

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Faculty Papers and Publications in Animal  
Science

Animal Science Department

---

October 1991

## EFFECTS OF SELECTION OF DATA ON ESTIMATES OF BREEDING VALUES BY THREE METHODS 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: <https://digitalcommons.unl.edu/animalscifacpub>



Part of the [Animal Sciences Commons](#)

---

Long, T. E.; Johnson, R. K.; and Keele, J. W., "EFFECTS OF SELECTION OF DATA ON ESTIMATES OF BREEDING VALUES BY THREE METHODS FOR LITTER SIZE, BACKFAT, AND AVERAGE DAILY GAIN IN SWINE" (1991). *Faculty Papers and Publications in Animal Science*. 50.

<https://digitalcommons.unl.edu/animalscifacpub/50>

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 SELECTION OF DATA ON ESTIMATES OF BREEDING VALUES BY THREE METHODS 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 was calculated based on individual phenotype (SP), an index of individual phenotype and full- and half-sib family averages (SI), or Best Linear Unbiased Prediction (BLUP). Traits considered were litter size (LS), backfat (BF), and ADG. Estimated breeding values were calculated using all data and after deletion of the poorest 5, 10, 15, or 20% of the records for BF and ADG, or 4.8, 8, 13, or 21% of the records for LS. When all data were used, expected genetic gain from BLUP was greater than for SP by 22, 7, and 31% and greater than for SI by 10, 4, and 21% for LS, BF, and ADG, respectively. Expected genetic gain was 4, 0, and 3% lower for LS, BF, and ADG, respectively, for selection on breeding values estimated by SI after the poorest 20% of the records were deleted compared with selection on estimates by SI using all the data. Genetic gain using BLUP on data with the poorest 20% of the records deleted was reduced by 5, 2, and 8% for LS, BF, and ADG, respectively, compared with genetic gain using BLUP on all the data. The advantage in genetic gain of BLUP, with 20% of the poorest records deleted, over SP was 15, 5, and 21% for LS, BF, and ADG, respectively. Although BLUP is affected to a greater degree by deletion of records than is SP or SI, selection of swine using BLUP on field data would improve response to selection over the use of SP or SI. Key Words: Pigs, Selection Index, BLUP, Selection, Genetic Gain

J. Anim. Sci. 1991. 69:2787-2794

## Introduction

With field data, genetic analyses of selected records are common. Testing space may be limited, slowly growing pigs are often removed before the end of the test period, and sows with small litters may be culled with no recording of their reproductive performance.

Several researchers (Henderson, 1975; Polak and Quaas, 1981; Goffinet, 1983; Gianola et al., 1988; Fernando and Gianola, 1989; Im et al., 1989) have considered the theoretical effects of selection of records on predictions of

breeding values and have suggested alternative ways of addressing this problem. Many different methods of genetic evaluation are currently used by various swine breeding organizations in the United States, but most of these methods do not account for selection of records to be analyzed. Knowledge of the magnitude of the effect of selection of records on estimates of breeding values would be useful in assessing current procedures for genetic evaluation of swine. The purpose of this study was to evaluate the effects of selection of records on estimates of breeding values calculated by three methods for litter size, backfat, and ADG. Methods evaluated were selection on own phenotype, an index of own plus full- and half-sib records, and a mixed-model procedure (Henderson, 1973).

## Materials and Methods

Data and traits investigated, number of fully formed pigs at birth (LS), backfat probe

<sup>1</sup>Journal Paper No. 9012 of the Nebraska Agric. Exp. Sta.

<sup>2</sup>Anim. Genet. and Breeding Unit, Univ. of New England, Armidale, NSW, 2351, Australia.

<sup>3</sup>Anim. Sci. Dept., Univ. of Nebraska.

<sup>4</sup>Roman L. Hruska U.S. Meat Anim. Res. Center, Clay Center.

Received May 29, 1990.

