COMPARISON OF GENETIC VALUES FOR REPRODUCTIVE PERFORMANCE IN BEEF CATTLE

by

KATHERINE ALISON DONOGHUE

(Under the Direction of J. Keith Bertrand and Ignacy Misztal)

ABSTRACT

The first objective of this study was to compare methods for handling censored fertility records for the trait days to calving (DC). Censored records were assigned penalty values on a within contemporary group basis under the first method (DCPEN). Under the second method (DCSIM), censored records were drawn from their respective predictive distributions. The PM of residual variances were significantly higher for DCPEN than DCSIM, especially at higher levels of censoring. Similar trends in variance components were observed in a study using field data for the trait. Little difference was observed between the two methods for correlations between true breeding values and posterior means of animal effects for sires, indicating that no major re-ranking of sires would be expected. The results from both studies suggest that either censored data handling technique could be successfully used in a genetic evaluation for DC.

The second objective of this study was to develop an evaluation for fertility that uses information from both natural service (NS) and artificial insemination (AI) matings. Firstly, mating records were used to examine the relationship between probability of calving to first insemination (CFI) in both types of data. Some differences were observed between PM for herd-year variance, which may be a reflection of the higher incidence of extreme category problem in the AI data. A high genetic correlation was observed between the two traits (0.821). This result, along with the observed lack of heterogeneity for additive variance, implies that an analysis of CFI with a common additive variance for AI and NS data could be undertaken.

In a separate study, the relationship between DC and two measures of fertility in AI data; CFI and calving success (CS); was examined. High genetic correlations were observed between DC-CFI (-0.681) and DC-CS (-0.751), indicating a strong, negative relationship between DC and both measure of fertility in AI data. The magnitude of the correlations between DC and CS/CFI suggest that it may be possible to use a multi-trait approach to the evaluation of artificial insemination and natural service data, and report one genetic value that could be used for selection purposes.

INDEX WORDS: Beef cattle, Female fertility, Genetic evaluation

COMPARISON OF GENETIC VALUES FOR REPRODUCTIVE PERFORMANCE IN BEEF CATTLE

by

KATHERINE ALISON DONOGHUE
B. Rur. Sc. (Hons), University of New England, Australia, 1997
M.S., The University of Georgia, 2001

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2003

COMPARISON OF GENETIC VALUES FOR REPRODUCTIVE PERFORMANCE IN BEEF CATTLE

by

KATHERINE ALISON DONOGHUE

Co-Major Professors:	J. Keith Bertrand Ignacy Misztal
Committee:	Romdhane Rekaya Jerome Baker Larry Benyshek Samuel Aggrey

Electronic Version Approved:
Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2003

DEDICATION

I would like to dedicate this dissertation to Mum, Dad, Andrew and Anna for all their love, support and encouragement throughout my graduate studies.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to Keith Bertrand for his friendship, guidance and assistance. It has been a pleasure and honour to work with him for the duration of my studies. I would also like to thank Ignacy Misztal for his assistance, and for always having an interesting observation that kept things in perspective! Thank you also to Romdhane Rekaya for his friendship and assistance, especially in the last six months. Also, many thanks to the remaining members of my committee; Jerry Baker, Larry Benyshek and Samuel Aggrey for their contributions.

Special thanks must go to many members of the Animal Breeding group at UGA, both past and present, who have made my experience so enjoyable. In particular, I would like to mention Robyn Sapp, John Mabry, Oseni, Del Little, Olga Ravagnolo, Andres Legarra, Monchai and Mike Kelley. I would also like to make a special thank you to Dr. Andra Nelson for all her assistance, and for her friendship over the years. Thanks also to Linda Heusner for her friendship and assistance. Also, thanks to my friends back home who have always supported me in my various endeavours.

I would like to express my appreciation to the Angus Society of Australia, and to Carel Teseling in particular, for providing the data for these studies. I would also like to acknowledge the advice and support of Drs. Hans Graser and David Johnston from AGBU. Appreciation is extended to Meat and Livestock Australia for generously supporting my studies in the USA.

Finally, very special thanks must to go to my family; without your love and support, I could never have finished. Thank you Mum and Dad for everything; for raising me to believe I could do anything that I wanted, and especially for encouraging me on the days that it was all too much! Thank you also to Andrew and Anna for your friendship, support and understanding.

TABLE OF CONTENTS

Page

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Fertility or reproductive performance is one of the most important components of production efficiency and genetic gain in beef production systems. A delay in conception due to poor fertility prolongs the calving interval, and causes a shift in calving pattern, which can lead to culling. However, reproductive traits in cattle are difficult to measure, report and interpret. This is particularly true for pasture mating situations, where information on females in extremely limited. In these situations, the only information readily available is whether or not a cow produces a calf, and when she calves.

Breeding value estimation for reproductive traits is difficult, in part because the expression of reproductive potential is often constrained by the management system. Reproductive data is of a complex nature, and is the culmination of many events that occur throughout the breeding season. Evaluation of genetic merit for reproduction requires information on the complete reproductive history of each animal, which is often unavailable. Thus, while genetic values for growth and carcass traits are reported in national genetic evaluations for most breeds, very few breeds report genetic values for fertility. In the past, correlated traits (such as scrotal circumference) have been used to indirectly select for female fertility.

Genetic values for the trait of days to calving are reported for many breeds in Australia using data from natural service matings. The trait is defined as the difference, in days, between the beginning of the breeding season and calving date for each breeding season. Censored animals (cows not calving) are assigned a projected value on a within breeding management group basis. The highest days to calving record within each breeding management group is identified, and a constant number of days (21 days) are added to this record to generate the projected value for all non-calvers. In the USA, genetic values for the trait of heifer pregnancy (or calving success) are currently reported for some North American breeds. These data is primarily composed of artificial insemination (AI) data from heifers.

