

December 1990

# Effects of errors in pedigree on three methods of estimating breeding value for litter size, backfat and average daily gain in swine

T. E. Long

*University of Nebraska-Lincoln*

R. K. Johnson

*University of Nebraska-Lincoln, rjohnson5@unl.edu*

J. W. Keele

*University of Nebraska-Lincoln*

Follow this and additional works at: <http://digitalcommons.unl.edu/animalscifacpub>



Part of the [Animal Sciences Commons](#)

---

Long, T. E.; Johnson, R. K.; and Keele, J. W., "Effects of errors in pedigree on three methods of estimating breeding value for litter size, backfat and average daily gain in swine" (1990). *Faculty Papers and Publications in Animal Science*. 89.  
<http://digitalcommons.unl.edu/animalscifacpub/89>

This Article is brought to you for free and open access by the Animal Science Department at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Papers and Publications in Animal Science by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

## Effects of errors in pedigree on three methods of estimating breeding value for litter size, backfat and average daily gain in swine

T. E. Long, R. K. Johnson, and J. W. Keele

University of Nebraska–Lincoln

Estimated breeding value (EBV) was calculated based on either individual phenotype (SP), an index of individual phenotype and full- and half-sib family averages (SI) or Best Linear Unbiased Prediction (BLUP). Calculations were done with correct data or data with 5, 10, 15 or 20% of the records per generation containing pedigree errors. Traits considered were litter size (LS), backfat (BF) and average daily gain (ADG). When data were correct, BLUP resulted in an advantage in expected genetic gain over SP of 22, 7.2 or 30.8% for LS, BF and ADG, respectively, and over SI of 9.6, 3.8 or 21.4%. When sire and dam pedigrees were incorrect for 20% of the pigs each generation, genetic gain using SI was reduced by 7, 2.5 or 6.5% and genetic gain using BLUP was reduced by 9.3, 3.2 or 12.4% for LS, BF and ADG, respectively. With 20% of the pedigrees in error, the advantages in genetic gain of using BLUP over SP, the method unaffected by errors in pedigree, were 10.5, 3.8 and 14.6% for LS, BF and ADG, respectively. These results suggest that, although BLUP is affected to a greater degree by pedigree errors than SP or SI, selection of swine using BLUP still would improve response to selection over the use of SP or SI.

# EFFECTS OF ERRORS IN PEDIGREE ON THREE METHODS OF ESTIMATING BREEDING VALUE FOR LITTER SIZE, BACKFAT AND AVERAGE DAILY GAIN IN SWINE<sup>1</sup>

T. E. Long<sup>2</sup>, R. K. Johnson<sup>3</sup> and J. W. Keele<sup>4</sup>

University of Nebraska, Lincoln 68583-0908 and  
U.S. Department of Agriculture, Clay Center, NE 68933

## ABSTRACT

Estimated breeding value (EBV) was calculated based on either individual phenotype (SP), an index of individual phenotype and full- and half-sib family averages (SI) or Best Linear Unbiased Prediction (BLUP). Calculations were done with correct data or data with 5, 10, 15 or 20% of the records per generation containing pedigree errors. Traits considered were litter size (LS), backfat (BF) and average daily gain (ADG). When data were correct, BLUP resulted in an advantage in expected genetic gain over SP of 22, 7.2 or 30.8% for LS, BF and ADG, respectively, and over SI of 9.6, 3.8 or 21.4%. When sire and dam pedigrees were incorrect for 20% of the pigs each generation, genetic gain using SI was reduced by 7, 2.5 or 6.5% and genetic gain using BLUP was reduced by 9.3, 3.2 or 12.4% for LS, BF and ADG, respectively. With 20% of the pedigrees in error, the advantages in genetic gain of using BLUP over SP, the method unaffected by errors in pedigree, were 10.5, 3.8 and 14.6% for LS, BF and ADG, respectively. These results suggest that, although BLUP is affected to a greater degree by pedigree errors than SP or SI, selection of swine using BLUP still would improve response to selection over the use of SP or SI. (Key Words: Pigs, Selection Index, Best Linear Unbiased Prediction, Pedigree.)

*J. Anim. Sci.* 1990. 68:4069-4078

## Introduction

A number of procedures have been developed for estimating breeding values of livestock. Application of them to field data, however, can produce biased evaluations when pedigrees contain errors and procedures utilize information from relatives. The result would be less accurate genetic evaluations and slower genetic progress than otherwise would be

possible. Van Vleck (1970a,b) demonstrated that incorrect identification of sires, in cattle data, can bias estimates of heritabilities, evaluations of sires and estimates of genetic progress due to selection. Bias increased as the fraction of records with errors increased. Swine are a litter-bearing species, and swine data contain a larger percentage of full-sibs than cattle data, giving swine data a different relationship structure. Crossfostering of pigs is routine and results in potentially more pedigree errors than in cattle data. Knowledge of the magnitude of the effect of pedigree errors on estimates of breeding values by different procedures would be useful. The purpose of this study was to evaluate the effects of errors in swine pedigrees on estimates of breeding values calculated by individual's phenotype, index of individual plus full and half-sib records or the mixed model procedure of Henderson (1963, 1973), hereafter referred to as BLUP.

<sup>1</sup>Journal Paper No. 9013 of the Nebraska Agric. Exp. Sta.

<sup>2</sup>Present address: Anim. Gen. and Breed. Unit, Univ. of New England, Armidale, New South Wales, 2351, AUST.

