USEFULNESS OF INFORMATION ON MATES OF SIRES IN ARTIFICIAL INSEMINATION

R. H. MILLER AND E. L. CORLEY
United States Department of Agriculture, Animal Husbandry Research Division, ARS
Beltsville, Maryland

ABSTRACT

DHIA data representing 24,853 daughter-dam comparisons for 263 Holstein A.I. sires were studied to determine the usefulness of information on a sire's mates (dams of daughters) in sire evaluation. Although there was a mean bias of 41 lb of milk per sire, due to selection of mates, the relative ranking of the bulls was the same (rank correlation = +0.998) when daughters' deviations alone were used as when both daughter and dam deviations were employed for ranking. If point estimates of a bull's breeding value are desired, failure to employ the mates' deviations will, on the average, result in an overestimate of the sire's merit. Analyses of variance indicated that the bulls' mates were equally as variable as the daughters of the bulls. A heritability estimate for milk yield of 0.24 was obtained by daughter-dam regression.

Until recently the production of a bull's mates was used as an integral part of sire evaluation procedures. Within the past 5 yr attention in sire appraisal has shifted to use of the yield of the contemporary herdmates of the bull's daughters.

The information on the bull's mates was used to accomplish two objectives in the estimation of the bull's breeding value. First, the mates' yield served to correct for any tendencies to use bulls on cow groups differing in average genetic merit. Secondly, the mates' production was employed to correct for the environmental levels of the herds in which the progeny made their records. Dissatisfaction with the daughter-dam procedure grew from the realization that use of the production of the sire's mates only partially achieved the latter objective. Since, on the average, the productive spans of dam and daughter overlap only partially, year-to-year trends in environmental levels are not wholly removed. Also, the use of daughter-dam pairs was not restricted to records made in the same herd.

As the use of artificial insemination grew, need for increased attention to environmental effects on production became more apparent. At the same time it was widely felt that there was no extensive selectivity in mating bulls to different levels of cows. Often the choice of semen was left to the inseminator and in these situations selection was frequently on the basis of conception rate. Where selection did occur on the basis of production information, natural-service progeny data were usually employed. Natural-service progeny performance was found to have a low correlation with a proof based on A.I. progeny (2).

Johansson and Robertson (4) showed that an index including daughters, mates, and the contemporary herd averages of each provides a more accurate assessment of the bull's breeding value than an index which excludes a portion of this information. The relative advantages of the complete index in practice are not clear and, therefore, this procedure has not been widely employed.

VanVleck et al. (6) employed an index containing deviations of mates and progeny evaluate three Holstein sires thought to have been used in a selective manner in New York. Mean superiority of mates over their contemporaries ranged from -285 to +599 lb of milk. They found that breeding value estimates from the complete index differed only slightly from estimates based on daughter deviations alone.

The crucial question relative to the use of the mates' production in sire evaluation is the distribution of the mean selection differential among bulls. If any selection among cows is being practiced in A.I. herds, a positive mean deviation of mates from their contemporary herd average is expected. How this selection affects sire evaluation depends upon whether cows chosen to produce herd replacements are mated at random among possible service sires. If allotment is essentially random, estimates of breeding value based solely on daughter devia-
tions are biased, but the relative standing of the bulls should not change. If cows to produce female replacements are further subdivided according to breeding value and the different groups mated to particular sires, there will be a change in ranking when daughters' and mates' deviations are used, as compared to rank on daughters' deviations alone.

The present investigation was undertaken to determine the average bias introduced when an A.I. sire's mates are not a random sample of cows within the herd, and to assess the importance of this bias in sire evaluations.

**USE OF MATES' INFORMATION IN THE SIRE INDEX**

In the succeeding discussion the following model is assumed to be operative:

\[ Y_{ijk} = \mu + h_i + s_j + e_{ijh} \]

where \( Y_{ijk} \) is the first available 305-2×-M.E. lactation record for the \( k^{th} \) daughter of the sire \( j \) in herd \( i \). Hence, \( h_i \) is an effect common to all animals in the \( i^{th} \) herd, \( s_j \) is a contribution present for all progeny of sire \( j \), and \( e_{ijh} \) is an element peculiar to the individual daughter. The elements \( s_j \) and \( e_{ijh} \) are assumed to be normally and independently distributed with variances \( \sigma_s^2 \) and \( \sigma_e^2 \). Herd-by-sire interaction is assumed to be absent.