The first objective of this study is to refine the evaluation of the trait of days to calving. An alternative approach to adding a fixed number of days to censored records would be to assume a truncated normal distribution for the uncensored records, and randomly draw from this truncated distribution to obtain a record for censored females on a contemporary group basis. This would allow the data to determine the value of censored records for each censored female. Comparison of these methods for handling censored records will be investigated in simulated and field data.

The second objective of this study is to develop an evaluation procedure for fertility in beef cattle that uses information from both natural service and AI matings. For breeds that have both natural service and AI data in their databases, it is natural to seek an evaluation that could combine information from both sources. The trait of heifer pregnancy may be defined for both natural service and artificial insemination (AI) data. However, unlike days to calving, it is unable to distinguish between early and late calvers in the calving season. Despite this obvious advantage of days to calving, only natural service mating data is used under the current definition of the trait. Thus, it appears that a multi-trait evaluation may be necessary if information from both natural service and AI matings are to be evaluated to produce one genetic value. The decision as to which genetic value to report would require further investigation of both traits, in order to determine which genetic value would predict reproductive performance with greater accuracy.

CHAPTER 2

REVIEW OF LITERATURE

Fertility or reproductive performance is one of the most important components of production efficiency and genetic gain in beef production systems. A delay in conception due to poor fertility prolongs the calving interval, and causes a shift in calving pattern, which can lead to culling. However, reproductive traits in cattle are difficult to measure, report and interpret. This is particularly true for pasture mating situations, where information on females in extremely limited. In these situations, the only information readily available is whether or not a cow produces a calf, and when she calves.

Breeding value estimation for reproductive traits is difficult, in part because the expression of reproductive potential is often constrained by the management system. Reproductive data is of a complex nature, and is the culmination of many events that occur throughout the breeding season. Evaluation of genetic merit for reproduction requires information on the complete reproductive history of each animal, which is often unavailable. Thus, while genetic values for growth and carcass traits are reported in national genetic evaluations for most breeds, very few breeds report genetic values for fertility. In the past, correlated traits (such as scrotal circumference) have been used to indirectly select for female fertility. Currently, genetic values for three measures of female fertility are being reported in national genetic evaluations; days to calving, stayability and heifer pregnancy.

Calving date/Days to calving

Calving interval has been used as the preferred measure of reproduction in dairy cattle. However, because a fixed breeding season is generally used in beef herds, calving interval has limited value as a selection criterion. Calving dates are generally available in field data, and their use requires minimal modification of existing performance programs. It is defined as the day of the year on which the cow calves, and allows comparison between cows when breeding is of the same duration, and starts on the same date. In early studies, calving date was found to be preferable to the alternative measure of calving interval. Later studies, which included the records of open cows in analyses, found calving date to be heritable, as well as having a clear economic interpretation.

In an early study, Bourdon and Brinks (1983) demonstrated the superiority of calving date over calving interval. In their study, calving interval was more susceptible to the bias caused by the use of a fixed breeding season, due to its strong dependence on previous calving date. In general, cows that calve early in the season will experience an ample postpartum period before the breeding season. As a rule, they will rebreed early, but are unable to register a subsequent calving interval of less than 365 days (Bourdon and Brinks, 1983). Cows that calve later in the season, however, have a shorter period between calving and breeding, and, therefore, the opportunity to record a shorter calving interval. In their study, calving interval decreased 0.86 days and calving date was delayed 0.11 days for each 1-day delay in previous calving date. The authors also noted that an additional advantage of calving date was its clearly identifiable economic value; calves born earlier in the calving season weigh more at weaning; while the economic interpretation of calving interval is difficult.

Numerous studies have addressed the issue of which parity should be used to measure calving date. Meacham and Notter (1987) used a sire model to estimate genetic parameters for calving date, calving interval and percent return using first and second calving records. All traits had a highly significant relationship with calving ease score recorded at first calving. Heritability estimates pooled across herds for calving interval, first and second calving dates and percent return were 0.04, 0.17, 0.07 and 0.11, respectively, and the estimated genetic correlation between first and second calving dates was 0.66. The authors noted that the lower heritability estimate for second calving date might reflect culling of open cows before the second calving, or real changes in the magnitude of genetic and environmental variation. The authors concluded that useful levels of genetic variation existed for first calving date, and that this trait could be used in sire selection as a measure of daughter's reproductive ability.

In order to make the best use of the data available for reproductive performance, information for open cows must be included in the evaluation. Notter (1988) noted that if data on open cows is ignored, the most genetically inferior, and possibly most informative, animals are ignored. Therefore, if sires differ markedly in the frequency of open daughters, consideration of open cows may be required to accurately estimate true sire difference in daughter's fertility.

Notter and Johnson (1988) obtained genetic parameter estimates for calving date with records for open cows included in the analysis using simulated data. They proposed a procedure using threshold theory to calculate penalties for open cows. Observed calving dates (CD) for cows that calved were transformed as $W = \ln(CD+1)$ to normalize the data, and calving dates for open cows were projected by considering cows that didn't calve to represent the upper tail of a truncated normal distribution of the transformed calving date. The authors found no carryover effects of prior calving date in cows calving within the first 21 days of the breeding season, but for cows calving after day 21, each 1-day increase in calving date was associated with an increase of 0.69 days in the next calving date. Adjustment for previous calving performance in this study reduced repeatability estimates of calving date from 0.26 to 0.24, while heritability estimates remained constant at 0.125, and the correlation between actual and adjusted calving date was 0.95. The authors suggested these results indicate that even though significant transient environmental effects existed between adjacent calving dates, adjustment for these effects didn't greatly affect overall rankings of females. Correlations between mean calving date and ability to conceive were found to be consistently higher for actual calving date than for the transformed data. The authors also found that selection against late calving was more effective than selection for early calving in identifying cows with genetic potential to conceive. The authors concluded that under pasture mating, unbiased estimates of ability to conceive cannot be obtained, but selection based on observed calving date yielded acceptable estimates, provided that open cows were included in the evaluation.