<sup>3</sup>Dept. of Anim. Sci. to whom requests for reprints should be made.

<sup>4</sup>Present address: Roman L. Hruska U.S. Meat Anim. Res. Center, Clay Center, NE 68933.

Received August 29, 1989.  
Accepted April 19, 1990.

### Materials and Methods

Data were from the University of Nebraska Gene Pool population, a 14-breed synthetic population that has been closed since 1965. From 1967 to 1977 the population contained two lines. One line was a randomly bred control and the other line underwent nine generations of selection for ovulation rate. The population then underwent two generations of random mating within line, after which the line selected for ovulation rate was partitioned into three lines. One line underwent relaxed selection (selected and mated randomly), one was selected for increased litter size and one was selected for decreased age at puberty. Eight generations of selection were completed. Sires and dams were used for one farrowing in the spring, then replaced, so generation interval was 1 yr. Management of the population and selection procedures have been described by Zimmerman and Cunningham (1975), Cunningham et al. (1979) and Johnson et al. (1984).

Three traits were considered: litter size, number of fully formed pigs at birth (LS), average daily gain from weaning at 28 d to 90 kg (ADG) and backfat probe adjusted to 90 kg of live weight (BF). Data contained 736 sires, 1,393 dams and 2,099 records for LS, 321 sires, 688 dams and 1,953 records for BF, and 321 sires, 689 dams and 2,077 records for ADG. Data from all lines were used for litter size, but only data from the randomly bred control line were used in the analysis of BF and ADG to avoid any effects selection might have had on estimates of parameters and breeding values. Data were combined from all lines for analysis of LS because we felt that, with a small data set, sampling variance might be a much greater problem than effects of selection. Growth rate and backfat were adjusted to an endpoint of 90 kg within sex. The equation to adjust ADG of males was  $ADG_{adj.} = ADG + .0054(90 - WT)$ , and the equation to adjust ADG of females was  $ADG_{adj.} = ADG + .0045(90 - WT)$ . The equation to adjust BF of males was  $BF_{adj.} = BF + .0209(90 - WT)$ , and the equation to adjust females was  $BF_{adj.} = BF + .0192(90 - WT)$ .

Three methods of estimating breeding values (EBV) were chosen to evaluate the effects of errors in pedigree on EBV. The first was selection on phenotype (SP). With this method, EBV was calculated as the product of

heritability,  $h^2$ , and the deviation of an individual's phenotype from its contemporary group mean. For LS, contemporary group was defined as the year-line average, and, for BF and ADG, contemporary group was defined as the year average.

The second method was a selection index (SI) based on an individual's phenotype, full-sib and half-sib family averages. Family averages also were expressed as deviations from the contemporary group mean. Numbers of full- and half-sib family members varied, so a separate selection index was calculated for each animal.

The third method of estimating breeding values was Best Linear Unbiased Prediction (BLUP), developed by Henderson (1963, 1973). The linear model we assumed was:

$$y = XB + Z_1u_1 + Z_2u_2 + e,$$

where  $y$  is an  $N \times 1$  vector of observations,  $X$  is a known  $N \times p$  matrix,  $B$  is an unknown  $p \times 1$  vector of contemporary group effects,  $Z_1$  is the  $N \times q_1$  matrix relating litter environmental effects to observations,  $Z_2$  is the  $N \times q_2$  matrix relating additive genetic effects to observations,  $u_1$  is a  $q_1 \times 1$  vector of random litter environmental effects,  $u_2$  is a  $q_2 \times 1$  vector of random additive genetic effects, and  $e$  is an  $N \times 1$  vector of nonobservable random "errors". Random effects were assumed to have the following expectations and distributions:

$$\begin{aligned} E(u_1) &= E(u_2) = E(e) = 0, \\ \text{Cov}(u_1, u_2) &= \text{Cov}(u_1, e) = \\ \text{Cov}(u_2, e) &= 0, \text{ and} \\ u_1 &\sim (0, I_{v_1^2}), \quad u_2 \sim (0, Av_g^2), \\ e &\sim (0, Iv_0^2), \end{aligned}$$

where  $v_1^2$ ,  $v_g^2$  and  $v_0^2$  represent litter, additive genetic and error variances, respectively,  $A$  is the numerator relationship matrix, and  $I$  is the identity matrix. The goal was to evaluate how well estimated breeding values predicted subsequent progeny performance. Hence, only contemporary and ancestral relative information were used to calculate breeding values.

Variance components were estimated by the pseudo expectation approach of Schaeffer (1986). This method was chosen because it is computationally less demanding than other procedures such as REML. J. W. Keele (unpublished data) found the method of pseudo

expectation had less bias and smaller standard errors with these data than Henderson's Method 3. Heritability estimates were .18, .53 and .13 for LS, BF and ADG, respectively. Estimates of  $c^2$ , common environmental effects, were .002, .07 and .16 for LS, BF and ADG, respectively.

Errors in pedigree were studied at five different levels: 0, 5, 10, 15 or 20% of the records having pedigree errors. Two different situations were investigated. The first was when the assignment of sire was incorrect. This would simulate an error in recording a mating. To simulate this type of error, for each generation of data, litters were reassigned at random to other sires within that generation at the five different levels of incorrect identification. Sires to which litters were reassigned already existed in the data. Situations in which existing animals were mistaken for immigrants were not investigated. The second situation was when assignment of sire and dam was incorrect. This would simulate an error in recording a crossfostering event. To simulate this type of error, for each generation of data, individual pigs were reassigned randomly to other litters within that generation at the five different levels of incorrect identification. After a pig was identified incorrectly, the errors in relationships were maintained in the numerator relationship matrix throughout the experiment.