Accepted February 15, 1991.

adjusted to 90 kg live weight (BF), and ADG from weaning at 28 d to 90 kg (ADG), were the same as those described in a previous paper on the effects of errors in pedigree on estimates of breeding values in swine (Long et al., 1991). The three methods of estimating breeding values (EBV), selection on phenotype (SP), selection index (SI), and the mixed-model method of Henderson (1973), also were previously described. For SP, EBV was calculated as the product of heritability and the deviation of an individual's phenotype from its contemporary group mean. The individual's record and the average of its full- and half-sibs, all expressed as deviations from the contemporary group mean, were used to calculate estimates of breeding values by SI. The mixed-model procedure was an animal model that also included environmental effects of litters. For the rest of this paper the acronym BLUP (best linear unbiased prediction) will be used to represent the mixed-model method.

Data used were from the University of Nebraska Gene Pool population, a 14-breed composite population formed in 1965. Two lines, one selected for increased ovulation rate and the other randomly selected, were maintained from 1967 to 1977. After two generations of random selection and mating, the line selected for ovulation rate was partitioned into three lines. One of these lines was continued with random selection, a second line was selected for decreased age at puberty for eight generations, and the third line was selected for increased litter size for eight generations. The control line started in 1967 was maintained with random selection to the end of the experiment. Complete descriptions of these selection experiments and responses to selection are given by Lamberson et al. (1991). Records analyzed were those from all lines for the period beginning in 1967, the base generation for initiation of the selection experiments, to 1987.

The complete data set for LS contained records of 2,099 individuals, 736 sires and 1,393 dams. There were 1,953 individual BF records representing 321 sires and 688 dams, and ADG was recorded on 2,077 pigs from 321 sires and 689 dams.

Variance components were estimated by the pseudo-expectation approach (Schaeffer, 1986). This method was chosen because it is computationally less demanding than other procedures such as REML, and estimates

obtained with it had less bias and smaller SE than those from Henderson's Method 3 for these data (Keele et al., 1991). Estimates obtained by the pseudo-expectation approach are biased by selection (Van Raden and Jung, 1988), but no selection for ADG or BF was practiced in any of the lines. Selection was for LS in one of the four lines, but records of pigs from this line composed a small portion of the data, and, when variance components for LS were obtained with these data using REML, the estimates were the same as those from the pseudo-expectation method (L. D. Van Vleck, personal communication). Heritability estimates were .18, .53, and .13, and estimates of common environmental effects ( $c^2$ ) were .002, .07, and .16 for LS, BF, and ADG, respectively.

Selection of data for genetic evaluation of LS was accomplished by deleting records of pigs with 4, 5, 6, or 7 pigs/litter. This resulted in deletion of 4.8, 8, 13, or 21% of the data. For genetic evaluation of ADG, records of the slowest gaining 5, 10, 15, or 20% of the pigs were deleted each generation. Records for BF were deleted for the fattest 5, 10, 15, or 20% of the pigs each generation. If a pig whose record had been deleted was selected as a parent and produced progeny, that pig's identification was retained in the data to maintain relationships, but its record was not used in any subsequent calculations of breeding value.

Backfat also was analyzed after records of pigs with low ADG were deleted. Often in testing programs, slowly growing pigs are not probed for backfat. Deleting these pigs before estimation of breeding values would affect EBV for BF because ADG and BF are genetically correlated (average value of .22; Stewart and Schinckel, 1990). The effect of this selection might be larger when information from relatives is used to estimate breeding values.

Expected genetic gains were calculated for selections based on breeding values estimated by each method (Long et al., 1991). We assumed the breeding values estimated by BLUP using all data were the best estimates of true breeding values. Animals were ranked within generation and a fixed number (approximately 20%) were selected. For animals selected by each method, their EBV that had been calculated by BLUP using all the data were obtained and deviated from the sex-

generation mean to obtain genetic selection differentials. These genetic selection differentials were averaged over generations for each sex. Half the selection differential of females was taken as expected genetic gain for LS. The average of male and female selection differentials was used as expected genetic gain for ADG and BF.

Methods of estimating breeding values and the effects of deleting poor records before analyses also were compared by expressing phenotypic records of offspring as a deviation from contemporary group averages and regressing these deviations on EBV of parents. The EBV of parents were those obtained by each method and at each level of selection of

records. Phenotypes of offspring were not used to estimate breeding values of parents.

