Short Communication: Variance Estimates Among Herds Stratified by Individual Herd Heritability

C. D. Dechow,[*][1] **H. D. Norman,**[†] **and C. A. Pelensky**[*]
*Department of Dairy and Animal Science, The Pennsylvania State University, University Park 16802
†Animal Improvement Programs Laboratory, Agricultural Research Service, USDA, Beltsville, MD 20705-2350

ABSTRACT

The objectives of this study were to compare (co)variance parameter estimates among subsets of data that were pooled from herds with high, medium, or low individual herd heritability estimates and to compare individual herd heritability estimates to REML heritability estimates for pooled data sets. A regression model was applied to milk yield, fat yield, protein yield, and somatic cell score (SCS) records from 20,902 herds to generate individual-herd heritability estimates. Herds representing the 5th percentile or less (P5), 47th through the 53rd percentile (P50), and the 95th percentile or higher (P95) for herd heritability were randomly selected. Yield or SCS from the selected herds were pooled for each percentile group and treated as separate traits. Records from P5, P50, and P95 were then analyzed with a 3-trait animal model. Heritability estimates were 23, 31, 26, and 8% higher in P95 than in P5 for milk yield, fat yield, protein yield, and SCS, respectively. The regression techniques successfully stratified individual herds by heritability, and additive genetic variance increased progressively, whereas permanent environmental variance decreased progressively as herd heritability increased.

Key words: herd heritability, daughter-dam regression, daughter-sire regression

Dechow and Norman (2007) estimated individual herd heritabilities for milk yield, fat yield, protein yield, and SCS for 1,939 herds using regression techniques. Correlations of regression heritabilities with individual herd heritability estimates generated with REML ranged from 0.45 to 0.68 for a subset of the 45 largest herds. However, attempts to adjust records to a constant genetic variance using the individual herd heritability estimates did not appreciably improve the accuracy of genetic evaluations. Lack of improvements when

data were adjusted for individual herd heritabilities could indicate that the true genetic variance was not different between herds with high and low heritability.

Data quality issues, such as sire misidentification rate, would also be expected to impact herd heritability estimates regardless of the herd's genetic variance. Individual herd heritabilities were generated for 20,902 herds and were merged with sire-misidentification rate as determined by DNA maker analysis for a subset of 230 herds (Dechow et al., 2008). The correlation between sire misidentification rate and a principal component for all measures of herd heritability was −0.50.

The first objective of this study was to compare (co)variance parameter estimates for milk yield, fat yield, protein yield, and SCS among subsets of data that were pooled from herds with high, medium, or low individual herd heritability estimates. The second objective was to assess how individual herd heritability estimates compare with REML heritability estimates when data were pooled for herds with similar individual heritability estimates.

Individual herd heritabilities for 20,902 herds were estimated (Dechow et al., 2008). Briefly, the model used to generate individual herd heritabilities in ASREML (Gilmour et al., 2006) was

$$y_{ijklmno} = BD_k + b_l A_l + HYS_m + b_d D_n + b_s S_n \quad [1]$$
$$+ b_{sd}(S \times SD_o) + b_{do}F_o + b_{so}G_o + e_{ijklmno},$$

where $y_{ijklmno}$ = mature-equivalent milk yield, mature-equivalent fat yield, mature-equivalent protein yield, or SCS for the ith record of cow j of breed k, parity l, calving in herd-year-season m, in state n and herd o; BD = the fixed effect of breed, b_l = coefficient for fixed regression on age (A) nested within parity, HYS = fixed effect of herd-year-season, b_d = coefficient for fixed regression on dam record nested within state (D), b_s = coefficient for fixed regression on sire PTA (S) nested within state, b_{sd} = coefficient for fixed regression on the interaction between sire PTA and herd standard deviation (SD), b_{do} = coefficient for random regression on dam record nested within herd (F), b_{so} = coefficient

Table 1. The total number of records (n), cows, sires, and herds represented in each subset in addition to average mature-equivalent milk, fat, and protein yields and SCS

Subset[1]	n	Cows	Sires	Herds	Milk (kg)	Fat (kg)	Protein (kg)	SCS
P5	34,228	20,855	497	100	11,184	401	338	3.09
P50	37,473	20,111	474	260	11,434	419	342	2.90
P95	40,345	21,118	507	138	12,033	440	361	2.81

[1]P5 = herds in the 5th percentile or lower for herd heritability; P50 = herds in the 47th to the 53rd percentile for herd heritability; P95 = herds in the 95th percentile or higher for herd heritability.

for random regression on sire PTA nested within herd (G), and $e_{ijklmno}$ = effect of random residual. The random regression coefficients (b_{do} and b_{so}) were assumed to be correlated.