Buddenberg et al. (1990) compared estimates of variance components obtained from excluding and including records of open cows. Open cows were assigned a value based on the projected mean calving date of the open cows in an unrestricted breeding season, as described by Notter and Johnson (1988). Data were transformed as outlined by Notter and Johnson (1988), and the projected mean calving date for open cows was then obtained separately for each year based on the actual data and percentages of open cows for that year. Heritability estimates were obtained as paternal half-sib correlations. In general, the proportion of variance due to service sire and sire of dam increased when open cows were included in the analysis. Heritability estimates (open records excluded) for first-calf, second-calf and mature animals were 0.20, 0.04 and 0.03, respectively. The corresponding estimates for open records included were 0.39, 0.13 and 0.00, respectively, and confirmed that estimates from data excluding open cows are biased downward. The authors suggested that the lower heritability estimates for older animals was most likely the result of culling open cows each year. Service sire was the largest source of variation in calving date in both data sets, and sire of dam accounted for only a small portion (<10%) of variation. As a result of culling open cows, variation associated with service sire and sires of dams generally decreased with age. The authors concluded that attention should be given to selection against late calving date of first-calf heifers, and that the advantages of this selection would be lower birth weights, less dystocia and more recovery time between calving and breeding.

Days to calving has been investigated by researchers in Australia, and gives the same information as calving date when the cows to be compared went into breeding on the same day. Meyer et al. (1990) compared calving rate, number of calves, calving success and days to calving as measures of reproductive performance in Australian beef cattle. Calving rate was defined as the number of calves a cow produced divided by the number of opportunities to do so; calving success was scored as 0 (non-calver) and 1(calver); and days to calving was calculated as the difference in days between the beginning of the breeding season and calving date, for each breeding season. Cows not calving were assigned a predicted value, derived from threshold theory, as suggested by

Notter and Johnson (1988). Days to calving was analyzed both as observed, and transformed to logarithmic values, while calving success was analyzed without adjustment for the categorical nature of the trait. Heritability estimates for calving rate for Zebu crosses, Herefords and Angus were 0.17, 0.07 and 0.02, respectively. Heritability estimates for days to calving for Zebu crosses, Herefords and Angus were 0.09, 0.05 and 0.08, respectively. Transformation to log scale had practically no effect on estimates or on the predicted difference between calvers and non-calvers. Heritability estimates for calving success for Zebu crosses, Herefords and Angus were 0.08, 0.08 and 0.02, respectively. The authors concluded that days to calving appeared the most suitable trait for incorporation into genetic evaluation, as it is readily measurable under pasture conditions, and allows information on all cows to be included in the analysis. As well, the authors noted that the expected deviation of the distribution from normality for this trait would be considerably less than for other traits considered, which would allow the application of standard methods of genetic evaluation for analyses.

Johnston and Bunter (1996) demonstrated that days to calving, as defined by Meyer et al. (1990), was also a suitable measure of reproductive performance in a large field data set. Cows with open records were assigned a projected value on a within breeding management group basis. The highest days to calving record within each breeding management group was identified, and a constant number of days (21 days) were added to this record to generate the projected value for all non-calvers. The procedure proposed by Notter and Johnson (1988) to create penalty records was also considered. However, the procedure was deemed unsuitable because some of the predicted days to calving records for non-calvers were less than actual days to calving records. Calving success, scored as 0 (non-calvers) and 1 (calvers), was included in a bivariate analysis with days to calving, and was analyzed without any adjustment for the categorical nature of the trait. The genetic correlation estimate between the traits was –0.97, and the authors concluded that selecting for days to calving would be the same as selecting for calving success, with the added benefit of being able to distinguish between early and late calvers. Heritability estimates for days to calving in the first and second parities were 0.10 and 0.11, respectively, while repeatability and heritability estimates of 0.25 and 0.12 were obtained for a repeatability model. The genetic correlation between days to calving in the first and second parities was 0.85, supporting the use of a repeatability model. Genetic correlations between days to calving in the first parity and growth traits were generally unfavorable but not significantly different from zero, and thus the authors concluded that direct selection on reduced days to calving would be required to improve the trait.

Once a trait has been deemed suitable for incorporation into national genetic evaluations, the nature of the relationship between the trait and other traits is of primary interest. Meyer et al. (1991) used a subset of the data used by Meyer et al. (1990) to investigate covariances between days to calving, growth traits and male fertility traits. A weak but consistently favorable association (-0.30) was found between scrotal circumference and days to calving, while serving capacity and days to calving were found to be unrelated in this study. There seemed to be little favorable genetic association between growth and female fertility in the temperate breeds. Estimates of the genetic correlations were larger in Zebu crosses for yearling (-0.36) and weaning weight (-0.66). The authors failed to find any unfavorable genetic correlations between growth and days to calving, and concluded that joint selection for fertility and growth should improve genetic potential in both.

Rege and Famula (1993) studied factors affecting calving date in USA field data. They found that animals which as heifers calved in the first 21 days of the calving season had lower average subsequent calving dates, and gave birth to calves which were weaned earlier and had significantly heavier yearling weights than those that calved after the 42nd day of the season. Also, animals that calved late as heifers proceeded to calve later than initial early calvers in subsequent parities. Repeatability of calving date was estimated at 0.23, and heritability at 0.16. Genetic correlations between calving date and birth weight (-0.30), weaning weight (-0.05), postweaning gain (-0.64) and yearling weight (-0.60) were generally favorable. The nature of the relationship between calving date and maternal breeding value (BV) was also studied, with an increase in maternal BV associated with a delay in calving date. The authors suggested that there is an optimum level of milk production above which reproduction is jeopardized. Moreover, calving date of younger cows was more adversely affected by high maternal BV than was calving date of older cows, and late calving was associated more with high than with low milk production potential. The authors found that early initial calvers were superior to their late counterparts in subsequent reproductive performance. They concluded that since heifer calvings aren’t constrained by a previous calving; most heifers are bred and have the opportunity to calve early; differences in heifer performance are good indicators of genetic differences in calving date.