### Results and Discussion

Table 1 presents regression coefficients of phenotype of offspring on EBV of parents. Negative estimates of regressions of daughter's phenotype on sire's EBV were obtained for LS by all three methods (Long et al., 1991). Regressions of daughter's phenotype on dam's EBV for LS tended to be reduced by selection of records and were not significantly different from zero for all three methods when the poorest 15 or 20% of the records were deleted. Regressions for BF were not affected by

TABLE 1. REGRESSION COEFFICIENTS OF OFFSPRING PHENOTYPE ON PARENTS' ESTIMATED BREEDING VALUES (EBV) FOR LITTER SIZE, BACKFAT, AND AVERAGE DAILY GAIN WHEN SELECTION OF DATA HAS OCCURRED

Trait and method	% of Poorest records deleted				
	0	5	10	15	20
Litter size <sup>a</sup>					
SP <sup>b</sup>					
EBV <sub>sire</sub>	-.56*	-.25	-.38	-.19	-.51
EBV <sub>dam</sub>	.34*	.31*	.33*	.33	.12
SI					
EBV <sub>sire</sub>	-.44	-.01	-.14	.09	-.46
EBV <sub>dam</sub>	.30*	.32*	.35*	.29	.05
BLUP					
EBV <sub>sire</sub>	-.10	.10	.06	.14	-.03
EBV <sub>dam</sub>	.40**	.38**	.40**	.38**	.22
Backfat					
SP					
EBV <sub>sire</sub>	.60**	.60**	.69**	.62*	.64**
EBV <sub>dam</sub>	.38**	.35**	.39**	.40**	.35**
SI					
EBV <sub>sire</sub>	.60**	.64**	.73**	.69**	.73**
EBV <sub>dam</sub>	.39**	.36**	.40**	.42**	.38**
BLUP					
EBV <sub>sire</sub>	.59**	.63**	.70**	.66**	.72**
EBV <sub>dam</sub>	.41**	.39**	.41**	.45**	.41**
Avg daily gain					
SP					
EBV <sub>sire</sub>	.66**	.92**	1.16**	1.22**	1.26**
EBV <sub>dam</sub>	.19	.30	.32	.42	.37
SI					
EBV <sub>sire</sub>	.50**	.64**	.83**	.92**	.88**
EBV <sub>dam</sub>	.16	.24	.25	.37	.19
BLUP					
EBV <sub>sire</sub>	.52**	.63**	.76**	.84**	.96**
EBV <sub>dam</sub>	.31*	.36*	.43*	.50*	.27

<sup>a</sup>Levels for litter size are 0, 4.8, 8, 13, and 21%.

<sup>b</sup>SP = own phenotype, SI = index of individual plus full- and half-sibs, and BLUP = best linear unbiased prediction.

\*P < .05.

\*\*P < .01.

selection of data as much as those for the other two traits. For ADG, the regression of phenotype of offspring on EBV of sire increased as a higher percentage of poor records were deleted. The largest increase occurred for SP.

The larger regressions of phenotype of offspring on EBV of parents for ADG, relative to the two other traits, might be due to the low estimate of heritability for ADG. The estimate of .13, obtained using all data, was used in all analyses. If poorly gaining pigs did poorly for environmental reasons, such as a poorly milking dam, illness, and so on, rather than for genetic reasons, the estimate obtained for heritability of ADG would be lower when all records were used than for data with the poorest records deleted. The expectation of the coefficient for the regression of phenotype of offspring on EBV of parents when BV is estimated from SP is  $1/2$  times the ratio of true heritability to the estimated heritability (Long et al., 1991). Estimated heritability is in the denominator; a low estimate would result in a regression coefficient greater than .5.