Higher levels of pedigree errors contained the same errors as lower levels; e.g., data with 10% errors were created by adding 5% more errors to the data with 5% errors. Thus, part-whole correlations existed among these sets of data. Two independent replicates of data with pedigree errors were done and averages of the replicates are presented.

Performance of the three procedures was evaluated by three methods. First, Spearman rank correlation coefficients among estimates of breeding value were calculated. It was assumed that the best rankings were obtained using BLUP on data to which pedigree errors had not been added. Rank correlations were calculated between this ranking and the ranking of animals obtained from the other methods on correct and incorrect data to ascertain how differently animals were being ranked relative to the assumed correct rankings.

The second method was to calculate the regression of offspring's phenotype on parent's

EBV. Because progeny information was not used in the calculation of breeding value, this regression coefficient has an expectation of  $1/2$ , assuming the model is correct and parameters have been estimated exactly.

Breeding values were estimated as  $\hat{u}_2 = \hat{h}^2 (P - \hat{\mu})$  for the SP method. Here,  $P$  is the animal's phenotypic record,  $\hat{\mu}$  is the contemporary group mean and  $\hat{h}^2$  is the estimate of heritability. It is easy to show that the expected value of the regression of offspring phenotype on parents' estimated breeding value equals  $1/2h^2/\hat{h}^2$ , and because the expected value of  $\hat{h}^2$  is  $h^2$ , the expected value of the regression coefficient is  $1/2$ . The expected value of the regression is more difficult to derive for breeding values estimated by SI, but it too can be shown to equal  $1/2$ . Equations from Henderson (1988) can be used to find the expectation of this regression for breeding values estimated by BLUP. He derived the following:

$$\text{Cov}(\hat{u}, \hat{u}_m) = (G - C_{22}) G^{-1} H,$$

where  $u$  is a random variable in the model, estimated by  $\hat{u}$ ,  $u_m$  is a random variable not included in the model that we wish to predict, estimated by  $\hat{u}_m$ ,  $G$  is the variance-covariance matrix of traits in the model,  $C_{22}$  is a matrix of some  $g$ -inverse of the coefficient matrix of the mixed-model equations and  $H = \text{Cov}(u, u_m^1)$ . In our case,  $u$  is a vector of parent breeding values and  $u_m$  is a vector of progeny breeding values. For a single trait,  $G = \sigma_g^2$  and  $H = \frac{1}{2}\sigma_g^2$ . Then, from Henderson (1988),

$$\begin{aligned} \text{Cov}(\hat{u}, \hat{u}_m) &= (G - C_{22}) G^{-1} H \\ &= (G - C_{22}) \left(\frac{1}{\sigma_g^2}\right) \left(\frac{1}{2}\sigma_g^2\right) \\ &= (G - C_{22}) \left(\frac{1}{2}\right) \\ V(\hat{u}) &= (G - C_{22}). \end{aligned}$$

The regression of offspring phenotype ( $u_m + e$ ) on estimated breeding value of parents is

$$\begin{aligned} b &= \\ \frac{\text{Cov}(\hat{u}, u_m + e)}{V(\hat{u})} &= \frac{(G - C_{22})\left(\frac{1}{2}\right)}{(G - C_{22})} = \frac{1}{2}. \end{aligned}$$

The third method of evaluation was to examine expected genetic gain under the

assumption that the breeding value was known. This was assumed to be the estimate from BLUP on correct data. Animals were ranked within generation by the three methods for the various proportions of incorrect pedigrees. A fixed number of the top ranking animals were chosen each generation. For LS, the top ranking eight females were chosen. Number of females per generation-line subclass averaged 40, resulting in an average selection intensity of approximately 20%. For BF and ADG, the top ranking three males and eight females were chosen. This represented an average selection intensity of approximately 20%/generation for both males and females. These EBV for selected individuals were averaged within sex each generation and deviated from the average BLUP EBV for that sex-generation combination to obtain genetic selection differentials. Because LS is a sex-limited trait, half the selection differential for females was taken as an estimate of expected genetic gain. Thus, it was assumed that sons and daughters of selected females were retained as breeders. For BF and ADG, selection differentials for the two sexes were averaged to obtain an estimate of expected genetic gain. Estimates of expected genetic gain from each generation were then averaged to obtain an estimate of expected genetic gain for selection on EBV for various proportions of incorrect pedigrees in the data.

#### Results and Discussion

Spearman rank correlation coefficients of best ranking with rankings obtained from the three methods when errors in sire identification occurred are presented in Table 1. When pedigrees were correct, the greatest differences among rankings of animals by the three methods were found for ADG, where rank correlations with BLUP were .78 and .85 for SP and SI, respectively. The smallest differences among ranks were found for BF, and LS was intermediate. Given the heritability estimates for the three traits, this was expected. As heritability increases, more weight is given to an individual's phenotype than information from relatives with SI and BLUP, and rankings from the three methods should correspond to a greater degree for traits with high heritabilities than they would when heritabilities are low.