The study by Johnston and Bunter (1996) investigated the relationship between calving success and days to calving, but was unable to account for the categorical nature of calving success, due to computational limitations. Johnston et al. (2001) estimated the nature of the relationship between days to calving and calving success, using a new analytical procedure that accounted for the categorical nature of calving success. Days to calving and calving success were defined as described by Johnston and Bunter (1996), and only records from the first parity were retained for analysis for both traits. In addition, calving success records were only used from breeding management groups where variation existed, so that calving success records were removed for all animals in breeding management groups where all cows calved. Variance components were estimated using the Bayesian approach via the Gibbs sampler. Heritability estimates for days to calving and calving success (on the underlying scale) were 0.12 and 0.04, respectively, and the genetic correlation estimate between the two traits was –0.66. The authors suggested that, based on these results, selection for reduced days to calving would result in correlated increases in calving success. The correlation between estimated breeding values (EBV) for both traits was –0.96, indicating that shorter days to calving was favorably associated with an increased probability of a successful calving. The regression coefficient for days to calving EBV was –0.6 percent success/day. Thus, for each 1-day shorter days to calving EBV, there was a 0.6% increase in calving success EBV. The authors concluded that, from a selection point of view, days to calving and calving success are genetically similar, with the former having a higher heritability.

Various methods have been used to incorporate records of open cows in the analysis of calving date and days to calving (Notter and Johnson, 1988; Johnston and Bunter, 1996). An alternative approach would be to use survival analysis to evaluate reproductive traits. Such analyses could model days to calving with a hazard rate or probability of calving past time t, given the individual has not calved prior to t. Studies in dairy cattle have shown that survival analysis is useful for evaluating longevity (Ducrocq, 1994) and fertility traits such as days open (Eicker et al., 1996) but little research has been undertaken using the survival model for analysis of beef fertility traits. Although survival analyses offer several advantages over the linear model, e.g., better statistical modeling of censored data, the high computational requirements associated with applying these non-linear analyses hinders their use with an animal model and large data sets. Despite this drawback, survival analysis offers the potential for better evaluation of fertility traits in beef cattle in the future.

Stayability

Another trait of primary interest to the beef industry is the length of the productive life of females, sometimes termed “stayability”. Snelling et al. (1995) conducted within-herd genetic analyses of stayability, where traits considered were probabilities of a female having 2,5,8 and 11 calves, given that she calved once. The number of calves born to each dam was used to assign binary stayability observations to dams old enough to have had the required number of calves, coded as 1 (success) and 0 (failure). Observations of failure on culled cows not yet old enough to have had the required number of calves were not used. Three variations of nonlinear procedures for mixed-model analysis of binary data were used to estimate variances and predict genetic merit; animal and sire model marginal maximum likelihood, and animal model Method R, with only the former yielding heritability estimates for all traits in all herds. The heritability estimates for probability of having 2,5,8 and 11 calves, given that she calved once, were 0.09, 0.11, 0.07 and 0.20, respectively, for herd one, and 0.02, 0.14, 0.09 and

0.07, respectively, for herd two. Comparing accuracies of the 4 traits, the predictions for probability of having 5 calves, given that she calved once, had the highest mean accuracy in both herds. The authors concluded that this result, along with higher heritability estimates, offset the greater number of records available at earlier ages.

Van der Westhuizen et al. (2001) estimated variance components for stayability, longevity and calving success, and investigated the nature of the relationship between the traits using a sire model. Stayability was defined as the probability of an animal surviving to a specific age (36, 48, 60, 72 and 84 months), given the opportunity to reach that age, and coded as 1 (cow survived) and 0 (last record). Calving success was coded as 1 (successful calving) and 0 (otherwise), and longevity was calculated from the age at which the last data set was recorded. Variance components and genetic values were obtained using GFCAT, a set of programs for the analysis of "mixed" model threshold models. Heritability estimates for stayability at 36, 48, 60, 72 and 84 months of age were 0.06, 0.10, 0.06, 0.03 and 0.11, respectively. Heritability estimates for calving success and longevity were 0.03 and 0.08, respectively. Product-moment correlations between stayability at different ages were found to be low, and the authors concluded that there would be little to no improvement in level of stayability when selection was applied at another level. In general, they concluded that heritability estimates and correlations between traits were of such a low magnitude that selection for these characteristics would result in limited genetic improvement, and also indicated that sires had little influence on the stayability, longevity or calving success of their daughters. However, the authors did not address whether these results would hold for evaluation under the animal model.

Heifer pregnancy

Evans et al. (1999) evaluated the feasibility of producing expected progeny differences (EPD) for heifer pregnancy using yearling bull scrotal circumference and yearling heifer pregnancy observations. Heifer pregnancy was defined as the observation that a heifer conceives and remains pregnant to palpation, given that she was exposed at breeding, and scored as 1 (successful pregnancy) and 0 (failure to maintain pregnancy up to 120 days). Heifer pregnancy was analyzed using a maximum *a posteriori* probit threshold model to predict BV on the underlying scale, while variance components were estimated using Method R. Age of dam and age of heifer had significant effects on heifer pregnancy; heifers from 2-year-old dams were 10% less likely to conceive and remain pregnant than heifers born from mature dams, and for every 20-day increase in heifer age, there was a corresponding 10% increase in the probability a heifer will conceive and remain pregnant. The heritability estimate for heifer pregnancy was 0.138, and the estimate of the genetic correlation between heifer pregnancy and scrotal circumference was not significantly different from zero. The authors concluded that heifer pregnancy data could be used to develop BV for heifer pregnancy.