The large regressions of offspring phenotype on EBV of sire for ADG could result from the use of the estimate of heritability for ADG over the entire range of the data. If the relationship between phenotype and breeding value is curvilinear for ADG, deleting records from one tail of the distribution would affect this regression. If a large percentage of pigs at the lower end of the distribution had better breeding values than were predicted by their phenotypes, one would expect an increase in the regression of phenotype of offspring on EBV of sire when records of those pigs were deleted. These regressions compared to their expectations are as follows: SP with complete data,  $b_{O \cdot SIRE} EBV = .66 = 1/2h^2/\hat{h}^2 = .5h^2/.13$ ,  $h^2 = .172$ ; SP when poorest 10% of the records were deleted,  $b_{O \cdot SIRE} EBV = 1.16 = 1/2h^2/\hat{h}^2 = .5h^2/.13$ ,  $h^2 = .302$ ; SP when poorest 20% of the records were deleted,  $b_{O \cdot SIRE} EBV = 1.26 = 1/2h^2/\hat{h}^2 = .5h^2/.13$ ,  $h^2 = .328$ . These results suggest that there was a curvilinear relationship between phenotype and breeding

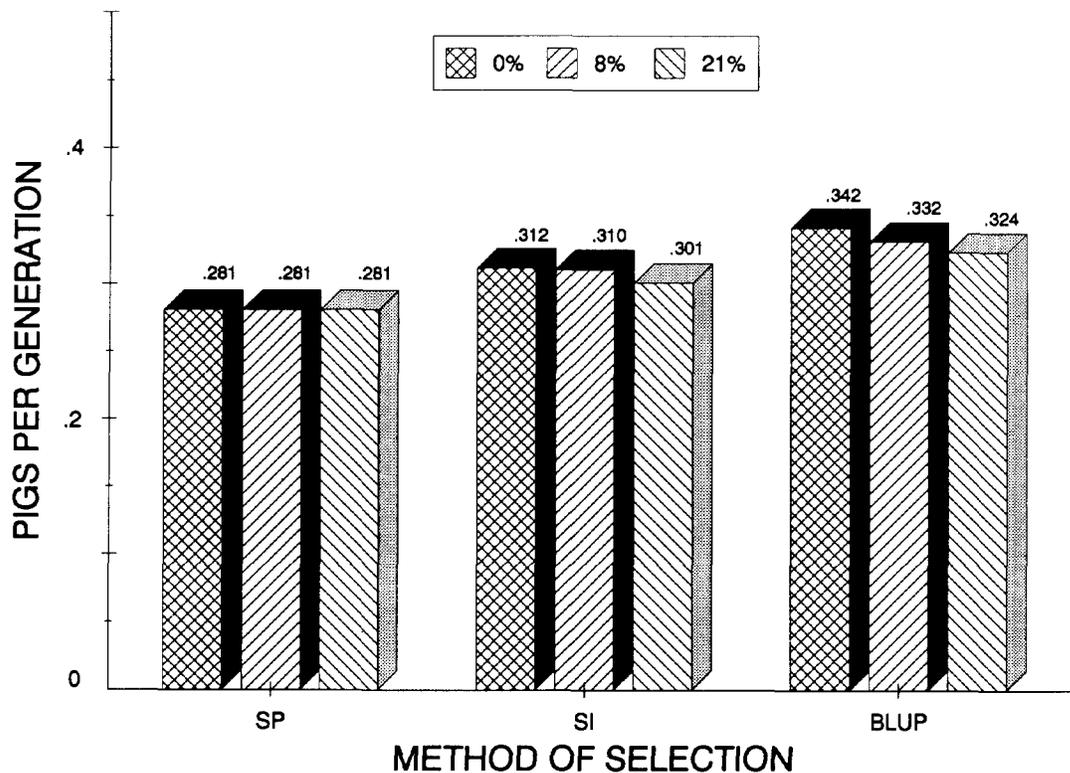


Figure 1. Expected genetic gain from selection on phenotype (SP), selection index (SI), and best linear unbiased prediction (BLUP) for litter size after deleting the poorest 0, 8, or 21% of the litter size records.

value for ADG. The regression of phenotype of offspring on EBV of dam for ADG was only significantly different from zero for BLUP when 0, 5, 10, and 15% of the records were deleted. Selection of data caused no discernible pattern on regressions of phenotype of offspring on EBV of parents for LS or BF. Thus, the relationship between phenotype and breeding value for both LS and BF was linear.