Increasing amounts of errors in record of paternity did not affect ranking of animals for

TABLE 1. SPEARMAN RANK CORRELATION COEFFICIENTS OF BLUP ESTIMATES OF BREEDING VALUE WITH THOSE FROM SP AND SI FOR LITTER SIZE, BACKFAT AND ADG WHEN DATA CONTAIN SIRE PEDIGREE ERRORS

Trait/method	% Error				
	0	5	10	15	20
Litter size					
SP <sup>a</sup>	.81	.81	.81	.81	.81
SI <sup>b</sup>	.88	.87	.87	.86	.86
BLUPID <sup>c</sup>	1.00	.99	.97	.96	.95
Backfat					
SP	.93	.93	.93	.93	.93
SI	.96	.96	.96	.96	.95
BLUPID	1.00	.99	.99	.99	.98
Avg daily gain					
SP	.78	.78	.78	.78	.78
SI	.85	.85	.84	.84	.83
BLUPID	1.00	.98	.97	.95	.93

<sup>a</sup>Selection on phenotype.

<sup>b</sup>Selection index.

<sup>c</sup>BLUP analysis on incorrect data.

selection by SP. An individual's phenotype is the only criterion for ranking with SP, and errors in pedigree would have no effect on estimates of breeding value. Estimates of breeding value from SI and BLUP are affected by errors in paternity, as can be seen by a reduction in the correlation coefficients with increasing frequency of errors. A reduction in the correlation coefficient indicates that the method is ranking animals differently from the rankings by BLUP on correct data, and the greater the reduction the greater the deviation. For all three traits, BLUP was affected to the greatest degree by errors in paternity. Correlations for LS went from 1.00 to .95, those for BF went from 1.00 to .98 and those for ADG from 1.00 to .93. Estimates from SI were affected by errors in sire pedigree, but to a lesser degree than those from BLUP. Correlations for LS went from .88 to .86, those for BF went from .96 to .95 and those for ADG went from .85 to .83. Although BLUP was affected to the greatest degree by errors in paternity, at 20% incorrect sire identifications, the correlation coefficient for BLUP with BLUP on correct data was highest of the methods for all traits.

Ranking of animals from a BLUP analysis on incorrect data was more highly correlated with ranking by BLUP on correct data than rankings from the other methods. This suggests that, although BLUP takes into account

TABLE 2. SPEARMAN RANK CORRELATION COEFFICIENTS OF BLUP ESTIMATES OF BREEDING VALUE WITH THOSE FROM SP AND SI FOR LITTER SIZE, BACKFAT AND ADG WHEN DATA CONTAIN SIRE AND DAM PEDIGREE ERRORS

Trait/method	% Error				
	0	5	10	15	20
Litter size					
SP <sup>a</sup>	.81	.81	.81	.81	.81
SI <sup>b</sup>	.88	.86	.84	.83	.81
BLUPID <sup>c</sup>	1.00	.95	.91	.88	.85
Backfat					
SP	.93	.93	.93	.93	.93
SI	.96	.96	.95	.95	.94
BLUPID	1.00	.99	.98	.97	.96
Avg daily gain					
SP	.78	.78	.78	.78	.78
SI	.85	.84	.82	.82	.81
BLUPID	1.00	.97	.93	.91	.89

<sup>a</sup>Selection on phenotype.  
<sup>b</sup>Selection index.  
<sup>c</sup>BLUP analysis on incorrect data.

all relationships among relatives and can produce biased estimates of breeding value when errors in those relationships occur, BLUP will still rank animals closer to the correct rankings than will SI or SP. Estimates of BV for ADG by SI and BLUP were affected to the greatest degree by errors in paternity because ADG had the lowest estimate of heritability.

Spearman rank correlation coefficients of best rankings with rankings obtained from the three methods when errors in sire and dam identification occurred are presented in Table 2. Trends are similar to results in Table 1, but they differ in magnitude. Because both sire and dam errors occurred, methods taking relationships among animals into account, such as SI and BLUP, were more affected than when only sire errors occurred. This point is apparent for LS when 20% of the sire and dam pedigrees were incorrect. Relative to the best rankings, estimates based on SI were no better than those from SP. Estimates from BLUP also were affected to a greater degree than when

TABLE 3. REGRESSION COEFFICIENTS OF OFFSPRING PHENOTYPE ON PARENT'S ESTIMATED BREEDING VALUE (EBV) FOR LITTER SIZE, BACKFAT AND ADG WITH SIRE PEDIGREE ERRORS

