Prediction of Transmitting Abilities for Holstein Type Traits

P. M. VANRADEN[1] and E. L. JENSEN
Department of Dairy Science
University of Wisconsin
Madison 53706

T. J. LAWLOR and D. A. FUNK[2]
Holstein Association
Brattleboro, VT 05301

ABSTRACT

Heritabilities and genetic and phenotypic correlations among 14 linear type traits were estimated from Holstein Association data by multiple trait REML. Data used for parameter estimation were records of 779,391 daughters of 871 sires included in the January 1988 sire evaluation. Each daughter was represented by her appraisal closest to 30 mo of age. Highest heritability was .37 for stature, and lowest was .10 for foot angle. Gains in reliability from using correlated traits in multiple trait prediction were large for some traits (up to 60% for foot angle for cows). Final score variance parameters were estimated from 953,596 records, which were 43% of records included in the national sire evaluation. Sire models that adjusted or did not adjust for merit of mates were compared. Heritability of final score was .27 with adjustment for merit of mates by subtraction of predicted transmitting ability of dam from daughter's record compared with .29 if mate was ignored. Evaluations for type for several popular older sires were reduced moderately by adjustment for merit of mates, but estimated genetic trend increased slightly. An improved genetic grouping procedure that considers group effects as inherited was adapted for use in sire models. Parameter estimates and models presented were implemented by the Holstein Association for computing July 1988 genetic evaluations for linear traits and final score.

(**Key words**: variance estimation, genetic evaluation, type traits)

INTRODUCTION

Profitability in dairy cattle depends on milk production and on several other yield and non-yield traits. Since 1983, the Holstein Association has collected data on 14 conformation traits scored visually (9). A final type score also is assigned as a composite of the traits observed. Genetic evaluations are published semi-annually for linear traits, final score, and an overall index of production and type that combines evaluation for fat yield, protein yield, and final score. Starting with July 1989 evaluations, information from linear traits also is included in the index of overall merit.

Animals should be ranked more accurately for overall merit if linear trait evaluations rather than final score evaluations are included in the overall index because 1) individual traits can be analyzed with their own heritabilities instead of being combined and then analyzed with an average heritability, 2) genetic and phenotypic correlations among individual traits can be included in the evaluation to improve accuracy, and 3) if economic values change or if the current economic function is inaccurate (12), a new economic function can be applied directly to individual trait evaluations to rerank animals. The final score system does not offer such flexibility.

To improve evaluations of all type traits, we investigated 1) (co)variance estimates by multiple trait REML for use in national genetic evaluation programs, 2) increased accuracy attained by using these covariances in multiple trait rather than single trait evaluation of linear traits, 3) impact of correction for merit of mates on estimates of final score variance parameters

VANRADEN ET AL.

TABLE 1. Summary of data for analysis.

	Linear traits	Final score
Records available	1,241,310	2,229,524
Sires with at least 1 daughter	54,053	115,559
Young sires born in 1980	5,872	5,601[1]
Young sires with the most progeny[2]	400	425
Older sires providing the most contemporaries	400	425
Bulls added for relationship ties	71	76
Sires used in variance estimation	871	926
Appraisals used in variance estimation	799,391	953,596
Appraisals with predicted transmitting ability of dam	0	698,659
Herd classes	58,191	129,419

[1]Fewer young sires for final score because of age restriction of 42 mo.

[2]Progeny required: 24 for linear traits, 22 for final score.

and on predictions of sire genetic merit, and 4) an improved genetic grouping procedure for sire models.

MATERIALS AND METHODS

Data

Data available were 1,241,310 linear trait records and 2,229,524 final score records used by the Holstein Association to evaluate sires in January 1988. Each cow was represented by her record closest to 30 mo of age. Final scores of cows scored after 42 mo of age were discarded. Linear scores for cows older than 42 mo were retained because of the more recent introduction of this program and large number of cows that had no opportunity to be scored at a younger age. Records were adjusted for age and stage of lactation using unpublished Holstein Association factors that since have been updated (2).