Corresponding to each \( Y_{ijk} \) there is an observation \( X_{ijh} \) which refers to the first available lactation record of the cow mated to sire \( j \) to produce the \( k^{th} \) daughter in herd \( i \). Hence, the appropriate model is

\[ X_{ijh} = \mu + h_i + s_j + e_{ijh} \]

VanVleek et al. (6) combined deviations of daughters and mates in the following way:

\[ I_i - \mu = b_1 d_{ijh} + b_2 c_{ijh} \]

where \( d_{ijh} \) is the mean deviation of the daughters, \( c_{ijh} \) is the mean deviation of the sire's mates, and \( b_1 \) and \( b_2 \) are the weights for the mean deviation of the sire's mates, and \( d_{ijh} \) and \( c_{ijh} \) are the average deviations of daughters and mates.

The values used by VanVleek et al. (6) for \( b_1 \) and \( b_2 \) were as follows:

\[ b_1 = \frac{nh^2}{(n - 1 - h^2)h^2 + 4} \]
\[ b_2 = \frac{nh^2}{2[(n - 1 - h^2)h^2 + 4]} \]

where \( n \) is the number of progeny and \( h^2 \) is the heritability of lactation yield. These index weights assume that environmental correlations are absent.

If \( d_{ijh} \) represents the contemporary herdmate deviation of the \( k^{th} \) daughter of sire \( j \), it can be shown that the variance is

\[ \frac{(m_{ijh} + 1)}{m_{ijh}} \sigma_s^2 + \sigma_e^2 \]

where \( m_{ijh} \) is the number of contemporaries. Thus, the variability of the individual deviations is not uniform. In lieu of \( d_{ijh} \), the weighted deviation

\[ d^w_{ijh} = \sqrt{\frac{m_{ijh}}{m_{ijh} + 1}} \]

\[ d_{ijh} = w_{ijh} d_{ijh} \]

can be employed. In like manner the deviation \( c_{ijh} \)

\[ c_{ijh} = v_{ijh} c_{ijh} \]

where \( P_{ijh} \) is the number of contemporaries of the mate of sire \( j \) corresponding to daughter \( k \).

Using the above notation an alternative index is:

\[ I_i - \mu = b_1' d_{ijh} + b_2' c_{ijh} \]

where

\[ b_1' = \frac{\sum d^w_{ijh} \cdot e_{ijh}}{\sum d^w_{ijh}} \]
\[ b_2' = \frac{\sum c^w_{ijh} \cdot e_{ijh}}{\sum c^w_{ijh}} \]

It can be shown that the solutions for the \( b' \) values are as follows:

\[ b_1' = \frac{2h^2}{[h^2 + \frac{\sum w^2}{\sum w^2} (4 - h^2)]} - \frac{h^2 (\sum w^2)}{\sum w^2} \]
\[ b_2' = \frac{\sum w^2}{\sum w^2} (4 - h^2) \]

It should be pointed out that the estimates \( I_i \) from the above equation are not on the same scale as those of VanVleek et al. (6). The \( I_i \) values are comparable to those obtained by doubling the \( I \) estimates.

As an example of the use of the above index, the regression coefficients \( b_1' \) and \( b_2' \) were computed for a sire with 50 progeny. Six different arrangements of herdmate numbers of daughters and dams were used. These results are tabulated in Table 1.

Table 1 indicates that the weights for the individual deviations vary considerably, but that the over-all regression coefficients for a sire are not altered by varying numbers of herdmates.

\[ \sigma_s^2 = \Sigma w^2, \sigma_e^2 = \Sigma w^2, \sigma_{ijh}^2 = \Sigma w^2, \sigma_{ijh}^2 = \Sigma w^2, v_{ijh}^2 = \Sigma w^2, \]

* (The summation subscripts have been omitted; it is understood that \( \Sigma w^2 = \Sigma w^2, \Sigma w^2 = \Sigma w^2, v_{ijh}^2 = \Sigma w^2, \text{ etc.} \)
TABLE 1

Effect of number of herdmates on weights for daughters and sire index regression coefficients

<table>
<thead>
<tr>
<th>Avg no. herdmates</th>
<th>Average weights</th>
<th>$b'_1$</th>
<th>$b'_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daughters Dams</td>
<td>Daughters Dams</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15.1 3.0</td>
<td>0.93</td>
<td>0.84</td>
<td>1.542</td>
</tr>
<tr>
<td>15.1 20.0</td>
<td>0.93</td>
<td>0.97</td>
<td>1.542</td>
</tr>
<tr>
<td>2.2 10.4</td>
<td>0.81</td>
<td>0.86</td>
<td>1.543</td>
</tr>
<tr>
<td>8.4 8.8</td>
<td>0.91</td>
<td>0.90</td>
<td>1.542</td>
</tr>
<tr>
<td>15.0 10.4</td>
<td>0.97</td>
<td>0.86</td>
<td>1.544</td>
</tr>
<tr>
<td>15.1 15.0</td>
<td>0.93</td>
<td>0.97</td>
<td>1.542</td>
</tr>
</tbody>
</table>