Individual herd heritability from daughter-dam regression was $2(b_d + b_{do})$. Individual herd heritability for daughter-sire regression was:

$$\frac{\{[b_s + b_{sd}(SD_o) + b_{so}]SD_{US}\}^2/R}{(SD_o)^2},$$

where b_s, b_{sd}, SD_o, and b_{so} are as defined in model 1; SD_{US} = genetic standard deviation assumed for USDA-DHIA Holstein genetic evaluations (655 kg for mature-equivalent milk); and R = average sire PTA reliability for all cows in a herd.

A principal component for heritability was generated for each herd using 8 herd heritability measures (daughter-dam and daughter-sire estimates for each of the 4 traits), and herds were stratified according to the principal component for herd heritability. Mature-equivalent milk, fat, and protein yields, and SCS from Holstein cows calving between August 2000 and August 2005 were extracted from the national dairy database if the cow calved in a randomly chosen herd with a principal component of heritability that fell into one of the following percentiles: lowest 5% (**P5**), 47 to 53% (**P50**), or above 95% (**P95**). The target number of cows in each subset was 20,000, and only cows with a first-lactation record available, an age at calving ranging from 20 to 120 mo, and that were from a sire with at least 10 daughters in the subset were retained. The total number of herds, cows, records, and sires represented in each subset in addition to average milk yield, fat yield, protein yield, and SCS are reported in Table 1.

Records from all herds within a subset were pooled and each subset was treated as a separate trait with heritabilities and correlations among subsets estimated using the following 3-trait animal model in ASREML:

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{Z}\mathbf{a} + \mathbf{W}\mathbf{p} + \varepsilon,$$

where **Y** = a vector of records from P5, P50, and P95 for either milk yield, fat yield, protein yield, or SCS;

β = a vector of fixed herd-year-season of calving effects with 6 bimonthly calving seasons, and linear and quadratic effects for age at calving nested with parity groups 1, 2, and ≥3; **a** = a vector of random animal effects; **p** = a vector of random permanent environmental effects; **X**, **Z**, and **W** are the corresponding incidence matrices; and ε = a vector of random errors. All traits were also analyzed with single-trait models that included an additional random effect for sire by herd interaction.

The average daughter-dam and daughter-sire heritability estimates for P5, P50, and P95 are reported in Table 2. The principal component for herd heritability was successful in stratifying herds according to herd heritability for all traits. Daughter-dam heritability estimates were higher than daughter-sire heritability estimates for all traits and in all subsets. Average daughter-dam heritability estimates ranged from 0.21 for protein yield in P5 to 0.53 for fat yield in P95. Daughter-sire herd heritability estimates ranged from an average of 0.09 for fat yield, protein yield, and SCS for P5 to 0.40 for protein yield in P95.

Estimates from ASREML for heritability, repeatability, additive genetic variance, and genetic correlations among yield and SCS in P5, P50, and P95 are displayed in Table 3. Heritability estimates were closer in magnitude to daughter-sire heritability estimates than daughter-dam heritability estimates. For all 4 traits, heritability estimates were lowest for P5, intermediate for P50, and highest for P95. Heritability estimates for P95 were 23, 31, 26, and 8% higher than heritability estimates from P5 for milk yield, fat yield, protein yield, and SCS, respectively. Repeatability differences between P95 and P5 were not as dramatic as those for heritability, ranging from 6 to 12%.

It is likely that the higher heritability estimates observed for P95 are due, in part, to lower sire misidentification rates. Dechow et al. (2008) reported a correlation of −0.50 between sire misidentification rate in 230 herds and a principal component for 12 measures of herd heritability. If the level of parent misidentification in P5 was higher than in P95, the estimate of genetic variance would be expected to be lower in P5 even if true genetic variance was not different. The higher per-

Table 2. The average daughter-dam and daughter-sire herd heritability estimates for herds in each subset

Subset[1]	Daughter-dam				Daughter-sire			
	Milk	Fat	Protein	SCS	Milk	Fat	Protein	SCS
P5	0.24	0.29	0.21	0.23	0.10	0.09	0.09	0.09
P50	0.36	0.41	0.32	0.26	0.19	0.19	0.17	0.11
P95	0.46	0.53	0.44	0.27	0.37	0.37	0.40	0.12

[1]P5 = herds in the 5th percentile or lower for herd heritability; P50 = herds in the 47th to the 53rd percentile for herd heritability; P95 = herds in the 95th percentile or higher for herd heritability.

manent environmental variance estimate in P5 likely reflects a large component of genetic variance that was not detected due to misidentification. The relatively small differences in repeatability when compared with heritability are supportive of the hypothesis that true genetic variance is similar between high and low heritability herds, but that estimated genetic variance is lower due to sire misidentification. Heterogeneous variance adjustments standardize genetic variance across herds and have resulted in small improvements in the accuracy of genetic evaluations (Wiggans and VanRaden, 1991). Attempts to standardize genetic variance based on daughter-dam and daughter-sire heritability estimates generally resulted in minimal improvements in evaluation accuracy (Dechow and Norman, 2007). If differences in herd heritability were due primarily to sire misidentification rates and not true differences in genetic variance, then variance adjustments based on herd heritability would not be expected to improve evaluation accuracy.