Trait/method		% Error									
		0		5		10		15		20	
		$\hat{b}$	SE	$\hat{b}$	SE	$\hat{b}$	SE	$\hat{b}$	SE	$\hat{b}$	SE
Litter size											
SP <sup>a</sup>	EBV <sub>sire</sub>	-.56	.29	-.56	.29	-.56	.29	-.56	.29	-.56	.29
	EBV <sub>dam</sub>	.34	.14	.34	.14	.34	.14	.34	.14	.34	.14
SI <sup>b</sup>	EBV <sub>sire</sub>	-.44+	.26	-.43	.26	-.41	.26	-.39	.26	-.42	.26
	EBV <sub>dam</sub>	.30	.12	.29	.12	.29	.12	.30	.12	.30	.12
BLUPID <sup>c</sup>	EBV <sub>sire</sub>	-.10	.12	-.12	.12	-.10	.12	-.10	.12	-.12	.12
	EBV <sub>dam</sub>	.40	.12	.40	.12	.40	.12	.42	.12	.40	.12
Backfat											
SP	EBV <sub>sire</sub>	.60	.05	.60	.05	.60	.05	.60	.05	.60	.05
	EBV <sub>dam</sub>	.38	.05	.38	.05	.38	.05	.38	.05	.38	.05
SI	EBV <sub>sire</sub>	.60	.05	.60	.05	.61	.05	.61	.05	.61	.05
	EBV <sub>dam</sub>	.39	.05	.38	.05	.39	.05	.39	.05	.40	.05
BLUPID	EBV <sub>sire</sub>	.59	.04	.59	.04	.60	.04	.61	.04	.60	.04
	EBV <sub>dam</sub>	.41	.04	.40	.04	.40	.04	.41	.05	.41	.05
Avg daily gain											
SP	EBV <sub>sire</sub>	.66	.17	.66	.17	.66	.17	.66	.17	.66	.17
	EBV <sub>dam</sub>	.19	.17	.19	.17	.19	.17	.19	.17	.19	.17
SI	EBV <sub>sire</sub>	.50	.16	.53	.16	.53	.16	.55	.16	.58	.16
	EBV <sub>dam</sub>	.16	.15	.19	.16	.15	.15	.14	.16	.15	.16
BLUPID	EBV <sub>sire</sub>	.52	.14	.52	.14	.52	.13	.54	.13	.58	.14
	EBV <sub>dam</sub>	.31	.14	.33	.14	.28	.14	.26	.14	.27	.14

<sup>a</sup>Selection on phenotype.  
<sup>b</sup>Selection index.  
<sup>c</sup>BLUP analysis on incorrect data.

TABLE 4. REGRESSION COEFFICIENTS OF OFFSPRING PHENOTYPE ON PARENT'S ESTIMATED BREEDING VALUE (EBV) FOR LITTER SIZE, BACKFAT AND ADG WITH SIRE AND DAM PEDIGREE ERRORS

Trait/method		% Error									
		0		5		10		15		20	
		ĥ	SE	ĥ	SE	ĥ	SE	ĥ	SE	ĥ	SE
Litter size											
SP <sup>a</sup>	EBV <sub>sire</sub>	-.56	.29	-.56	.29	-.56	.29	-.56	.29	-.56	.29
	EBV <sub>dam</sub>	.34	.14	.34	.14	.34	.14	.34	.14	.34	.14
SI <sup>b</sup>	EBV <sub>sire</sub>	-.44+	.26	-.44	.26	-.39	.26	-.40	.26	-.33	.26
	EBV <sub>dam</sub>	.30	.12	.31	.12	.34	.12	.32	.12	.30	.12
BLUPID <sup>c</sup>	EBV <sub>sire</sub>	-.10	.12	-.12	.12	-.11	.12	-.15	.12	-.13	.12
	EBV <sub>dam</sub>	.40	.12	.42	.12	.45	.12	.40	.12	.36	.12
Backfat											
SP	EBV <sub>sire</sub>	.60	.05	.60	.05	.60	.05	.60	.05	.60	.05
	EBV <sub>dam</sub>	.38	.05	.38	.05	.38	.05	.38	.05	.38	.05
SI	EBV <sub>sire</sub>	.60	.04	.62	.05	.62	.05	.63	.05	.64	.05
	EBV <sub>dam</sub>	.39	.05	.39	.05	.39	.05	.40	.05	.40	.05
BLUPID	EBV <sub>sire</sub>	.59	.04	.61	.05	.61	.05	.62	.05	.62	.05
	EBV <sub>dam</sub>	.41	.04	.41	.05	.40	.05	.41	.05	.42	.05
Avg daily gain											
SP	EBV <sub>sire</sub>	.66	.17	.66	.17	.66	.17	.66	.17	.66	.17
	EBV <sub>dam</sub>	.19	.17	.19	.17	.19	.17	.19	.17	.19	.17
SI	EBV <sub>sire</sub>	.50	.16	.52	.16	.55	.16	.57	.16	.59	.16
	EBV <sub>dam</sub>	.16	.15	.15	.16	.16	.16	.17	.16	.21	.17
BLUPID	EBV <sub>sire</sub>	.52	.14	.54	.14	.57	.14	.59	.14	.56	.14
	EBV <sub>dam</sub>	.31	.14	.28	.14	.27+	.14	.22	.15	.18	.15

<sup>a</sup>Selection on phenotype.

<sup>b</sup>Selection index.

<sup>c</sup>BLUP analysis on incorrect data.

just sires were incorrect but, even when 20% of the sire and dam pedigrees were incorrect, BLUP ranked animals closer to best rankings than the other two methods.

Table 3 presents regression coefficients of offspring phenotype on parent's estimated breeding value. Progeny data were not used in estimation of breeding value. Thus, these regression coefficients indicate how well EBV predicted subsequent progeny performance. Deviations in values of these regressions from 1/2 indicate inadequacy of the model and biases in the estimates of heritability.

Negative estimates were obtained for regressions of daughter's phenotype on sire's EBV for LS. Because LS is a sex-limited trait, sire's EBV from SP and SI were calculated as half the paternal granddam's SP or SI EBV, respectively. The regression of daughter's phenotype on phenotype of paternal granddam was negative ( $-.047 \pm .026$ ;  $P > .07$ ); and, because sire's EBV was calculated as half the paternal granddam's EBV, a negative estimate of this regression was obtained. Regressions of daughter's phenotype on dam's EBV were

positive for all three methods, with BLUP having the highest value, but all were below .5. Values less than .5 would occur if heritability of litter size was overestimated. The magnitude of the deviation from .5 would depend on the amount of error in the estimate of heritability. This could explain, in part, why these regressions were less than .5. Surprisingly, pedigree errors did not affect regressions of daughter's litter size on EBV of dam when estimates were obtained by SI or BLUP.