Table 2 presents regressions of phenotype of offspring on EBV of parents for BF when selection of records was based on ADG. Selection of records based on ADG had little effect on these regressions. Thus, there probably was a low genetic association between these two traits.

Figures 1, 2, and 3 illustrate expected genetic gain for selection intensity of about 20% when the assumption is made that BLUP on all data gives the true average breeding value. Figure 1 presents expected genetic gain for LS from selection for EBV calculated by

each method when 0, 8, and 21% of the poorest records were deleted before analysis. Calculating breeding values by BLUP using all records resulted in an advantage in expected genetic gain for litter size of 22 and 9.6% over SP and SI, respectively. Deleting the poorest records did not affect expected genetic gain when selection was on phenotype. Deleting poor records did, however, reduce expected genetic gain for SI and BLUP. Expected genetic gain was 3.7% less for SI and 5.4% less for BLUP after deleting 21% of the poorest records than when all data were used to estimate breeding values. However, even after deleting the poorest 21% of the records, genetic gain from selection on BLUP still was 15.3% higher than for SP, the method unaffected by deleting poor records.

Expected genetic gain for ADG from the three methods when records on 0, 10, or 20% of the most poorly gaining pigs were deleted before analysis is presented in Figure 2. Of the

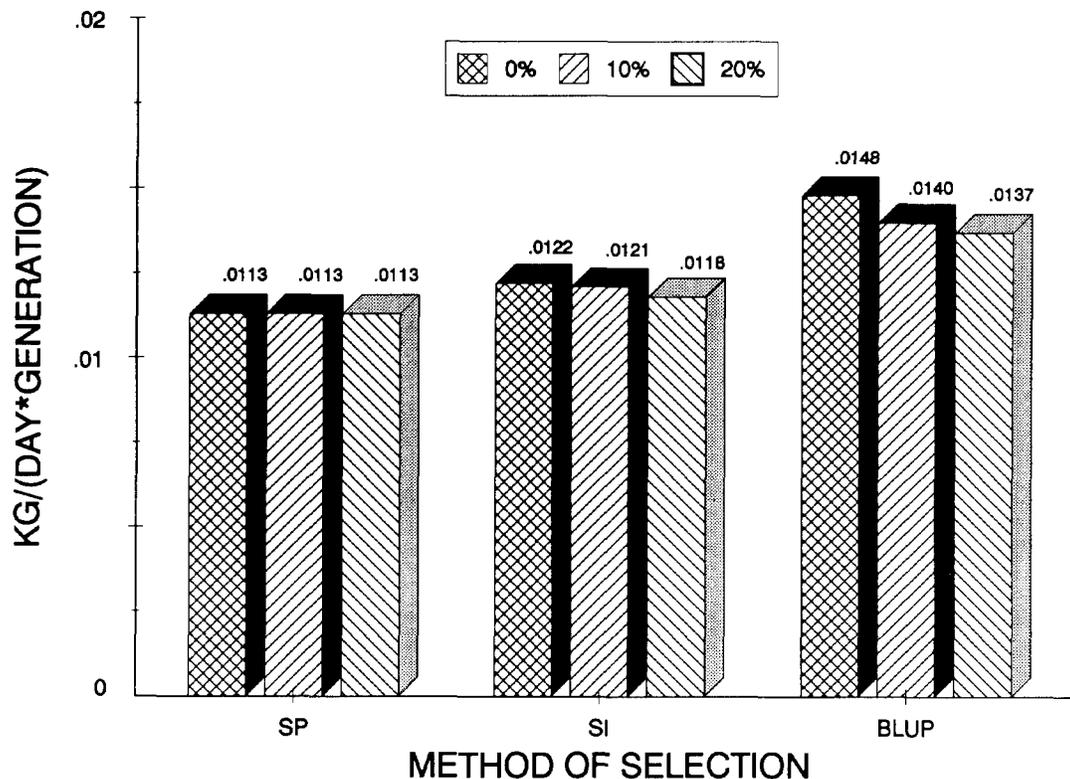


Figure 2. Expected genetic gain from selection on phenotype (SP), selection index (SI), and best linear unbiased prediction (BLUP) for average daily gain (ADG) after deleting the poorest 0, 10, or 20% of the ADG records.

