire misidentification on the reproof effect. The reproof effect was found to vary with the percentage of parentage errors in the first and subsequent proofs. When the initial progeny test scheme had a 5% parentage error and the data used for the subsequent proof had 30% parentage error, with 80% of the daughters sired by other graduate bulls and 20% of the daughters sired by bulls with genetic merit equal to that of the cow population, the reproof effect for protein was -0.24 kilograms. An estimate of 0.23 kilograms of protein is a difference in a sire BV of 0.46 kilograms. This value is considerably higher than that found for the reproof effect. Sires evaluated following their initial proof may have in excess of 30% misidentified progeny in the commercial population. The next step of the research will involve estimating within-herd heritabilities for each trait and determining the association between the estimates and the level of sire misidentification as outlined by Dechow *et al.* (2007). Negative correlations suggest that the information on within-herd heritabilities can be used to identify herds that provide inaccurate data for sire evaluation.

CONCLUSIONS

A comparison of progeny group means of daughters that had paternity assigned via DNA verification versus mating records found that estimated BVs are, on average, biased downwards when all progeny are not DNA-verified. There is evidence that the effect increases with increasing sire BV. Higher genetic merit sires are likely to have greater bias than lower genetic merit sires.

REFERENCES

Anonymous (2013). http://www.nzael.co.nz/news/2013/2/2141742705/nzael-releases-revised-economic-values. Accessed 15 May 2013.

Banos G., Wiggans, G.R. and Powell R.L. (2001) *J. Dairy Sci.* **84**:2523.

Dechow C.D., Norman H.D., Zwald N.R., Cowen C.M., Meland O.M. (2007) *J. Dairy Sci.* **91**:1640.

Geldermann H., Pieper U. and Weber W.E. (1986) *J. Anim. Sci.* **63**:1759.

Harris B., Pryce J.E. and Montgomerie, W.A. (2007) *Proc. Assoc. Advmt. Anim. Breed Genet.* **17**:434.

Johnson D.L. (2010) A study of prior proofs and subsequent reproofs of graduate progeny-tested bulls. LIC Internal Report.

Ron M., Blanc Y., Band M., Ezra E. and Weller J.I. (1996) *J. Dairy Sci.* **79**:676.

Van Vleck L. D. (1970) *J. Dairy Sci.* **53**: 1697.