Doyle et al. (2000) investigated the nature of additive genetic relationships between heifer pregnancy, subsequent rebreeding and stayability. Heifer pregnancy was defined as described by Evans et al. (1999), and stayability as described by Snelling et al. (1995). Subsequent rebreeding was defined as the observation of a 2-year-old conceiving and remaining pregnant to palpation, given pregnancy as a yearling and exposure during the breeding season, and was coded as 1 (rebred animals) and 0 (non-pregnant females). All traits were analyzed using a maximum *a posteriori* probit threshold model to predict genetic merit on the underlying scale, while Method R was used to estimate variance components. The average heritability estimates for heifer pregnancy, subsequent rebreeding and stayability were 0.21, 0.19 and 0.15, respectively. The authors noted that, for the trait of subsequent rebreeding, only 87 of the 162 sub-samples produced point estimates within the parameter space, which they attributed in part to the small number of observations available, and the 50% repeated sub-sampling procedure of Method R. Three additive genetic groups formed on heifer pregnancy estimated BV (low, intermediate and high) were used in the analysis of stayability. The authors found differences between these groups, providing evidence for the existence of a nonlinear relationship between heifer pregnancy and stayability. The authors concluded that the difference found between the middle and high heifer pregnancy genetic groups suggested higher heifer fertility appeared favorably related to higher sustained fertility. In conclusion, the authors noted that heifer pregnancy and stayability were heritable and should respond favorably to selection, however subsequent rebreeding did not appear to be heritable. It should be noted, however, that variance components were estimated using Method R, which is not recommended for use with small data sets, as in this study. Thus, no conclusions regarding heritability of the traits can be made, and further research in this area is necessary

Other measures

Calving rate is an alternative measure of reproductive performance that has received attention by researchers. Ponzoni (1992) compared the merits of calving rate and calving day in the context of a comprehensive breeding objective. Calving day in this study was analogous to calving date, and calving rate was defined as the number of calves born per cow present in the herd. In this study, reproductive rate made the greatest contribution to genetic gain in economic units, regardless of which of the 2 traits was in the breeding objective. Genetic gain in reproductive rate and total gain in economic units were greater when calving rate was included in the breeding objective. This result was attributed to the greater phenotypic variance of calving rate under the economic and genetic assumptions made in this study. However, Ponzoni (1992) concluded that from a genetic point of view, the difference between using calving rate or calving day would be small, compared with the effect of completely ignoring reproduction.

While from a genetic point of view, calving rate, as defined by Ponzoni (1992), may be superior to calving date, from a production perspective, calving rate and calving success have some of the same deficiencies as calving interval. Both measures are historic, and do not indicate when cows calve in the calving season. Calving rate as defined by Meyer et al. (1990), can only be used after a number of calvings have taken place, and, thus, can't be used directly on heifers as a measure of future production.

Another potential trait for selection is pregnancy rate, as it has been shown that pregnancy rate measured in the first parity is the same trait as lifetime pregnancy rate. Morris and Cullen (1994) estimated genetic correlations between pubertal traits of males or females and lifetime pregnancy rate. Yearling pregnancy rate was considered normal, and coded as 1 (success) and 0 (failure). Lifetime pregnancy rate was calculated as the number of pregnancies divided by number of mating years, up to the fifth mating year. Heritability estimates for yearling and lifetime pregnancy rate and calving date were 0.04, 0.04 and 0.04, respectively. The phenotypic and genetic correlations between yearling and lifetime pregnancy rate were 0.84 and 0.92, respectively, indicating that they are the same trait. Genetic correlations of standardized age at first estrus with yearling or lifetime pregnancy rate were all negative and, hence, desirable in direction. For scrotal circumference, genetic correlation estimates with yearling and lifetime pregnancy rate were 0.53 and 0.34, respectively. From this study it appears that pubertal traits are favorably correlated with lifetime pregnancy rate.

Morris et al. (2000) estimated genetic parameters for age at first estrus, calving date and pregnancy rates using experimental data. Heritabilities for standardized age at first estrus and calving date were 0.27 and 0.09, respectively. Genetic correlations of standardized age at first estrus with calving date and pregnancy rate were 0.57 and –0.36, respectively. The pregnancy rate for the line selected for reduced age at puberty was 5% higher than the line selected for increased age at puberty, and the mean calving date was 3 days earlier. Thus, the authors concluded that selecting for reduced age at puberty leads to earlier calving dates and higher pregnancy rates in beef females.

Researchers have also attempted to identify physiological parameters, such as endocrine factors, that are related to fertility, and are heritable. Mialon et al. (2000) found a favorable genetic correlation between age at puberty and postpartum intervals in experimental data. The length of postpartum anoestrus was estimated based on weekly blood progesterone assays and on twice daily detection of estrus behavior. Estimates of heritability and repeatability for the interval from calving to first observed estrus were 0.12 and 0.38, respectively. Corresponding values for the interval from calving to the first positive progesterone test were 0.35 and 0.60, respectively. The genetic and phenotypic correlations between the two measures of postpartum interval were 0.98 and 0.65, respectively. The genetic relationships between postpartum intervals and body weight and body condition score at time of calving were negative; cows that were genetically heavier at calving with more body reserves had shorter postpartum intervals. A favorable positive genetic correlation between age at puberty and postpartum intervals was found, in that heifers which were younger at puberty also had shorter postpartum intervals. While the favorable relationships of the postpartum intervals with weight at calving and age at puberty may benefit beef producers, it is unlikely that direct selection on either trait will be possible, due to the difficulty in measuring both traits outside of experimental populations.

Age at first calving has also been studied as a potential measure of reproductive performance. A reduced age at first calving would increase the number of calves born in the herd. An advantage of this measure is that it can be computed without the need for additional data, as the birth date of the cow and her first calving are generally known. The biggest disadvantages are that it only represents one component in the reproductive life of a cow, and that it is only recorded in heifers. Furthermore, in a variable seasonal environment, age at first calving reflects management decisions to a greater extent than genetic merit. Bourdon and Brinks (1982) reported a low heritability estimate (0.07) for age at first calving, and favorable correlations with growth traits.