Number of sires represented by at least one daughter was 54,053 for linear traits and 115,559 for final score. A subset of sires was chosen so that variance components could be estimated by REML. The subset was chosen to minimize bias due to selection while retaining as large a fraction of total records as possible. This subset contained 1) young sires born in 1980 that could have been selected on pedigree but not progeny information, 2) older sires that provided the most contemporaries for daughters of young sires, and 3) ancestors that provided relationship ties among the bulls already chosen. Progeny required for young sires was 24 for linear traits and 22 for final score. Numbers of bulls in these groups are in Table 1. Number

of records for variance estimation was 779,391 for linear traits (64% of total data available) and 953,596 for final score (43% of data).

Estimation of (Co)Variance Components

Variances and covariances were estimated from the model:

$$y_{ijklm} = h_{im} + g_{jm} + s_{jkm} + e_{ijklm} \quad [1]$$

where y_{ijklm} is a particular observation for trait m, h_{im} is effect of herd class i for trait m, g_{jm} is effect of genetic group j for trait m, s_{jkm} is effect of sire k nested within genetic group j for trait m, and e_{ijklm} is a random residual. Herd classes were groups of cows scored closest to 30 mo of age during the same classifier visit to a herd.

Sire and error effects were treated as random with variances $S \otimes A$ and $E \otimes I$, respectively, where S and E are covariance matrices among the 14 traits for sire and error effects, A is the matrix of additive genetic relationships among sires included in the analysis, I is the identity matrix, and $\otimes$ denotes a Kronecker product. Mixed model equations were derived as in Foulley et al. (1) and Henderson (3). Variance and covariances were estimated by multiple trait REML using procedures similar to those of Jensen and Mao (4) and Klei et al. (5). The program simultaneously diagonalizes sire coefficient and relationship matrices before beginning iteration and diagonalizes S and E at each iteration (11). Estimates of S and E are guaranteed to be positive definite.

Use of Covariances in Evaluations

Multiple trait genetic evaluations that incorporate the estimated genetic and environmental correlations should rank animals more accurately than would single trait evaluations. Accuracies of multiple trait evaluations are higher, because data from correlated traits provide information not contained in data of the trait being evaluated. If all traits are measured on each animal, multiple trait evaluations can be computed by transforming data to canonical (uncorrelated) traits, evaluating canonical traits by single trait methods, and then back-transforming these solutions (1).

Let u represent the 14 transmitting abilities of a particular sire, u* represent transmitting abilities for canonical traits of the same sire, and P^{-1} be the canonical transformation matrix. Then $u = Pu^*$ and $u^* = P^{-1}u$ (1). A similar correspondence exists between reliabilities (squared correlations of predicted with true transmitting ability) on original and canonical scale. Let r and r* represent reliabilities of a particular sire on the original and transformed scales, respectively. Because variance of predictors on the canonical scale are diagonal, r and r* are simple linear functions of each other. Let s_{ii} refer to diagonal i of S and s_{jj}^* refer to diagonal j of S*, where S* is the sire variance matrix for canonical traits. If T is defined as a matrix with individual elements $t_{ij} = p_{ij}^2 \, s_{jj}^* / s_{ii}$, then $r = Tr^*$, and $r^* = T^{-1}r$.

These formulas allow transmitting abilities and reliabilities by multiple trait procedures to be obtained easily. Cow evaluations were computed by transforming data and evaluations of sire and dam to canonical scale and then using procedures similar to those of the Modified Contemporary Comparison (8). Animal model evaluation of these data was not feasible because of memory limitations of the computer available. Predicted transmitting abilities and reliabilities on canonical scale then were back-transformed to original scale. Reliabilities from single trait and multiple trait analyses were compared for cows with one appraisal, many herdmates, and no information from relatives. Single trait reliabilities equal heritabilities for such cows.

Gains in reliability also were calculated for bulls with 20, 50, or 1000 daughters. Traits that provided the most information for evaluation of other traits were determined by computing all other traits were determined by computing all

possible pairwise multiple trait reliabilities for an example cow.

