Estimation of Genetic Parameters for Somatic Cell Score in Holsteins[1]

Y. DA, M. GROSSMAN,[2] and I. MISZTAL[3]
Department of Animal Sciences
University of Illinois
Urbana 61801

G. R. WIGGANS
Animal Improvement Programs Laboratory
Agricultural Research Service, USDA
Beltsville, MD 20705-2350

ABSTRACT

Genetic parameters of somatic cell scores for Holstein cows were estimated using an animal model and REML for two data sets. Set 1, with 13,017 records from 5278 cows, was used to obtain variance components, heritability, and repeatability for two lactation measures: the simple average and the weighted average of test day data. Set 2, with 14,418 records from 4806 cows, was used to obtain genetic correlations for the simple average between lactations 1 and 2, between lactations 1 and 3, and between lactations 2 and 3. Simple and weighted average of test day somatic cell scores had the same heritabilities (.12) and repeatabilities (.35). Phenotypic variances were about 1.2, and herd-sire interaction variances were small (.002). Genetic correlation for somatic cell score was about .55 between lactations 1 and 2 and between lactations 1 and 3 and .65 between lactations 2 and 3. Phenotypic correlation was .20 between lactations 1 and 2, .16 between lactations 1 and 3, and .31 between lactations 2 and 3.

(**Key words**: somatic cell scores, genetic parameters, variance components, Holsteins)

Abbreviation key: **SA** = simple average of SCS for test day data, **SCS** = somatic cell score, **WA** = weighted average of SCS for test day data.

INTRODUCTION

Somatic cell count is used to monitor mastitis, which causes serious economic loss. In the US alone, the cost of mastitis to the dairy industry is about 2 billion dollars annually (14). Somatic cell count also is used to monitor milk quality. The US Public Health Service and the FDA have established SCC of 750,000/ml as the maximum acceptable concentration in milk (2). Many management practices have been identified to control this disease. Genetic selection, however, is the approach to improve resistance to mastitis in future generations.

Genetic parameters, which are functions of (co)variance components, provide information about the genetic nature of a trait and are needed for genetic evaluations and selection strategies. As a measure for statistical and genetic analyses, however, SCC has several deficiencies: its distribution is not normal, and its relationship with milk yield is not linear (21). Somatic cell score (**SCS**), which is the $\log_2$ transformation of SCC [$SCS = \log_2(SCC/100) + 3$], corrects the problems of SCC (21) and has been accepted by the National Cooperative DHI Program (20) as a standard recording scale for SCC. In another study (7), a log transformation of total number of somatic cells was suggested to avoid effect of dilution or stage of lactation.

Several authors recently have reported estimates of genetic parameters for the somatic cell trait, using various expressions of the trait, methods to combine test day records, and

Received September 27, 1991.
Accepted March 23, 1992.
[1]Supported in part by the Illinois Agricultural Experiment Station, Hatch Project 35-367.
[2]Reprint requests.
[3]Supported by the Holstein Association.

models and methods for estimation. Expressions of the somatic cell trait include SCC (19), SCS (1), $\log_e$ SCC (7, 8, 13, 22), and $\log_{10}$ SCC (5). Methods to combine test day records include arithmetic average of SCS (5, 7, 19), $\log_e$ (sum of SCC) (6, 7), lactation average of adjusted individual test day records (1, 22), and geometric mean of $\log_e$ SCC (13). Estimation has been from sire models (7, 8, 13, 19, 22) and an animal model (5), using Henderson's methods 1 (8, 13) and 3 (19) and REML (1, 5, 7, 19). Heritabilities ranged from .01 (5) to .47 (22). The variety of expressions for the trait and models for estimation might be a cause for the wide variation in estimates of genetic parameters.

A standard method to combine test day SCS into one measure for a lactation has not been determined. In addition to methods reviewed earlier, a weighted average (WA) of test day SCS has been proposed (26). Estimates of genetic parameters for WA, however, are not available.