DATA

The data employed in the present study were lactation records of Holstein-Friesian cows whose sires were summarized in the National Dairy Herd Improvement Program in 1961. The first available records of all progeny of specified A.I. sires were abstracted from the files, along with the contemporary herdmate average for each record. Sires were chosen which had a relatively large number of daughters, to insure obtaining a usable number of daughter-dam pairs.

From the same files the first available lactation record with herdmate data for dams of the progeny of these sires was selected. When a dam had more than one daughter, all such daughters were used, regardless of whether the daughters were all by the same sire or by different sires.

All daughters were the result of artificial service. Both natural-service and artificially sired mates were used, although the proportion of artificially sired animals was higher than would normally be true, due to the nature of the basic file employed.

Initially, 263 A.I. sires with a total of 70,467 progeny were selected. From these, a total of 24,853 daughter-dam pairs representing 245 sires was obtained. The mean milk yields of daughters and mates were 12,560 and 12,580 lb, respectively. Distributions of the data by years and area are given in Tables 2 and 3, respectively.

TABLE 2

Distribution of records by year

<table>
<thead>
<tr>
<th>Year</th>
<th>Daughters</th>
<th>Dams</th>
</tr>
</thead>
<tbody>
<tr>
<td>1946</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>1947</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>1948</td>
<td>0</td>
<td>56</td>
</tr>
<tr>
<td>1949</td>
<td>0</td>
<td>176</td>
</tr>
<tr>
<td>1950</td>
<td>4</td>
<td>494</td>
</tr>
<tr>
<td>1951</td>
<td>19</td>
<td>1,067</td>
</tr>
<tr>
<td>1952</td>
<td>40</td>
<td>2,068</td>
</tr>
<tr>
<td>1953</td>
<td>211</td>
<td>3,393</td>
</tr>
<tr>
<td>1954</td>
<td>579</td>
<td>3,642</td>
</tr>
<tr>
<td>1955</td>
<td>1,068</td>
<td>3,107</td>
</tr>
<tr>
<td>1956</td>
<td>1,230</td>
<td>2,118</td>
</tr>
<tr>
<td>1957</td>
<td>2,634</td>
<td>2,250</td>
</tr>
<tr>
<td>1958</td>
<td>6,679</td>
<td>3,134</td>
</tr>
<tr>
<td>1959</td>
<td>10,954</td>
<td>3,151</td>
</tr>
<tr>
<td>1960</td>
<td>1,435</td>
<td>176</td>
</tr>
</tbody>
</table>

*Regional definitions:
1—Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut
2—New York, New Jersey, Pennsylvania
3—Ohio, Indiana, Illinois, Michigan, Wisconsin
4—Minnesota, Iowa, Missouri, North Dakota, South Dakota, Nebraska, Kansas
5—Delaware, Maryland, Virginia, West Virginia, North Carolina, South Carolina, Georgia, Florida
6—Kentucky, Alabama, Mississippi
7—Arkansas, Oklahoma, Texas
8—Montana, Idaho, Wyoming, Colorado, Utah
9—Washington, Oregon, California

RESULTS

The first phase of the analysis dealt with the determination of the effect of use of information on mates on the relative rank of the sires. Data from a total of 207 sires with ten or more daughter-dam pairs were used to study rank correlation between a bull's standing on daughter deviations alone as compared to rank on the basis of both deviations of daughters and mates.
The first comparison was between rank with the index

\[ I_s - \mu_o = b \bar{d}_r, \]

versus rank on the index

\[ I - \mu_o = b \bar{d}_r + b_c \bar{c}_r, \]

where

\[ b = \frac{nh^2}{(n-1)h^2 + 4} \]

and the remaining statistics are as previously described. A heritability value for milk yield of 0.25 was assumed.

The rank correlation obtained between relative standing of the bulls on the basis of the two indices was 0.998. This was despite the fact that many relatively large average deviations of mates were encountered, the range being +1,600 to −950 lb. However, as developed in the previous section, the bull’s proof is biased only by a fraction \((b_c)\) of the mean deviation of his mates. A survey of the bias values \((-b_c \bar{c}_r)\) showed that the error was rather consistent from sire to sire, with a mean value of +41 lb per bull. Table 4 shows the distribution of the bias introduced by selection of mates of these sires.