Adjustment for Merit of Mates

Data for final score were analyzed with and without correction for merit of mates. Correction for merit of mates was by subtracting from each record the predicted transmitting ability of the dam. Norman et al. (7) discarded records that did not contain dam evaluations. In our study, unevaluated dams were assigned a value equal to the average of dams in that herd class that were evaluated, or if no dams in a herd class were evaluated, 0 was assigned to all. This enforced an assumption of random mating if information was missing, which was the case for 27% of records, and allowed these records to contribute to sire evaluation. Adjustment for dam's final score evaluation may not completely account for nonrandom mating for individual linear traits with nonlinear economic values.

Heritabilities were estimated as $4\sigma_s^2 / (\sigma_s^2 + \sigma_e^2)$ where σ_s^2 is sire variance and σ_e^2 is error variance. This formula assumes that σ_s^2 is one-fourth of additive genetic variance and σ_e^2 is environmental variance plus three-fourths of additive genetic variance. Subtraction of dam's evaluation removes less than one-fourth of additive genetic variance from σ_e^2, and resulting heritability estimates should be biased upward slightly.

Improved Grouping Procedure

Genetic merit of all sires in the population were predicted from a model that included an improved genetic grouping procedure and an interaction of sire with herd. For variance estimation, selected sires were nested within groups based on birth year as by Henderson (3). For routine evaluations, the grouping procedure proposed by Thompson (10) was used. Ratios of error to interaction variance were not estimated in this study but a value of 4 was assigned for all traits.

The model for evaluation then was:

$$y_{ijklm} = h_{ijm} + c_{ikm} + \Sigma q_{kr}g_{rm} + s_{km} + e_{ijklm}$$

where y_{ijklm} is the observation for trait m, h_{ijm}

is herd class effect of classification j within herd i for trait m, c_{ikm} is interaction effect between herd i and sire k for trait m, q_{kr} relates sire k to unknown ancestor group r (r = 1,...,n), g_{rm} is effect of unknown ancestor group r for trait m, s_{km} is effect of sire k for trait m, and e_{ijklm} is random residual. For each sire, Σq_{kr} = 1 as in Westell et al. (13) because all ancestor paths eventually trace back to unknown ancestors.

Westell et al. (13) used this grouping procedure for an animal model and showed that group equations can be constructed jointly with the inverse of the relationship matrix (A^{-1}). Algebra identical to that of Westell et al. (13) can be used to incorporate Thompson's grouping procedure (10) for a sire model. Coefficients of the following 4 × 4 matrix are added into either A^{-1} or into appropriate group equations for each bull evaluated:

$$
\begin{array}{c}
\begin{array}{cccc} \text{Bull} & \text{Sire} & \text{MGS} & \text{MGD} \end{array} \\
\begin{array}{c} \text{Bull} \\ \text{Sire} \\ \\ \text{MGS} \\ \text{MGD} \end{array}
\left[
\begin{array}{cccc}
1 & -.5 & -.25 & -.25 \\
-.5 & .25 & .125 & .125 \\
& & & \\
-.25 & .125 & .0625 & .0625 \\
-.25 & .125 & .0625 & .0625
\end{array}
\right] (1/p)
\end{array}
$$

where MGS is maternal grandsire of the bull, MGD refers to a group of unknown maternal grandams, and p is proportion of the bull's genetic variance not attributed to known ancestors. Values of p are 11/16 if sire and MGS are both known, 3/4 if only sire is known, 15/16 if only MGS is known, and 1 if neither is known.

Coefficients for sire or MGS are assigned to group equations if sire or MGS are unknown, and MGD coefficients always are assigned to group equations. Each bull's pedigree information is then 1/2 of sire effect (including fixed and random portions) plus 1/4 of MGS effect (including fixed and random portions) plus 1/4 of fixed MGD group effect. The nested grouping procedure of model [1] assumes that only random portions of sire and MGS effects are inherited, which is not reasonable biologically.