The highest phenotypic variance for all traits was observed for P95. Several studies have described a tendency toward higher heritability estimates as phenotypic variance increases (Lofgren et al., 1985; Vinson, 1987; Van Tassell et al., 1999). However, phenotypic variance was not generally higher for P50 than for P5.

Dechow and Norman (2007) speculated that daughter-sire PTA regression could be depressed in the presence of genotype by environment interaction. Sire by herd interactions might be expected to be higher in low daughter-sire heritability herds if genotype by environment interactions were responsible for lower individual herd heritability estimates. However, estimates of sire by herd interaction ranged from 0.84% (SCS) to 2.15% (milk yield) in P5, 1.16% (milk yield) to 1.98% (protein yield) in P50, and from 1.78% (milk yield) to 2.72% (fat yield) in P95. Genetic correlations among P5, P50, and P95 were also high, ranging from 0.86 to 1. There was little evidence that genotype by environment interaction depressed daughter-sire herd heritability estimates.

Genetic parameter estimates for pooled samples of records from herds that were stratified into high, medium, and low herd heritability groups indicated that

Table 3. Heritability (diagonals), genetic correlation (above diagonals), repeatability (r), phenotypic variance (V_p), additive genetic variance (V_a), and permanent environmental variance (V_{pe}) estimates for milk yield, fat yield, protein yield, and SCS stratified by level of individual herd heritability[1]

	Herd heritability			r	V_p	V_a	V_{pe}
	P5	P50	P95				
Milk							
P5	0.13	0.90	0.86	0.41	2,348,889	299,248	668,964
P50		0.22	0.89	0.45	2,336,537	517,076	534,600
P95			0.36	0.49	2,906,776	1,041,789	378,172
Fat							
P5	0.08	0.89	0.97	0.40	29,335	2,285	9,569
P50		0.20	0.93	0.45	31,847	6,503	7,911
P95			0.39	0.52	40,411	15,882	5,104
Protein							
P5	0.10	0.91	0.90	0.42	17,375	1,810	5,435
P50		0.20	0.90	0.45	15,917	3,245	3,959
P95			0.36	0.50	21,060	7,643	2,953
SCS							
P5	0.06	0.94	1.00	0.27	1.38	0.09	0.29
P50		0.10	0.99	0.32	1.48	0.14	0.33
P95			0.14	0.33	1.52	0.21	0.30

[1]P5 = herds in the 5th percentile or lower for herd heritability; P50 = herds in the 47th to the 53rd percentile for herd heritability; P95 = herds in the 95th percentile or higher for herd heritability.

the regression methods used to estimate individual herd heritabilities were effective. Daughter-dam heritability estimates were generally much higher than REML heritability estimates, whereas the magnitude of daughter-sire and REML heritability estimates were of a similar magnitude. Additive genetic variance increased progressively as herd heritability increased, whereas permanent environmental variance decreased as herd heritability increased.

ACKNOWLEDGMENTS

REFERENCES

Dechow, C. D., and H. D. Norman. 2007. Within-herd heritability estimated with daughter-parent regression for yield and somatic cell score. J. Dairy Sci. 90:482–492.

Dechow, C. D., H. D. Norman, N. R. Zwald, C. M. Cowan, and O. M. Meland. 2008. Relationship between individual herd-heritability estimates and sire misidentification rate. J. Dairy Sci. 91:1640–1647.

Gilmour, A. R., B. J. Gogel, B. R. Cullis, S. J. Welham, and R. Thompson. 2006. ASReml User Guide: Release 2.0. VSN International Ltd, Hemel Hempstead, UK.

Lofgren, D. L., W. E. Vinson, and R. E. Pearson. 1985. Heritability of milk yield at different herd means and variances for production. J. Dairy Sci. 68:2737–2739.

Van Tassell, C. P., G. R. Wiggans, and H. D. Norman. 1999. Method R estimates of heritability for milk, fat, and protein yields of United States dairy cattle. J. Dairy Sci. 82:2231–2237.

Vinson, W. E. 1987. Potential bias in genetic evaluations from differences in variation within herds. J. Dairy Sci. 70:2450–2455.

Wiggans, G. R., and P. M. VanRaden. 1991. Method and effect of adjustment for heterogeneous variance. J. Dairy Sci. 74:4350–4357.