TABLE 2. REGRESSION COEFFICIENTS OF OFFSPRING PHENOTYPE OF PARENTS' ESTIMATED BREEDING VALUE (EBV) FOR BACKFAT WHEN SELECTION OF DATA WAS BASED ON AVERAGE DAILY GAIN

Trait and method	% of Poorest records deleted				
	0	5	10	15	20
Backfat	6				
SP <sup>a</sup>					
EBV <sub>sire</sub>	.60**	.57**	.56**	.56*	.55**
EBV <sub>dam</sub>	.38**	.37**	.38**	.36**	.34**
SI					
EBV <sub>sire</sub>	.60**	.58**	.58**	.58**	.56**
EBV <sub>dam</sub>	.39**	.37**	.38**	.36**	.36**
BLUP					
EBV <sub>sire</sub>	.59**	.57**	.58**	.57**	.55**
EBV <sub>dam</sub>	.41**	.40**	.42**	.39**	.38**

<sup>a</sup>SP = own phenotype, SI = index of individual plus full and half sibs, and BLUP = best linear unbiased prediction.

\*\*P <.01.

three traits, the largest advantage of BLUP over the other two methods was for ADG when all records were used. Selection on EBV estimated by BLUP resulted in a 30.8 and 21.4% advantage in expected genetic gain over

selection based on estimates from SP and SI, respectively. Deleting the poorest 20% of the records decreased expected genetic gain by 7.5 and 3.2% for BLUP and SI, respectively. However, expected genetic gain from selection

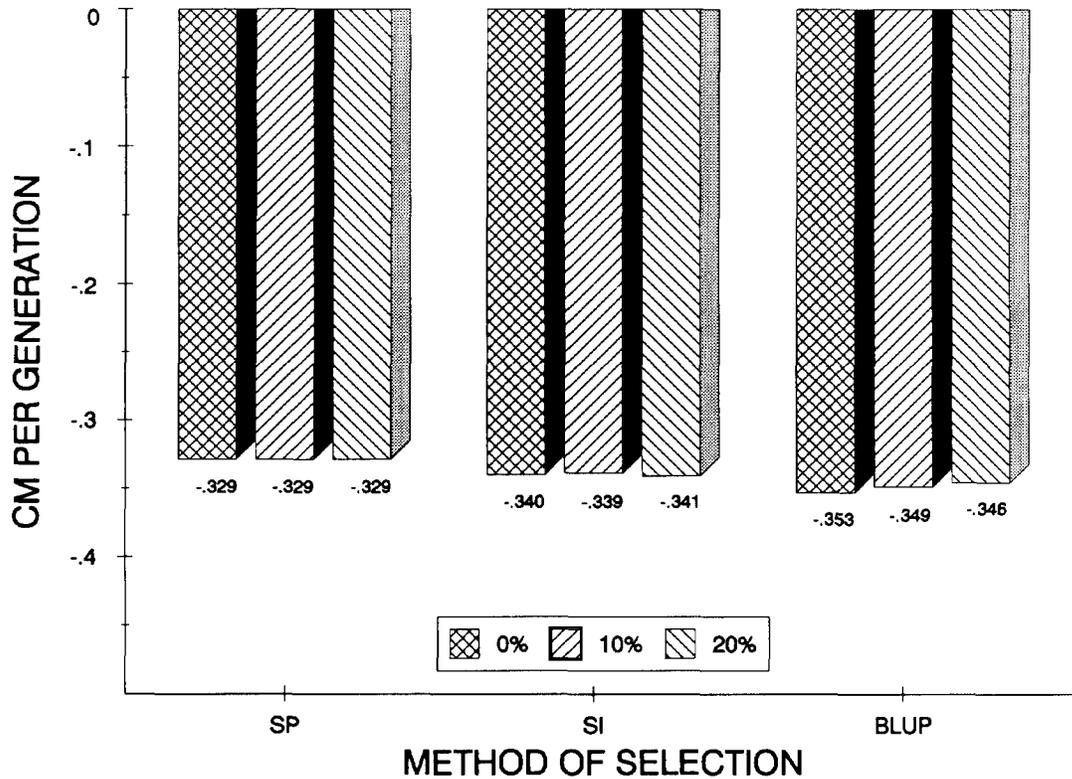


Figure 3. Expected genetic gain from selection on phenotype (SP), selection index (SI), and best linear unbiased prediction (BLUP) for backfat after deleting the poorest 0, 10, or 20% of the backfat records.

on BLUP EBV was still 21% higher than for selection on those from SP.