RESULTS

Estimation of (Co)Variance Components

Means, within-herd phenotypic SD, genetic SD, and heritability estimates for linear traits are in Table 2. For all traits, SE of heritabilities were approximately .01. Highest heritability was .37 for stature, and lowest was .10 for foot angle. Estimates of phenotypic and genetic correlations among traits are in Table 3. Heritabilities did not differ greatly from earlier estimates calculated from smaller data sets by Henderson's method 3 (6, 9) or by multiple trait REML (5). Nearly all phenotypic correlations were within .05 of those reported by Lawstuen et al. (6). Genetic correlations were generally within .2 of those reported by Klei et al. (5) and Lawstuen et al. (6); however, a few genetic correlations differed by more than .4. Examples are genetic correlations of udder depth with rump angle, which was estimated as –.13 in this study but as .35 in the study by Lawstuen et al. (6), and dairy form with foot angle, which was

TABLE 2. Means, SD, and heritabilities of linear traits.

Trait	Mean	SD		Heritability
		Phenotypic	Genetic	
Stature	31.6	7.2	4.4	.37
Strength	29.8	6.5	3.3	.26
Body depth	31.7	6.4	3.6	.32
Dairy form	28.8	6.8	3.3	.23
Rump angle	24.8	5.0	2.7	.29
Thurl width	27.6	6.2	3.0	.24
Rear leg set	28.5	6.3	2.5	.16
Foot angle	23.6	6.0	1.9	.10
Fore attachment	25.4	6.5	2.8	.18
Rear udder height	24.1	6.7	2.8	.18
Rear udder width	23.6	6.6	2.6	.16
Udder cleft	28.9	5.6	2.2	.15
Udder depth	24.9	4.5	2.2	.25
Teat placement	26.7	5.8	2.7	.21

estimated as –.10 in this study but as .42 in the study by Klei et al. (5).

Convergence to three decimal places for all heritabilities and correlations was achieved in only three rounds of iteration, probably because of the large number of daughters per sire. Less than 3 h of central processing unit time on an IBM 3081 was required for variance estimation, which was divided about equally between absorption of herd class equations and diagonalization of coefficient matrices.

Multiple Trait Prediction

Use of multiple trait mixed models to predict individual transmitting abilities allows data from all correlated traits to contribute to predictions of a given trait. Advantages of multiple trait prediction are greatest if traits of low or moderate heritability are strongly correlated to traits of higher heritability and if phenotypic correlations differ greatly from genetic correlations. Two genetic correlations were above .9, which indicates that the traits involved measure nearly the same genes. However, phenotypic correlations were lower than genetic correlations, which indicates that some information is gained by scoring both.

Comparisons of reliability from single trait and multiple trait prediction are in Tables 4 and 5. Substantial gains occurred for cows (Table 4) for predictions of foot angle (60%), udder traits (10 to 44%), strength (36%), and thurl width (29%). Advantages of multiple trait over single trait prediction were largest for cows and decreased for sires as number of daughters increased. Advantages could be overstated if true parameters differ from estimates.

Adjustment for Merit of Mates

Table 6 gives estimates of variance components and heritabilities for final score with and without adjustment for mate genetic merit. Heritability of final score was .27 with adjustment for merit of mates compared with .29 if mate's merit was ignored. Adjustment for merit of mates removed from σ_e^2 the variation due to predicted transmitting abilities of dams. Expected and actual declines in σ_e^2 were 2 and 1.5%, respectively. If sires were mated randomly to cows within herd classes, σ_s^2 was expected to be unchanged by adjustment for

TABLE 3. Genetic (above diagonal) and phenotypic (below diagonal) correlations among linear type traits.