In general, prediction of genetic values with multiple lactation records may use a unitrait repeatability model, in which different lactations are treated as manifestations of the same trait, or a multitrait model, in which different lactations are treated as different traits. Genetic correlations between SCS from different lactations can help to determine whether SCS should be evaluated as the same trait or as different traits. If genetic correlation is near unity, then values for SCS from different lactations should be treated as repeated records of the same trait; otherwise, they should be treated as different traits. Conflicting evidence exists, however, as to whether genetic correlation is high between SCS from different lactations. Correlations between adjacent lactations ranged from .44 to .77 (22), between pairs of the first three lactations from .90 to .97 (13), between first and later lactations from .71 to .86, and between second and third lactations about unity (1).

Estimates of genetic parameters for SCS were different for different lactations (7, 13, 19, 22), and curves for SCS of the first lactation were different from those of later lactations (18, 26). Such findings may favor a two-trait model with SCS of first lactation as one trait and SCS of second and later lactations as the other trait. Estimates of heritabilities and genetic correlations for SCS are necessary for modeling SCS.

The purpose of this study was to estimate variance components, heritability, and repeatability of SCS for a unitrait repeatability model and to estimate genetic correlations for a multitrait model, treating SCS from different lactations as different traits, using the animal model and REML estimation procedure.

MATERIALS AND METHODS

Data

Test day SCS records were provided by the North Carolina Data Processing Center. Two data sets were used to estimate genetic parameters. Set 1 was used to obtain estimates of genetic parameters under a repeatability model using two lactation measures: the simple average (SA) and the WA (26) of test day data. Set 2 was used to estimate genetic correlations under a multitrait animal model, treating SCS from each of the first three lactations as three different traits. For each set, data were collected through an official DHI plan. First lactation records and the first four SCS test day data for each lactation were present, and sires of cows were known.

First lactation was required because selection decisions are often based on first lactation. As minimum information for a lactation, cows were required to have first four tests; thus, cows culled after the first four tests could be included in the sample. Set 1 had 13,017 records from the first four lactations of 5278 Holstein cows. In addition to a first lactation record, a later record was required for each cow to have more information about permanent environment effects. Requiring more than one record, however, may have resulted in some selection of cows. Such selection may have resulted from voluntary culling for low yield and also from involuntary culling for mastitis, each of which could affect SCS results, because SCS is correlated to both. It is doubtful, however, that much first lactation culling would be directly for SCS.

Pedigrees of cows in set 1 were traced back to grandparents and included 530 sires (334 were sires of cows with records, and each sire had at least 4 daughters) and 2336 dams without records but with either pedigree informa-

tion or at least two offspring. Parents not connecting animals were treated as unknown parents that were assigned to genetic groups according to the ages of their daughters. Set 2 had 14,418 records from the first three lactations of 4806 cows, 260 sires, and 4189 dams without records; pedigrees of cows were traced back to parents of cows. Dams without records were included because of the ties they provided and because of the minimal additional computations they required when a property of REML was used (11). Requiring the first three records for each cow led to balanced information on each trait and to computing efficiency. Cows in sets 1 and 2 were from North Carolina and Virginia, and set 2 also included cows from Florida, Georgia, and Texas.

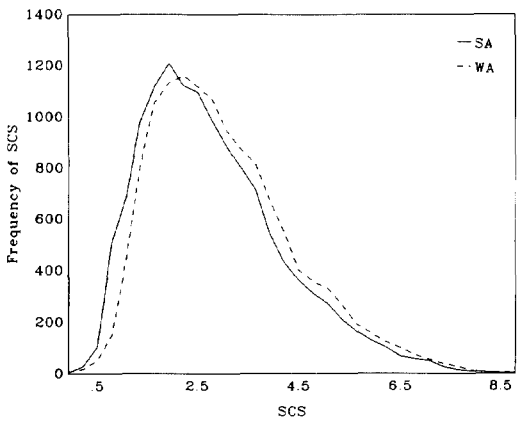

Figure 1. Distribution of values for somatic cell score (SCS) using simple average (SA) and weighted average (WA) of test day data.