For BF, the average of the regression coefficients from sire and dam EBV was close to .5, indicating adequacy of the model and accurate estimates of parameters. Errors in paternity had little effect on the regression of progeny's backfat on parent's EBV. For ADG, regressions on dam's EBV were less than regressions on sire's EBV, and average values were about .4. Again, errors in pedigrees of sires had very little effect on regressions.

Table 4 presents regression coefficients of offspring phenotype on parent's estimated breeding value when errors in sire and dam pedigrees occurred. Relative to the coefficients



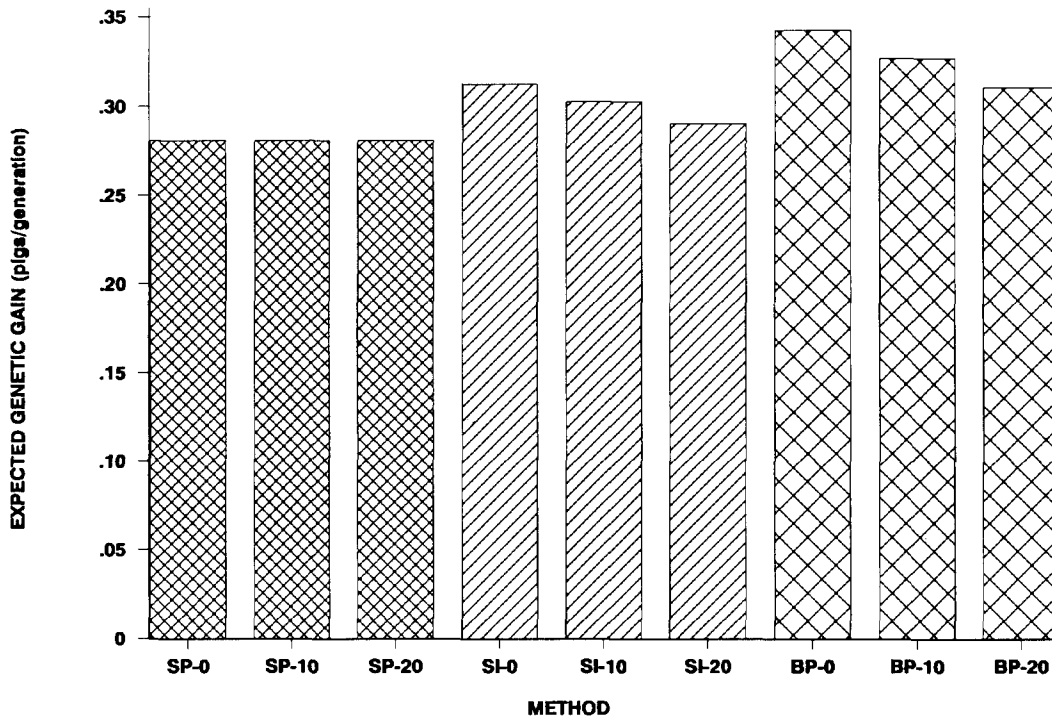


Figure 1. Expected genetic gain from selection on phenotype (SP), selection index (SI) and BLUP (BP) for number of fully formed pigs at birth (LS) when sire and dam pedigree errors have occurred in 0, 10 or 20% of the data.

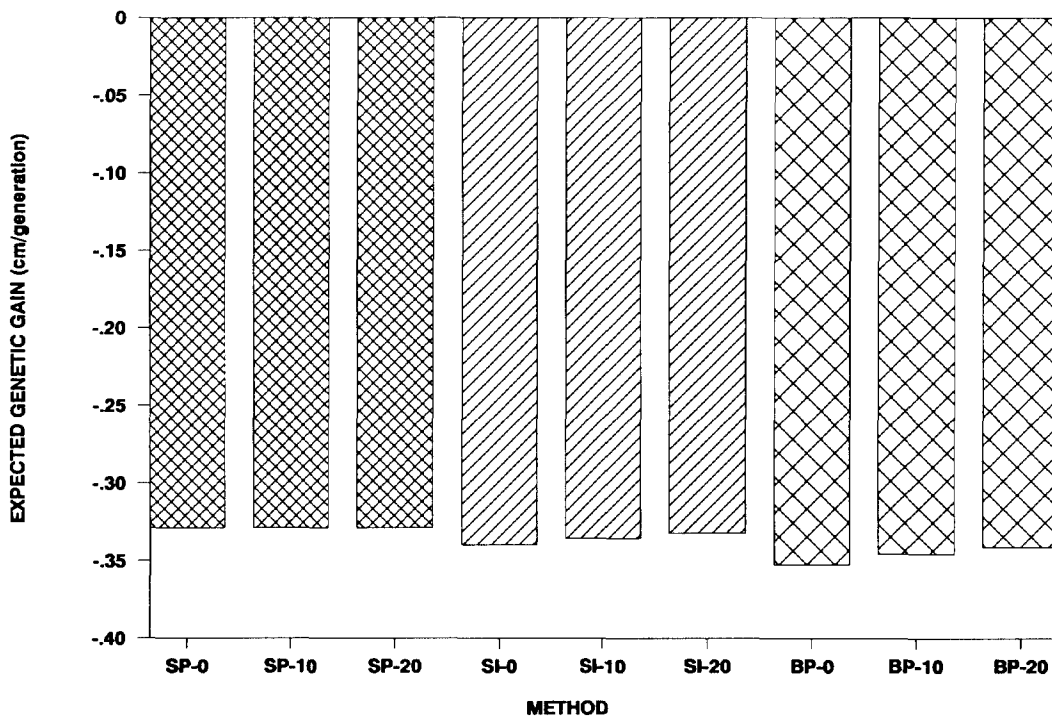


Figure 2. Expected genetic gain from selection on phenotype (SP), selection index (SI) and BLUP (BP) for backfat when sire and dam pedigree errors have occurred in 0, 10 or 20% of the data.