Conclusions

Reproduction is a complex trait, and, hence, there are many different measures of reproductive performance. Some of the more popular alternative measures include age at first estrus, age at first breeding, calving rate and pregnancy rate. While many researchers have identified these measures to be heritable, they are not widely used for several reasons. Some measures are historic, and fail to provide an indication of when cows calve in the calving season (calving and pregnancy rates), while others are heavily influenced by management (age at first calving). Other measurements cannot be measured feasibly in field data (postpartum intervals).

The traits of calving date and days to calving have been identified as suitable measures of reproductive performance. They are heritable traits, and allow producers to distinguish between early and late calvers in their herds. However, past studies have generally used records from the first and second parities. Thus further research to ascertain whether these results can be extrapolated for the entire reproductive life of the female is needed. As well, the method of prediction of records for open cows needs further refinement. The alternative approach of using survival analysis should be investigated in the future.

The trait of heifer pregnancy is currently used in genetic evaluation. However, it fails to identify when an individual will calve in the calving season, thus should be included along with some measure of calving date. As well, the relationship between heifer pregnancy and lifetime productivity, or stayability, has not been clearly defined. Further research to properly quantify this relationship is needed.

While several measures of reproductive performance are currently being incorporated into national genetic evaluation, further refinement is still needed. Given the nature of reproductive records, it is unlikely that one individual measure will be able to completely predict reproductive performance. Most likely several measures will need to be used together. The main limitations to genetic evaluation of fertility in the past, and currently, are the lack of records available from field data. The adoption of whole-herd reporting schemes by herds will help to alleviate this problem. In conclusion, there is much potential to make improvements to the evaluation of female reproductive performance of beef cattle in the future.

Literature Cited

Bourdon, R. M. and J. S. Brinks. 1983. Calving date versus calving interval as a reproductive measure in beef cattle. J. Anim. Sci. 57:1412-1417.

Bourdon, R. M. and J. S. Brinks. 1982. Genetic, environmental and phenotypic relationships among gestation length, birth weight, growth traits and age at first calving in beef cattle. J. Anim. Sci. 57:1412-1417.

Buddenberg, B. J., C. J. Brown and A. H. Brown. 1990. Heritability estimates of calving date in Hereford cattle maintained on range under natural mating. J. Anim. Sci. 68:70-74.

Doyle, S. P., B. L. Golden, R. D. Green and J. S. Brinks. 2000. Additive genetic parameter estimates for heifer pregnancy and subsequent reproduction in Angus females. J. Anim. Sci. 78:2091-2098.

Ducrocq, V. 1994. Statistical analysis of length of productive life for dairy cows of the Normande breed. J. Dairy Sci. 77:855-866.

Eicker, S. W., Y. T. Grohn and J. A. Hertl. 1996. The association between cumulative milk yield, days open, and day to first breeding in New York Holstein cows. J. Dairy Sci. 79:235-241.

Evans, J. L., B. L. Golden, R. M. Bourdon and K. L. Long. 1999. Additive genetic relationships between heifer pregnancy and scrotal circumference in Hereford cattle. J. Anim. Sci. 77:2621-2628.

Johnston, D. J. and K. L. Bunter. 1996. Days to calving in Angus cattle: Genetic and environmental effects, and covariances with other traits. Livest. Prod. Sci. 45:13-22.

Johnston, D. J., J. M. Henshall and B. Tier. 2001. Estimate of the genetic correlation between calving success and days to calving in Angus females. Proc. Assoc. Advmt. Anim. Breed. Genet., Queenstown, New Zealand. pp 353-355.

Mialon, M.-M., G. Renand, D. Krauss and F. Ménissier. 2000. Genetic variability of the length of postpartum anoestrus in Charolais cows and its relationship with age at puberty. Genet. Sel. Evol. 32:403-414.

Meacham, N. S. and D. R. Notter. 1987. Heritability estimates for calving date in Simmental cattle. J. Anim. Sci. 64:701-705.

Melton, B. E. 1995. Conception to consumption: The economics of genetic improvement. In: Proc. Beef Improvement Federation 27th Research Symposium and Annual Meeting, Sheridan, Wyoming. pp 40-47.

Meyer, K., K. Hammond, P. F. Parnell, M. J. Mackinnon and S. Sivarajasingam. 1990. Estimates of heritability and repeatability for reproductive traits in Australian beef cattle. Livest. Prod. Sci. 25:15-30.

Meyer, K., K. Hammond, M. J. Mackinnon and P. F. Parnell. 1991. Estimates of covariances between reproduction and growth in Australian beef cattle. J. Anim. Sci. 69:3533-3543.

Morris, C. A., J. A. Wilson, G. L. Bennett, N. G. Cullen, S. M. Hickey and J. C. Hunter. 2000. Genetic parameters for growth, puberty, and beef cow reproductive traits in a puberty selection experiment. New Zealand J. Agric. Res. 43:83-91.

Morris, C. A. and N. G. Cullen. 1994. A note on genetic correlations between pubertal traits of males or females and lifetime pregnancy rate in beef cattle. Livest. Prod. Sci. 39:291-297.

Notter, D. R. 1988. Evaluating and reporting reproductive traits. In: Proc. Beef Improvement Federation 20th Research Symposium and Annual Meeting, Albuquerque, New Mexico. pp 21- 42.

Notter, D. R. and M. H. Johnson. 1988. Simulation of genetic control of reproduction in beef cows. IV. Within-herd breeding value estimation with pasture mating. J. Anim. Sci. 66:280-286.

Ponzoni, R. W. 1992. Which trait for genetic improvement of beef cattle reproduction-calving rate or calving day? J. Anim. Breed. Genet. 109:119-128.

Rege, J. E. O. and T. R. Famula. 1993. Factors affecting calving date and its relationship with production traits of Hereford dams. Anim. Prod. 57:385-395.

Snelling, W. M., B. L. Golden and R. M. Bourdon. 1995. Within-herd genetic analyses of stayability of beef females. J. Anim. Sci. 73:993-1001.