Lactation Measures

A lactation measure combines individual test day SCS data into a single value for a lactation (26). Two lactation measures of test day SCS were used: SA and WA (26). The SA and WA were computed as

$$SA = \sum_{i=1}^{k} SCS_i/k$$

$$WA = \sum_{i=1}^{k} w_i SCS_i / \sum_{i=1}^{k} w_i,$$

where SCS_i is the SCS at test i: w_i is the statistical weight for SCS_i, obtained by correcting for the effect of lactation stage (26); and k is the number of tests in the lactation. Number of tests averaged 9.4 per lactation and ranged from 4 to 11. Distributions of values for SA and WA were skewed slightly to the right (Figure 1), probably because some cows were in the later stages of subclinical mastitis or in clinical mastitis. Sample statistics for SA and WA are in Table 1.

Model

An animal model (4, 15, 24, 25) was used to estimate genetic parameters. For unitrait estimations using WA, random effects in the model included herd-sire interaction, permanent environment and additive genetic effects, and residuals. Herd-sire interaction was dropped for SA because the result for WA showed that herd-sire interaction variance was small. For multitrait estimation, random effects included additive genetic effects and residuals. Factors affecting SCS were studied to determine which to include as fixed effects in the model, using a raw data set of more than 20,000 records, from which set 1 was derived. Means of SA and WA increased by year of calving (Figure 2) from 1980 through 1986, except for 1983, and were higher for calvings in June through September (defined as season 1) than for calvings in other months (defined as season 2) (Figure 3). The reason for the

TABLE 1. Somatic cell score (SCS) by lactation measure (n = 13,017).

Lactation measure[1]	$\bar{X}$	SD	Minimum	Maximum
SA	2.63	1.21	.14	8.20
WA	3.08	1.21	.40	8.45

[1]SA = Simple average of test day SCS, WA = weighted average of test day SCS; SA and WA averaged 9.4 tests per lactation.

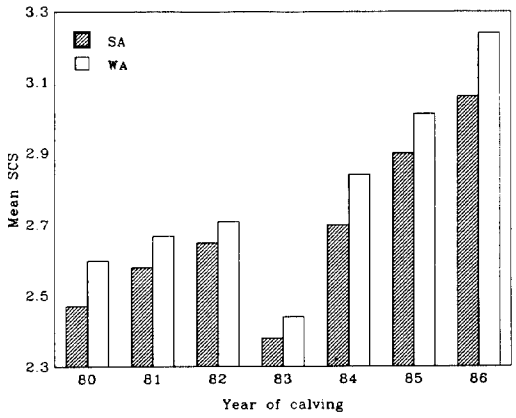

Figure 2. Mean somatic cell scores (SCS) for simple average (SA) and weighted average (WA) of test day SCS data by year of calving.

decrease in 1983 is unknown. The increase of SCS in recent years may reflect a positive genetic correlation with milk yield, which is also increasing. Means of SCS increased with increasing test, except for test 2, and with lactation number (Figure 4). Previous studies (18, 26) also found similar effects of season and lactation. For set 1, fixed effects included 714 herd-year-seasons, four lactations, and 10 genetic groups of unknown parents. For set 2, fixed effects included 1145 herd-year-seasons, and 4 genetic groups of unknown parents.

REML Estimation

The REML formulation for the animal model with groups (4) was used to estimate genetic parameters, and computer algorithms used sparse matrix techniques (10, 12). For set 1, three REML estimations were conducted. Estimation 1 used SA and the animal model without herd-sire interaction; estimation 2 used WA and the same model as estimation 1, and estimation 3 used WA and the animal model with herd-sire interaction. For set 2, one multitrait REML estimation was conducted using SA. To achieve computing efficiency, the multitrait estimation used a canonical transformation (3, 9, 17, 23) that requires the same incidence matrices for each trait, which is not the case for the three different lactations. The incidence matrix for lactation 2 was used as an approximate incidence matrix for lactations 1 and 3. This approximation should be close to using the original incidence matrices, because records for the same cows had the same herd effects, and most calving years and months for consecutive lactations were about 1 yr apart.