The index values \(I_o\) and \(I\) for the 207 sires are plotted in a scatter diagram in Figure 1. It is clear from Figure 1 that a straight line with a slope of 1.0 will fit these points very closely.

As a further illustration of the data, Table 5 presents averages for the three bulls with the largest positive deviations of mates and the three bulls which had the largest negative deviations. For each sire, Table 5 gives the number of progeny, the average unadjusted daughter deviation, the average unadjusted deviation of mates, the bias in the estimate of breeding value based on progeny, the index values for \(I_o\) and \(I\), and the relative ranks based on each of the indices. These examples demonstrate that large gross deviations of mates do not appreciably alter the comparison ratings of the sires.

As an indication of the most extreme situation possible in these data, the 14 sires which showed the greatest increase in rank using index \(I\) and the 14 sires which showed the greatest decrease in rank with index \(I\) were compared. Within this group of 28 bulls, the rank correlation between \(I_o\) and \(I\) was 0.925, which may be an indication of the lower limit obtainable in extreme situations.

No values for the index \(I\) were computed. However, the weighted deviations were used in the index

\[ \text{TABLE 4} \]

<table>
<thead>
<tr>
<th>Bias (lb)</th>
<th>No. of sires</th>
</tr>
</thead>
<tbody>
<tr>
<td>−20 or less</td>
<td>3</td>
</tr>
<tr>
<td>0− −20</td>
<td>17</td>
</tr>
<tr>
<td>0− +20</td>
<td>30</td>
</tr>
<tr>
<td>+20− +40</td>
<td>47</td>
</tr>
<tr>
<td>+40− +60</td>
<td>59</td>
</tr>
<tr>
<td>+60− +80</td>
<td>27</td>
</tr>
<tr>
<td>+80−100</td>
<td>17</td>
</tr>
<tr>
<td>+100 or more</td>
<td>7</td>
</tr>
</tbody>
</table>

distribution of the bias introduced by selection of mates of these sires.

The index values \(I_o\) and \(I\) for the 207 sires are plotted in a scatter diagram in Figure 1. It is clear from Figure 1 that a straight line with a slope of 1.0 will fit these points very closely.

As a further illustration of the data, Table 5 presents averages for the three bulls with the largest positive deviations of mates and the three bulls which had the largest negative deviations. For each sire, Table 5 gives the number of progeny, the average unadjusted daughter deviation, the average unadjusted deviation of mates, the bias in the estimate of breeding value based on progeny, the index values for \(I_o\) and \(I\), and the relative ranks based on each of the indices. These examples demonstrate that large gross deviations of mates do not appreciably alter the comparison ratings of the sires.

As an indication of the most extreme situation possible in these data, the 14 sires which showed the greatest increase in rank using index \(I\) and the 14 sires which showed the greatest decrease in rank with index \(I\) were compared. Within this group of 28 bulls, the rank correlation between \(I_o\) and \(I\) was 0.925, which may be an indication of the lower limit obtainable in extreme situations.

No values for the index \(I\) were computed. However, the weighted deviations were used in the index

\[ \text{TABLE 5} \]

<table>
<thead>
<tr>
<th>No. daughters</th>
<th>Unadjusted daughter deviation</th>
<th>Unadjusted mates’ deviation</th>
<th>Adj. mates’ deviation ((-b_c \bar{c}_r))</th>
<th>(I_o)</th>
<th>(I)</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>18</td>
<td>−380</td>
<td>1,600</td>
<td>112</td>
<td>−209</td>
<td>165</td>
</tr>
<tr>
<td>B</td>
<td>20</td>
<td>−400</td>
<td>1,590</td>
<td>111</td>
<td>−228</td>
<td>168</td>
</tr>
<tr>
<td>C</td>
<td>12</td>
<td>140</td>
<td>1,450</td>
<td>87</td>
<td>62</td>
<td>118</td>
</tr>
<tr>
<td>D</td>
<td>25</td>
<td>−140</td>
<td>−950</td>
<td>−76</td>
<td>−88</td>
<td>147</td>
</tr>
<tr>
<td>E</td>
<td>10</td>
<td>1,950</td>
<td>−350</td>
<td>−18</td>
<td>792</td>
<td>810</td>
</tr>
<tr>
<td>F</td>
<td>63</td>
<td>−220</td>
<td>−270</td>
<td>−27</td>
<td>−178</td>
<td>162</td>
</tr>
</tbody>
</table>

\[ \text{FIG. 1. Indexes of breeding value based on daughter herdmate and dam-herdmate deviations.} \]
\[ I' - \mu = b_1 \tilde{d} + b_2 \tilde{e} \]

i.e., the regression coefficients appropriate to the index \( I \) were used. The standard of comparison using daughter deviations alone was

\[ I'_d = b \tilde{d} \]

The rank correlation between \( I \) and \( I'_d \) was found to be 0.997, as compared to 0.998 using the unweighted deviations.