Van der Westhuizen, R. R., S. J. Schoeman, G. F. Jordaan and J. B. van Wyk. 2001. Heritability estimates derived from threshold analyses for reproduction and stayability traits in a beef cattle herd. S. Afr. J. Anim. Sci. 31:25-32.

CHAPTER 3

COMPARISON OF METHODS FOR HANDLING CENSORED RECORDS IN BEEF FERTILITY DATA: SIMULATION STUDY [1]

---

[1] Donoghue, K. A., R. Rekaya and J. K. Bertrand. Submitted to *Journal of Animal Science*, 5/2/2003.

ABSTRACT

A simulation study was conducted to compare methods for handling censored days to calving records in beef cattle data. Days to calving was defined as the days between the time a bull is turned out in the pasture and the subsequent parturition. Simulated data were generated to have similar data structure and genetic relationships to an available field data set. Records were simulated for 33,176 daughters of 4238 sires. Data were simulated using a mixed linear model that included the fixed effects of contemporary group and sex of calf, linear and quadratic covariates for age at mating, and random effects of animal and residual error. Two methods for handling censored records were evaluated, and two censoring rates of 12 and 20% were applied to assess the influence of higher censoring rates on inferences. Censored records were assigned penalty values on a within contemporary group basis under the first method (DCPEN). Under the second method (DCSIM), censored records were drawn from their respective predictive distributions. Bayesian approach via Gibbs sampling was used to estimate variance components and predict breeding values. Posterior means (PM) and standard deviations (SD) of additive genetic variance for DCPEN at 12 and 20% censoring were 23.23 (3.72) and 21.04 (3.64), respectively, while the same estimates for DCSIM at 12 and 20% censoring were 23.65 (3.30) and 21.93 (3.42), respectively. In all cases, the true value of the genetic variance was within the high posterior density interval HPD (95%). PM (SD) of residual variance for DCPEN at 12 and 20% censoring were 415.7 (4.67) and 440.0 (4.79), respectively, while the same estimates for DCSIM at 12 and 20% censoring were 371.0 (4.31) and 365.4 (4.43), respectively. The true value of the residual variance was within the HPD (95%) for DCSIM, but it was outside this interval for DCPEN at both censoring rates, indicating a systematic bias for this parameter. Bayes factor and Deviance Information Criteria were used for model comparisons, and both criteria indicated the superiority of the DCSIM method. However, little difference was observed between the two methods for correlations between true breeding values and posterior means of animal effects for sires, indicating that no major re-ranking of sires would be expected. This suggests that either censored data handling technique can be successfully used in a genetic evaluation for days to calving.

Key words: Beef cattle, Fertility, Censored records

Introduction

Fertility or reproductive performance has been reported to be at least twice as important, economically, as production traits under a conventional cow-calf operation (Melton, 1995). Johnston and Bunter (1996) demonstrated that days to calving, defined as the days between the time a bull is turned out in the pasture and the subsequent parturition, was a suitable measure of reproductive performance in a large field data set. In the same study, cows with censored records were assigned a projected value on a within contemporary group basis; therefore, all cows within a contemporary group that failed to calve received the same trait value.

An alternative approach to adding a fixed number of days to censored records would be to assume a truncated normal distribution for the uncensored records, and randomly draw from this truncated distribution to obtain a record for censored females on a contemporary group basis. This would allow the data to determine the value of censored records for each censored female.

The objective of this study was to compare methods for handling censored fertility records in beef cattle data. Two methods for handling censored records were evaluated and compared; assigning a penalty value on a within contemporary group basis and simulating records for these animals from their respective predictive distributions. Several data sets with different levels of censoring were simulated and analyzed using both methods. Bayes factor and Deviance Information Criteria were used to assess the plausibility of the two approaches.

Materials and Methods

Data Simulation

Given that the objective of the project was to assess the suitability of two different models, data were simulated using known parameters obtained from a field data set (Donoghue et al., 2003). The simulated data set had a similar structure to the field data set, and used the same pedigree file. Single days to calving (DC) records were simulated for 33,176 daughters of 4238 sires. Data were simulated using a mixed linear model, and included the fixed effects of contemporary group and sex of calf, linear and quadratic covariates for age at mating, and random effects of animal and residual error. Contemporary group was defined to include animals from the same herd who were mated in the same month and year to the same sire, and these effects were drawn from the uniform distribution U[210,510]. The effect of sex of calf was drawn from a normal distribution N(50,400). The random effect of animal was sampled from a normal distribution with mean zero and variance $\sigma^2_u=25$. The residual terms were sampled from a normal distribution with mean zero and variance $\sigma^2_e=370$. To assess the influence of higher censoring rates on inferences, two censoring rates of 12 and 20% were applied.

Animals were selected to be censored on a contemporary group basis; all animals within a CG were ranked by their DC records, and animals with the highest DC records were chosen to be censored. The number of cows within a particular CG that were censored was random. A summary of the real and simulated data sets can be found in Table 3.1.

Data Analyses

Penalty method (DCPEN). Penalty values were assigned to each censored record on a within contemporary group basis. As suggested by Johnston and Bunter (1996), the highest DC record within each contemporary group was identified, and a constant number of days (21) were added to this record to generate the projected value for all censored records within that group.

Simulation method (DCSIM). Trait values for censored records were simulated from their respective predictive distributions (truncated normal distributions). For all animals in the same contemporary group, the truncation point was the largest observed DC record. The predicted DC for a censored record was between the truncation point and positive infinity. Thus, an animal with a censored record could not receive a simulated record that was smaller than a non-censored record within her contemporary group. The number of days added to this truncation point for each of the censored records was determined by drawing samples at random from the truncated distribution, and depended upon the fixed effects in the model, as well as her relationships with other animals.

Data analysis. A single trait mixed linear model was used for analysis of DC. In matrix notation, the following model was adopted:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e}$$

where **y** is a vector of 33176 observations; **b** is the vector of fixed effects, **u** is the vector of additive genetic values of all animals and **e** is the vector of residual terms. **X** and **Z** are known incidence matrices. The vector **b** included 3568 contemporary group effects, two sex of calf effects, and linear and quadratic covariates for age at mating.