RESULTS AND DISCUSSION

For the unitrait repeatability model (Table 2), additive genetic variance (σ_a^2) was about .15 for SA and for WA, and variance of permanent environment effects (σ_p^2) was almost

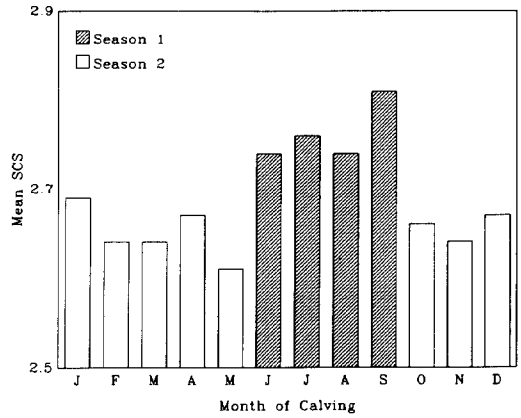

Figure 3. Mean somatic cell scores (SCS) by month of calving; season 1 contained June through September, and season 2 contained the other months.

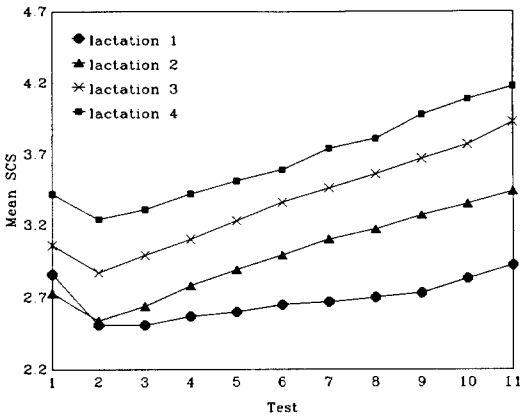

Figure 4. Mean somatic cell scores (SCS) by test for first four lactations.

TABLE 2. Estimates of additive genetic variance (σ_a^2), permanent environment variance (σ_p^2), herd-sire interaction variance (σ_c^2), residual variance (σ_e^2), phenotypic variance, h^2, and repeatability (r) for somatic cell scores, using two lactation measures: the simple average (SA) and the weighted average (WA).

Lactation measure[1]	σ_a^2	σ_p^2	σ_c^2	σ_e^2	σ_T^2	h^2	r
SA	.145	.273	NE[2]	.757	1.175	.123	.355
WA	.148	.276	NE	.778	1.202	.123	.353
WA[3]	.147	.275	.002	.778	1.202	.122	.353[4]

[1]SA = Simple average of test day SCS, WA = weighted average of test day SCS.

[2]NE = Not estimated.

[3]Variance components in this row are estimated using the animal model with herd-sire interaction.

[4]Computed as $r = (\sigma_a^2 + \sigma_p^2 + \sigma_c^2)/(\sigma_a^2 + \sigma_p^2 + \sigma_c^2 + \sigma_e^2)$; other values of r computed as $r = (\sigma_a^2 + \sigma_p^2)/(\sigma_a^2 + \sigma_p^2 + \sigma_e^2)$.

twice as large, about .27 for SA and for WA. Herd-sire interaction variance for WA (σ_c^2) was small (.002). Phenotypic variance (σ_T^2) was 1.2 for SA and for WA. Heritability for SCS was about .12 for SA and for WA. Repeatability was about .35 for SA and for WA. For the multitrait model (Table 3), genetic correlation for SCS was about .55 between lactations 1 and 2 and between lactations 1 and 3 and .65 between lactations 2 and 3. Phenotypic correlation for SCS was .20 between lactations 1 and 2, .16 between lactations 1 and 3, and .31 between lactations 2 and 3. Heritability from the multitrait model was .05 for lactation 1, .07 for lactation 2, and .11 for lactation 3.