Figure 3 presents the expected genetic gain for BF from the three methods when the records of the fattest 0, 10, or 20% of the pigs were deleted before analysis. For complete data, the advantage for BLUP was 7.2 and 3.8% over SP and SI, respectively. Deleting the records of the fattest pigs before analysis resulted in little (1.9 and 0%) reduction in expected genetic gain for BLUP and SI. Selection on EBV estimated by BLUP after deleting 20% of the poorest records resulted in 5.2% more expected genetic gain than for selection on EBV estimated by SP.

Expected genetic gain in BF when records were selected on ADG is presented in Figure 4. Deleting records of the pigs with lowest ADG and then estimating breeding value for BF caused a reduction in rate of genetic gain for all three methods. It affected SP because some of the slowly growing pigs also had a low amount of backfat, but they were elimi-

nated as candidates for selection because of their slow growth. It had a slightly greater effect when records of relatives were used to estimate breeding value. Rate of genetic gain was reduced 6.1, 8.1, and 7.6% for SP, SI, and BLUP, respectively.

In comparisons of expected genetic gain, we assumed that EBV by BLUP using all data gave the true average breeding values. Of course, this assumption is true only if the model is correct and the genetic parameters used are correct. Genetic parameters were estimated from these same data; thus, they are the ones that best fit these data. It might seem that the results were destined to favor BLUP, but the same estimates of parameters were used for all three methods, so relative comparisons of expected genetic gain by the three methods are considered to be appropriate.

Regressions of phenotype of offspring on EBV of parents do not rely on the assumption that BLUP EBV give average true genetic values because records of these offspring were