in Table 3, the addition of errors in dam pedigree to errors in the sire pedigree did not appear to greatly affect these regressions.

Figures 1, 2 and 3 illustrate expected genetic gain assuming that estimates from BLUP on correct data give the true breeding value of an animal. Figure 1 presents the expected genetic gain for LS in pigs/generation. For correct data, calculating breeding values using BLUP resulted in an advantage in expected genetic gain for litter size of 22 or 9.6% over SP or SI, respectively. Errors in pedigree did not affect expected genetic gain from selection on phenotype. Errors in pedigree did, however, reduce expected genetic gain for SI and BLUP. In comparing correct data to data with 20% of the sire and dam pedigrees in error, expected genetic gain using SI was reduced by 7.0%, and, using BLUP, was reduced by 9.3%. Even with a reduction of 9.3% in expected genetic gain using BLUP on incorrect data, BLUP gave a 10.5% advantage in expected genetic gain over SP.

Figure 2 presents expected genetic gain for BF in centimeters per generation. For correct

data, BLUP resulted in an advantage of 7.2 or 3.8% over SP or SI, respectively. Errors in pedigree in 20% of the data resulted in a 3.2 or 2.5% reduction in expected genetic gain for BF with BLUP and SI, respectively. Data with 20% sire and dam pedigree errors resulted in a 3.8% advantage for BLUP in expected genetic gain over SP, the method unaffected by errors in pedigree.

Expected genetic gains for ADG in kg/(d-generation) when sire and dam pedigree errors occurred are presented in Figure 3. The largest advantage of BLUP over the other two methods was for ADG; BLUP resulted in a 30.8 or 21.4% advantage over SP or SI, respectively. When 20% sire and dam pedigree errors occurred, expected genetic gain was reduced by 12.4 or 6.5% for BLUP or SI, respectively. Using BLUP to estimate breeding value for ADG when 20% of the data contained sire and dam pedigree errors resulted in an advantage of 14.6% in expected genetic gain over SP.

For data without introduced errors, results from this study agree with findings of Belon-

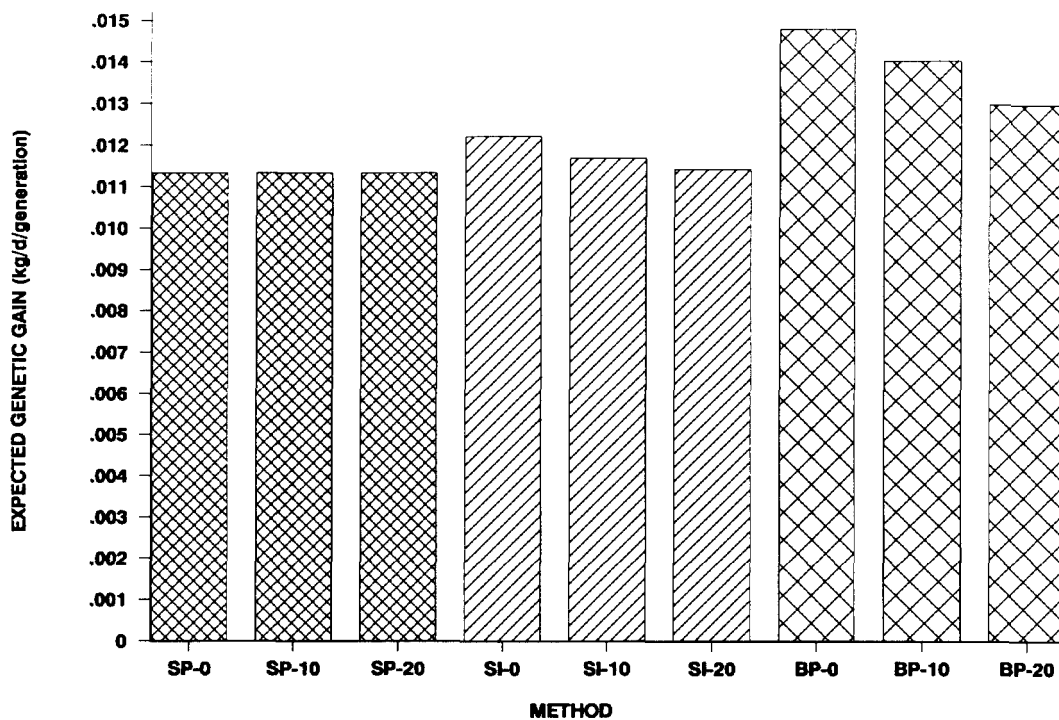


Figure 3. Expected genetic gain from selection on phenotype (SP), selection index (SI) and BLUP (BP) for ADG when sire and dam pedigree errors have occurred in 0, 10 or 20% of the data.

sky and Kennedy (1988). They simulated a closed 100-sow herd over a 10-yr period to compare selection on phenotype to selection on BLUP for traits with different heritabilities. Values chosen for heritability were .10, .30 and .60. They found advantages in using BLUP over selection on phenotype of 55% for  $h^2 = .10$ , 25% for  $h^2 = .30$  and 10% for  $h^2 = .60$ . Using data from an experimental research herd, where estimates of  $h^2$  were .13 for ADG, .18 for LS and .53 for BF, we found a similar trend. The advantage of BLUP over SP for correct data was 30.8% for ADG, 22% for LS and 7.2% for BF. For all traits, expected genetic gain from selection on BLUP was greater than expected genetic gain from selection on SP, but that advantage decreased as heritability increased.