Heritability estimates from the unitrait animal model were close to estimates (from .09 to .13) by Banos and Shook (1) and Schutz et al. (19). Results show that estimates of genetic parameters were about the same for SA and WA, probably because there were relatively few missing tests; consequently, phenotypic correlation between SA and WA was high (.995). The small herd-sire interaction variance indicates that its effect on genetic evaluations of SCS is negligible. A larger estimate for herd-sire interaction variance, about .02, was obtained by Banos and Shook (1), but that estimate is still small relative to additive and permanent environment effects. The high repeatability, relative to heritability, is because of the relatively large permanent environment variance, which implies that substantial improvement in accuracy of prediction could be achieved by using repeated observations.

Genetic correlations between SCS from three lactations were moderately high and close to those by Shook et al. (22), but lower than those by Banos and Shook (1) and by Monardes and Hayes (13). Because genetic correlations were not near unity, our results favor a multitrait model for the genetic evaluation of SCS. Genetic and phenotypic correlations indicate that lactations 2 and 3 are more similar than lactations 1 and 2 or lactations 1 and 3. A multitrait model, therefore, with first and later lactations as different traits is appropriate for the genetic evaluation of SCS.

TABLE 3. Estimates of genetic parameters of somatic cell scores using a multitrait animal model treating lactations as different traits.[1]

	Lactation 1	Lactation 2	Lactation 3
Lactation 1	.05	.55	.54
Lactation 2	.20	.07	.65
Lactation 3	.16	.31	.11

[1]Heritabilities are diagonal elements, genetic correlations are upper off-diagonal elements, and phenotypic correlations are lower off-diagonal elements.

The requirement for a later record in set 1 may have resulted in a selected sample for SCS. The requirement for three lactations in set 2 could have caused more selection than in set 1. The lower heritability for lactation 1 could be an indication that the sample was selected because selection can reduce heritability (16). The increased estimates of heritability for lactations 2 and 3, however, suggest that selection for SCS was not strong and that those estimates of heritabilities and genetic correlations are approximately correct. Estimates for genetic correlation should have been affected by selection even less than estimates for heritability.

CONCLUSIONS

Based on variance components estimated using a unitrait repeatability animal model and REML, heritability for SCS was relatively low, and repeatability was about three times as high as heritability, implying that repeated records should be used to improve prediction accuracy under a unitrait evaluation. Herd-sire interaction variance was negligible for the genetic evaluation of SCS. Based on variance and covariance components estimated using a multitrait animal model and REML, genetic correlations for SCS between the first three lactations were moderately high, and phenotypic correlations were low. Genetic and phenotypic correlations between lactations 1 and 2 and between lactations 1 and 3 were lower than those between lactations 2 and 3, implying that first lactation could be treated as a different trait from later lactations.

ACKNOWLEDGMENTS

The authors thank two anonymous reviewers and W. J. Koops, M. M. Schutz, R. D. Shanks, and P. M. VanRaden for valuable comments and suggestions and Cindy Ernst for help in preparing the data. Computing support was provided by the National Center for Supercomputer Applications at the University of Illinois at Urbana-Champaign and by the Cornell National Supercomputer Facility, which is funded in part by the National Science Foundation, New York State, and IBM Corporation.

REFERENCES

1 Banos, G., and G. E. Shook. 1990. Genotype by environment interaction and genetic correlations among parities for somatic cell count and milk yield. J. Dairy Sci. 73:2563.

2 Bennett, R. 1992. Lead, follow, or get out the way: the new PMO SCC policy. Page 52 *in* Proc. 31st Annu. Mtg. Natl. Mastitis Counc., Arlington, VA. Natl. Mastitis Counc., Arlington, VA.

3 Da, Y., and M. Grossman. 1991. Multitrait animal model with genetic groups. J. Dairy Sci. 74:3183.

4 Da, Y., M. Grossman, and I. Misztal. 1989. Prediction error variance and restricted maximum likelihood estimation for animal model with relationship grouping. J. Dairy Sci. 72:2125.

5 Emanuelson, U., B. Danell, and J. Philipsson. 1988. Genetic parameters for clinical mastitis, somatic cell count, and milk production estimated by multiple-trait restricted maximum likelihood. J. Dairy Sci. 71:467.