Trait	Stature	Strength	Body depth	Dairy form	Rump angle	Thurl width	Rear leg set	Foot angle	Fore attachment	Rear udder height	Rear udder width	Udder cleft	Udder depth	Teat placement
Stature	1.00	.75	.82	.28	-.06	.64	-.07	.37	.30	.26	.31	.23	.32	.22
Strength	.56	1.00	.93	-.09	-.11	.72	-.19	.46	.25	.23	.36	.16	.06	.17
Body depth	.64	.75	1.00	.19	-.13	.71	-.10	.37	.20	.21	.33	.17	.00	.20
Dairy form	.15	-.12	.10	1.00	-.03	.08	.27	-.10	-.06	.12	.15	.18	-.13	.15
Rump angle	.06	.00	.01	-.01	1.00	-.24	-.10	-.11	-.21	-.23	-.19	-.16	-.13	-.12
Thurl width	.40	.44	.44	.03	-.05	1.00	-.04	.34	.27	.27	.40	.13	.10	.16
Rear leg set	-.05	-.11	-.06	.14	-.02	-.03	1.00	-.43	-.11	-.18	-.16	-.02	-.06	.00
Foot angle	.14	.19	.16	-.02	-.05	.15	-.18	1.00	.25	.21	.26	.20	.16	.19
Fore attachment	.12	.15	.12	.03	-.10	.12	-.05	.15	1.00	.59	.56	.44	.78	.58
Rear udder height	.13	.16	.16	.13	-.10	.16	-.07	.15	.43	1.00	.91	.39	.42	.31
Rear udder width	.17	.23	.22	.12	-.07	.23	-.08	.16	.39	.70	1.00	.40	.30	.39
Udder cleft	.05	.04	.05	.13	-.07	.06	.01	.08	.30	.30	.28	1.00	.43	.73
Udder depth	.12	.00	-.04	-.06	-.06	.03	-.03	.08	.42	.20	.13	.31	1.00	.40
Teat placement	.06	.06	.07	.09	-.06	.06	.01	.08	.40	.26	.26	.46	.32	1.00

VANRADEN ET AL.

TABLE 4. Single trait and multiple trait reliabilities for cows scored once, with many herdmates, and with no information contributed by relatives.

Trait	Reliability[1]		Increase in reliability	Trait providing most information
	Single	Multiple		
			(%)	
Stature	.37	.39	5	Body depth
Strength	.26	.34	36	Body depth
Body depth	.32	.36	13	Dairy form
Dairy form	.23	.25	9	Teat placement
Rump angle	.29	.30	3	Thurl width
Thurl width	.24	.31	29	Body depth
Rear leg set	.16	.17	6	Foot angle
Foot angle	.10	.16	60	Strength
Fore attachment	.18	.26	44	Udder depth
Rear udder height	.18	.22	22	Rear udder width
Rear udder width	.16	.21	31	Rear udder height
Udder cleft	.15	.20	33	Teat placement
Udder depth	.25	.29	16	Fore attachment
Teat placement	.21	.23	10	Dairy form

[1]Single trait reliability = heritability.

merit of mates. Actual σ_s^2 declined by 9%, indicating positive assortative mating within herd class.

Adjustments for merit of mates of individual bulls were computed from the entire data set (2,229,524 records). Largest adjustments were among popular, older bulls with extremely high evaluations. Evaluations of such bulls tended to decline; largest individual decline was about one-third of a genetic SD. However, among currently available bulls, changes were not great. Rank correlation between adjusted and unadjusted evaluations of these 1396 sires was .996. Estimated genetic trend was slightly higher with adjustment for merit of mates than without adjustment. Linear trait variances and covariances were estimated without adjustment for mates; however, published evaluations now include this adjustment.

Improved Grouping Procedure

Solutions for inherited group effects were examined and were monitored across iterations. Unknown MGD groups were defined separately

TABLE 5. Single trait and multiple trait reliabilities of sires with differing number of daughters.

Trait	20 daughters		50 daughters		1000 daughters	
	Single	Multiple	Single	Multiple	Single	Multiple
Stature	.72	.73	.86	.87	.99	.99
Strength	.63	.68	.81	.83	.99	.99
Body depth	.68	.70	.84	.85	.99	.99
Dairy form	.60	.62	.79	.80	.99	.99
Rump angle	.66	.67	.83	.83	.99	.99
Thurl width	.61	.64	.80	.81	.99	.99
Rear leg set	.51	.52	.72	.72	.98	.98
Foot angle	.39	.45	.62	.65	.97	.97
Fore attachment	.54	.60	.75	.77	.98	.98
Rear udder height	.54	.57	.75	.76	.98	.98
Rear udder width	.51	.55	.72	.74	.98	.98
Udder cleft	.49	.53	.71	.73	.98	.98
Udder depth	.63	.65	.81	.82	.99	.99
Teat placement	.58	.60	.78	.78	.99	.99

TABLE 6. Variance parameters for final score with and without adjustment for merit of mates.