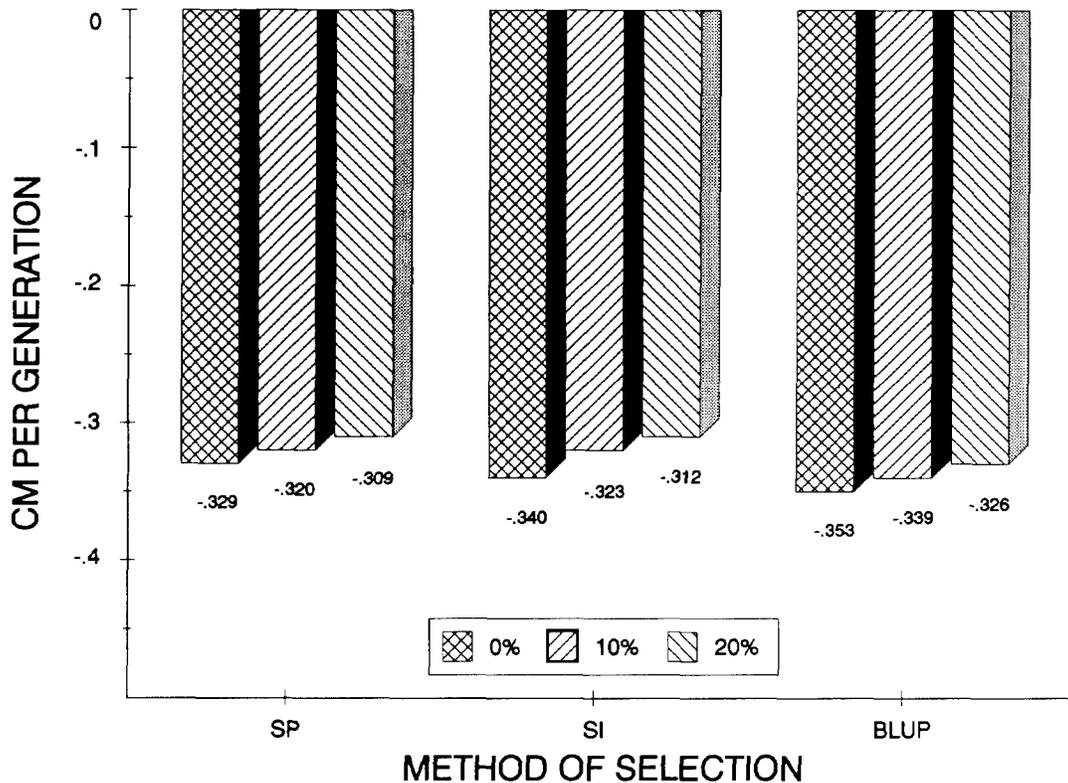


Figure 4. Expected genetic gain from selection on phenotype (SP), selection index (SI), and best linear unbiased prediction (BLUP) for backfat after deleting the poorest 0, 10, or 20% of the average daily gain (ADG) records.

not used to calculate EBV of parents. Thus, results in Tables 1 and 2 are unbiased and interpretations of them are valid.

#### Implications

Censoring data on performance will decrease the amount of genetic gain possible in swine when using either selection index or mixed-model procedures for genetic evaluations. Response to selection on best linear unbiased prediction estimates of breeding values was affected to a greater degree by censoring records than responses to selection on estimates obtained by selection index. Even so, selection of swine using best linear unbiased prediction on field data would improve response to selection over the use of selection on phenotype or selection index. This assumes that levels of selection of records are not greater in U.S. field data than levels investigated in this study. Producers should use as many data as possible to estimate breeding values to reap the maximum benefit from a method such as best linear unbiased prediction.

#### Literature Cited

- Fernando, R. L. and D. Gianola. 1989. Statistical inferences in populations undergoing selection and non-random mating. In: *Advances in Statistical Methods for Genetic Improvement of Livestock*. Springer-Verlag, New York.
- Gianola, D., S. Im and R. L. Fernando. 1988. Prediction of breeding value under Henderson's selection model: A revisit. *J. Dairy Sci.* 71:2790.
- Goffinet, B. 1983. Selection on selected records. *Genet. Sel. Evol.* 15:91.
- Henderson, C. R. 1973. Sire evaluation and genetic trends. In: *Proc. Anim. Breed. Genet. Symp. in Honor of Dr. Jay L. Lush*. pp 10-41. Am. Soc. Anim. Sci., Champaign, IL.
- Henderson, C. R. 1975. Best linear unbiased estimation and prediction under a selection model. *Biometrics* 32:69.
- Im, S., R. L. Fernando and D. Gianola. 1989. Likelihood inferences in animal breeding under selection: a missing-data theory view point. *Genet. Sel. Evol.* 21:399.
- Keele, J. W., T. E. Long and R. K. Johnson. 1991. Comparison of methods of estimating variance components in pigs. *J. Anim. Sci.* 69:(In press).
- Lamberson, W. R., R. K. Johnson, D. R. Zimmerman and T. E. Long. 1991. Direct responses to selection for increased litter size, decreased age at puberty or random selection following selection for ovulation rate. *J. Anim. Sci.* (In press).
- Long, T. E., R. K. Johnson and J. W. Keele. 1990. Effects of errors in pedigree on three methods of estimating breeding value for litter size, backfat and average daily gain in swine. *J. Anim. Sci.* 68:4069.
- Pollak, E. J. and R. L. Quass. 1981. Monte Carlo study of genetic evaluations using sequentially selected records. *J. Anim. Sci.* 52:257.
- Schaeffer, L. R. 1986. Pseudo expectation approach to variance component estimation. *J. Dairy Sci.* 69:2884.
- Stewart, T. S. and A. P. Schinckel. 1990. Genetic parameters for swine growth and carcass traits. In: L. D. Young (Ed.) *Genetics of Swine*. Publ. of the Tech. Committee of the North Central Regional Res. Project NC-103 ARS, USDA, Clay Center, NE.
- Van Raden, P. M. and Y. C. Jung. 1988. A general purpose approximation to Restricted Maximum Likelihood: The tilde-hat approach. *J. Dairy Sci.* 71:187.