The second portion of the analysis was a comparison of variability of the mates and their daughters. If cows to produce herd replacements were highly selected, then the variability among them should be less than that of their unselected progeny. Harvey and Lush (3), using Jersey ROM data, found that cows with tested progeny were equally as variable in fat yield as their herd contemporaries, and that the dams were actually more variable than their daughters. This was attributed to an absence of intense selection for fat yield.

Bereskin (1), in an analysis of Holstein-Friesian DHIA herds in Iowa, found little difference in milk yield variability between daughters and dams. He concluded that the dams were not highly selected.

Data used in the first portion of the analysis were analyzed to determine if there were differences in the level of variation of daughters and dams. Estimates of daughter-dam covariance were also obtained. Results are summarized in Table 6.

There were 13,921 pairs where both daughter and dam made their records in the same herd. Intra-herd-year differences were not removed, although calving dates in the data employed ranged from 1946-1960. The data were, however, largely concentrated in the period 1951-1960, as shown in Table 2. Harvey and Lush (3) pointed out that year differences within herds are small and also are present in the within-herd covariance only to a small extent. In the present data calving dates occurred in the same year for only 1,829 pairs out of 13,921.

From Table 6 it can be seen that the dams' production is only slightly less variable than that of the daughters. The within-herd standard deviation of daughters is 2,071 lb as compared to 2,047 lb for dams.

The intra-herd regression of daughters on dams is computed from Table 6 as 5,028/41,906, or 0.12. Doubling this value gives a heritability coefficient of 0.24. Bereskin (1) obtained values of daughter-dam regressions for milk yield of 0.124 and 0.128 when intra-herd-year-season variation was ignored and when correction for this variation was made, respectively. These regressions were based on averaged records of all lactations. Converted to a single record basis, the corresponding regressions were 0.105 and 0.110, respectively, as compared to 0.12 in the present analysis.

Harvey and Lush (3) obtained a daughter-dam regression value of 0.121 for Jersey fat yield, when intra-herd-year variation was removed by correction. When the yearly variation was ignored, the estimate of the intra-herd regression coefficient was 0.119. These regressions were derived from relationships among average records of cows.

**CONCLUSIONS**

The rank correlation values obtained in this study indicate that inclusion of the production of a sire's mates in the index of breeding value does not alter the relative standing of the individual bulls. These results show that there is no tendency for certain A.I. sires to be mated to groups of cows deviating markedly from the average merit of the population of cows used to produce herd replacements. Selection differentials for mates were such that estimates of breeding value were biased by + 41 lb of milk when yield of mates was ignored.

The selection practiced in choosing cows to produce herd replacements did not result in marked differences in variability between the progeny and the bulls' mates. Intraherd standard deviations in milk yield were 2,071 and 2,047 lb for progeny and mates, respectively.

Results of the present study indicate that use of information on an A.I. sire's mates does not increase the accuracy of ranking. If point estimates of a bull's breeding value are desired,

**TABLE 6**

<table>
<thead>
<tr>
<th>Source</th>
<th>d.f.</th>
<th>Daughters</th>
<th>Dams</th>
<th>Daughters × Dams</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among herds</td>
<td>2,795</td>
<td>26,599</td>
<td>26,391</td>
<td>15,364</td>
</tr>
<tr>
<td>Within herds</td>
<td>11,125</td>
<td>42,968</td>
<td>41,906</td>
<td>5,028</td>
</tr>
<tr>
<td>Standard deviation (lb)</td>
<td></td>
<td>2,071</td>
<td>2,047</td>
<td></td>
</tr>
</tbody>
</table>

* Data coded to 10-lb units.
failure to employ the mates' deviations will, on the average, result in an overestimate of the sire's merit.

ACKNOWLEDGMENTS

Appreciation is expressed to P. W. Keyes and A. H. Gardiner for programming assistance.

REFERENCES