Adjustment for mates	Error variance	Sire variance	Heritability
Yes	10.34	.747	.270
No	10.50	.821	.290

for AI bulls (daughters in 10 or more herds) and non-AI bulls. Solutions for AI bulls were .17 higher than for non-AI bulls for recent years and up to 1.11 higher for earlier years, which indicates that AI sampled bulls are selected more intensely. Unknown sires and MGS were grouped together except that MGS were assumed to be 2 yr older. Total number of groups was 11 for linear traits and 16 for final score data, which included more records and years. Oldest unknown male and female groups were partially confounded, and solutions for these converged slowly. A common group for earliest unknown parents might be appropriate.

CONCLUSIONS

Methods to improve predictions of genetic merit for type traits were investigated. Genetic variances and covariances for linear type traits measured by the Holstein Association were estimated by multiple trait REML. Estimates should be biased little by selection because 1) relationships among sires were included, 2) a large percentage of total records were used, and 3) many type traits are not highly correlated with milk production on which most selection is based. Substantially higher reliabilities for some linear trait evaluations resulted from using information from correlated traits in routine predictions of genetic merit by multiple trait methodology.

Adjustment for merit of mates now is included in evaluations of both linear traits and final score; however, changes in rank were small. An improved genetic grouping procedure that treats both fixed and random portions of genetic effects as inherited was used. Parameter estimates and models from this study were

implemented by the Holstein Association for computing July 1988 genetic evaluations for final score and linear traits. Further improvements might be use of an animal rather than sire model and inclusion of all scores in the evaluation rather than only those from records closest to 30 mo. Use of linear traits in an index of overall profitability should rank animals more accurately than would an index using final scores if appropriate economic values are used.

ACKNOWLEDGMENTS

REFERENCES

1 Foulley, J. L., S. Calomiti, and D. Gianola. 1982. Ecriture des equations du BLUP multicaracteres. Ann. Genet. Select. Anim. 14:309.
2 Funk, D. C., and L. B. Hansen. 1988. Adjustment factors for age and stage of lactation for linear type traits from Holstein classification. J. Dairy Sci. 71(Suppl. 1):143. (Abstr.)
3 Henderson, C. R. 1976. Multiple trait sire evaluation using the relationship matrix. J. Dairy Sci. 59:769.
4 Jensen, J., and I. L. Mao. 1988. Transformation algorithms in analysis of single trait and of multitrait models with equal design matrices and one random factor per trait: a review. J. Anim. Sci. 66:2750.
5 Klei, L., E. J. Pollak, and R. L. Quaas. 1988. Genetic and environmental parameters associated with linearized type appraisal scores. J. Dairy Sci. 71:2744.
6 Lawstuen, D. A., L. B. Hansen, and L. P. Johnson. 1987. Inheritance and relationship of linear type traits for age groups of Holsteins. J. Dairy Sci. 70:1027.
7 Norman, H. D., R. L. Powell, and J. R. Wright. 1987. Influence of genetic differences in merit of mates on sire evaluation. J. Dairy Sci. 70:141.
8 Powell, R. L. 1978. A procedure for including the dam and maternal grandsire in USDA-DHIA Cow Indexes. J. Dairy Sci. 61:794.
9 Thompson, J. R., K. L. Lee, A. E. Freeman, and L. P. Johnson. 1983. Evaluation of a linearized type appraisal system for Holstein cattle. J. Dairy Sci. 66:325.
10 Thompson, R. 1979. Sire evaluation. Biometrics 35:339.
11 VanRaden, P. M. 1986. Computational strategies for estimation of variance components. Ph.D. thesis. Iowa State Univ., Ames.
12 VanRaden, P. M. 1988. Economic value of body size in Holsteins. J. Dairy Sci. 71(Suppl. 1):238. (Abstr.)
13 Westell, R. A., R. L. Quaas, and L. D. Van Vleck. 1988. Genetic groups in an animal model. J. Dairy Sci. 71: 1310.