6 Heuven, H.C.M. 1987. Diagnostic and genetic analysis of mastitis field data. Ph.D. Diss., Univ. Wisconsin-Madison, Madison.

7 Heuven, H.C.M., H. Bovenhuis, and R. D. Politiek. 1988. Inheritance of somatic cell count and its genetic relationship with milk yield in different parities. Livest. Prod. Sci. 18:115.

8 Kennedy, B. W., M. S. Sethar, J. E. Moxley, and B. R. Downey. 1982. Heritability of somatic cell count and its relationship with milk and composition in Holsteins. J. Dairy Sci. 65:843.

9 Meyer, K. 1985. Maximum likelihood estimation of variance components for a multivariate mixed model with equal design matrices. Biometrics 41:153.

10 Misztal, I. 1990. Restricted maximum likelihood estimation of variance components in animal model using sparse matrix inversion and a supercomputer. J. Dairy Sci. 73:163.

11 Misztal, I., and Y. Da. 1990. Property of trace in restricted maximum likelihood estimation of variance components. J. Dairy Sci. 73:3583.

12 Misztal, I., T. J. Lawlor, T. Short, and P. M. VanRaden. 1992. Multiple-trait REML estimation of variance components of yield and type traits in dairy cattle using an animal model. J. Dairy Sci. 75:554.

13 Monardes, H. G., and J. F. Hayes. 1985. Genetic and phenotypic statistics of lactation cell counts in different lactations of Holstein cows. J. Dairy Sci. 68:1449.

14 National Mastitis Council. 1987. Current Concepts of Bovine Mastitis. 3rd ed. Natl. Mastitis Counc., Arlington, VA.

15 Quaas, R. L. 1988. Additive genetic model with groups and relationships. J. Dairy Sci. 71:1338.

16 Robertson, A. 1977. The effect of selection on the estimation of genetic parameters. Z. Tierz. Zuechtungsbiol. 94:131.

17 Schaeffer, L. R., and B. W. Kennedy. 1986. Computing solutions to the mixed model equations. Proc. 3rd World Congr. Genet. Appl. Livest. Prod., Lincoln, NE XII:391.

18 Schutz, M. M., L. B. Hansen, G. R. Steuernagel, and A. L. Kuck. 1990. Variation of milk, fat, protein, and somatic cell for dairy cattle. J. Dairy Sci. 73:484.

19 Schutz, M. M., L. B. Hansen, G. R. Steuernagel, and A. L. Kuck. 1990. Genetic parameters for somatic cells, protein, and fat in milk of Holsteins. J. Dairy Sci. 73:494.

20 Sechrist, R. S. 1985. Summary of NCDHIP policies.

Natl. Coop. Dairy Herd Improvement Prog. Handbook, Fact Sheet B-3. Natl. DHIA, Columbus, OH.
21 Shook, G. E. 1982. Approaches to summarizing somatic cell count which improve interpretability. Page 150 *in* Proc. 21st Annu. Mtg. Natl. Mastitis Counc., Louisville, KY. Natl. Mastitis Counc., Arlington, VA.
22 Shook, G. E., F. Ruvuna, and A.K.A. Ali. 1982. Genetic parameters for lactation average of somatic cell concentration in milk. Proc. 2nd World Congr. Genet. Appl. Livest. Prod., Madrid, Spain VIII:142.
23 Thompson, R. 1976. Estimation of quantitative genetic parameters. Page 639 *in* Proc. Int. Conf. Quant. Genet. E. Pollak, O. Kempthorne, and T. B. Bailey, ed. Iowa State Univ. Press, Ames.
24 Westell, R. A., R. L. Quaas, and L. D. Van Vleck. 1988. Genetic groups in an animal model. J. Dairy Sci. 71:1310.
25 Wiggans, G. R., I. Misztal, and L. D. Van Vleck. 1988. Animal model evaluation of Ayrshire milk yield with all lactations, herd-sire interaction, and groups based on unknown parents. J. Dairy Sci. 71:1319.
26 Wiggans, G. R., and G. E. Shook. 1987. A lactation measure of somatic cell count. J. Dairy Sci. 70:2666.