For data without errors, results also agree with findings of Mabry et al. (1987). They compared breeding values estimated by BLUP to individual phenotype from data involving 3,999 boars in 24 central test stations. They found a rank correlation of .84 and .81 for BF and ADG, respectively, when the analysis was across sale group and reference sires were included in the BLUP analysis. Methods were ranking animals differently, in agreement with our findings illustrated in Tables 1 and 2. Similarly, Carlson et al. (1984) found that prediction error variances were lower when breeding values of centrally tested boars were estimated by mixed model procedures rather than by phenotypic deviations from group means.

Keele et al. (1988) compared BLUP to selection on phenotype and a selection index based on individual phenotype plus full- and half-sib averages for days to 100 kg and backfat using 203,869 records from the Nebraska SPF Swine Accrediting Agency. They found that the correlation between a sire's EBV and its progeny average was 33% larger for days to 100 kg and 44% larger for backfat when BLUP was used to estimate breeding value than when individual phenotype was used. Although we did not find as large an advantage for using BLUP over SP for BF, expected genetic gain was 7.2% higher for BLUP than for SP.

When errors in pedigree occurred, methods that made the most use of information from relatives to calculate EBV were affected to the greatest degree. The greatest reduction in expected genetic gain due to pedigree errors

was found when breeding values for ADG were estimated by BLUP when 20% of the sire and dam pedigrees were in error. Even with this reduction of 12.4% relative to correct data, expected genetic gain from using BLUP for ADG on incorrect data remained 14.6% larger than expected genetic gain from using SP to improve ADG. This would indicate that, for swine field data, BLUP is a very robust procedure for estimating breeding values.

#### Implications

Although estimation of breeding values by mixed model procedures is affected to a greater degree by pedigree errors than the other two methods investigated, selection of swine using best linear unbiased predictions of breeding values still would improve response to selection over use of selection on phenotype or a selection index. This assumes that levels of pedigree errors in field data are less than levels investigated in this study. If swine pedigrees are less accurate than the scenarios investigated, there would be a point at which expected genetic gain from selection on breeding values estimated with relatives' records would be inferior to selection on phenotype. Therefore, in swine recording programs, accuracy in recording pedigrees must be stressed to reap the maximum benefit from mixed model procedures to estimate breeding values.

#### Literature Cited

- Belonsky, G. M. and B. W. Kennedy. 1988. Selection on individual phenotype and best linear unbiased predictor of breeding value in a closed swine herd. *J. Anim. Sci.* 66:1124.
- Carlson, J. P., L. L. Christian, M. I. Rothschild and R. L. Willham. 1984. An evaluation of four procedures to rank centrally tested boars. *J. Anim. Sci.* 59:934.
- Cunningham, P. J., M. E. England, L. D. Young and D. R. Zimmerman. 1979. Selection for ovulation rate in swine: Correlated response in litter size and weight. *J. Anim. Sci.* 48:509.
- Henderson, C. R. 1963. Selection index and expected genetic advance. In: W. D. Hanson and H. F. Robinson (Ed.) *Statistical Genetics and Plant Breeding*. pp 141-193. National Academy Press, Washington, DC.
- Henderson, C. R. 1973. Sire evaluation and genetic trends. *Proc. of the Animal Breeding and Genetics Symposium in Honor of Dr. J. L. Lush*. Am. Soc. Anim. Sci. and Am. Dairy Sci. Assoc., Champaign, IL.
- Henderson, C. R. 1988. Exact prediction error variances for full model computed from reduced model. *J. Dairy Sci.* 71:3128.
- Johnson, R. K., Dwane R. Zimmerman and R. J. Kittok.

1984. Selection for components of reproduction in swine. *Livest. Prod. Sci.* 11:541.
- Keele, J. W., R. K. Johnson, L. D. Young and T. E. Socha. 1988. Comparison of methods of predicting breeding values of swine. *J. Anim. Sci.* 66:3040.
- Mabry, J. W., L. L. Benyshek, M. H. Johnson and D. E. Little. 1987. A comparison of methods for ranking boars from different central test stations. *J. Anim. Sci.* 65:56.
- Schaeffer, L. R. 1986. Pseudo expectation approach to variance component estimation. *J. Dairy Sci.* 69:2884.
- Van Vleck, L. D. 1970a. Misidentification in estimating the paternal sib correlation. *J. Dairy Sci.* 53:1469.
- Van Vleck, L. D. 1970b. Misidentification and sire evaluation. *J. Dairy Sci.* 53:1697.
- Zimmerman, D. R. and P. J. Cunningham. 1975. Selection for ovulation rate in swine: Population, procedures and ovulation response. *J. Anim. Sci.* 46:937.