The vector **y** includes uncensored data points $\mathbf{y_0}$ (m x 1) and censored records $\mathbf{y_c}$ (n – m x 1), where n is the total number of observations, such that $\mathbf{y'} = [\mathbf{y'_o}, \mathbf{y'_c}]$. If we let $\mu_i = \mathbf{x'_i b} + \mathbf{z'_i u}$, (i = 1,…n), then the density of the conditional distribution of the uncensored and censored records, given the parameters, can be written as:

$$f(\mathbf{y} \mid \mathbf{b}, \mathbf{u}, \sigma_e^2) = (2\pi\sigma_e^2)^{-m/2} \exp\left[\frac{-1}{2\sigma_e^2}\sum_{i=1}^{m}(y_{oi}-\mu_i)^2\right] \times \prod_{i=m+1}^{n}\left[1-\Phi\left(\frac{c_i-\mu_i}{\sigma_e}\right)\right]$$

where y_{ci} is the unobserved DC, c_i is the value of the trait at censoring time, and Φ(.) is the standard normal cumulative distribution function.

Augmenting the posterior distribution with the unobserved calving dates corresponding to the censored observations (Tanner, 1996) simplifies the procedure. Let $\mathbf{w} = \{w_j\}$ with $w_j > c_j$, j = m+1, m+2, …n and let the augmented data vector be $\mathbf{y'_A} = [\mathbf{y'_o}, \mathbf{w'}]$. The joint density of calving dates, both observed and unobserved, given the parameters, is:

$$f(\mathbf{y}_A \mid \mathbf{b}, \mathbf{u}, \sigma_e^2) \propto (\sigma_e^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma_e^2}\left[\sum_{i=1}^{m}(y_{oi}-\mu_i)^2 + \sum_{i=m+1}^{n}(w_i-\mu_i)^2\right]\right\}$$

where $w_i \geq c_i$ for all appropriate i values. Taking into account the augmented data **w,** the new vector of unknown parameters of the model will be $\boldsymbol{\theta}'_A = (\mathbf{b'}, \mathbf{u'}, \mathbf{w'}, \sigma^2_u, \sigma^2_e)$. To complete the Bayesian formulation, the following prior distributions were assumed to the unknown parameters:

$p(\mathbf{b}) \sim N(0,10^6)$

$p(\mathbf{u}|\sigma^2_u) \sim N(0, A\,\sigma^2_u)$

where **A** is the matrix of additive relationships between animals and σ^2_u is the additive variance

$$p(\sigma_e^2) \sim U[0,10000]$$

$$p(\sigma_u^2) \sim U[0,1000]$$

where U[.] is a uniform distribution.

The augmented joint posterior density can be expressed as:

$$\mathrm{p}(\mathbf{b}, \mathbf{u}, \mathbf{w}, \sigma_u^2, \sigma_e^2 \mid \mathbf{y_o}) \propto \mathrm{p}(\mathbf{y_A} \mid \mathbf{b}, \mathbf{u}, \sigma_e^2)\mathrm{p}(\mathbf{u} \mid \sigma_u^2)$$

$$\propto (\sigma_e^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma_e^2}\left[\sum_{i=1}^{m}(\mathrm{y}_{oi} - \mu_i)^2 + \sum_{i=m+1}^{n}(\mathrm{w}_i - \mu_i)^2\right]\right\}$$

$$(\sigma_u^2)^{-q/2} \exp\left[-\frac{1}{2\sigma_u^2}\mathbf{u'A^{-1}u}\right]\prod_{i=m+1}^{n}\mathrm{I}(\mathrm{w}_i \geq \mathrm{c}_i)$$

where I(.) is an indicator variable denoting censoring of record i (i = m+1, m+2, …n).

Following the notation of Guo et al. (2001), let $\theta_A = (\theta'_L, \mathbf{w'}, \sigma^2_u, \sigma^2_e)'$, where $\theta'_L = (\mathbf{b'}, \mathbf{u'})'$, the full conditional distributions required for the implementation of the Gibbs sampling are easily obtained:

$\mathrm{w}_j|\theta_L, \mathbf{w_{\_j}}, \sigma^2_u, \sigma^2_e, \mathbf{y_o} \sim N(\mu_j, \sigma^2_e)\ I\ (\mathrm{w}_j > c_j);$ j = m+1, m+2, ….n

where $\mathbf{w_{\_j}}$ is **w** without w_j

$$\theta_L \mid \mathbf{w}, \sigma_u^2, \sigma_e^2, \mathbf{y_o} \sim \mathrm{N}(\hat{\mathbf{s}}, \mathbf{C}^{-1}\sigma_e^2)$$

$$\sigma_u^2 \mid \theta_L, \sigma_e^2, \mathbf{w}, \mathbf{y_o} \sim (\mathbf{u'A^{-1}u})\chi^{-2}_{q-2}$$

$$\sigma_e^2 \mid \theta_{L,}\sigma_u^2, \mathbf{w}, \mathbf{y_o} \sim (\mathbf{y_A} - \mathbf{Xb} - \mathbf{Zu})'(\mathbf{y_A} - \mathbf{Xb} - \mathbf{Zu})\chi^{-2}_{n-2}$$

where $\mathbf{C^{-1}}$ is the inverse of the coefficient matrix and χ_v^{-2} is the inverted chi distribution with v degree of freedom.

For each level of censoring (12 and 20%), variance components were estimated and breeding values were predicted using the two methods for handling censored records. As well, variance components and breeding values were obtained for the complete non-censored data set. For each analysis, five replicates were simulated. Results from the average of these analyses were compared to the true variances, fixed effect and breeding value solutions for the simulated data.

Model Comparison

Bayes Factor. The Bayes factor, defined by Newton and Raftery (1994), was used to assess the feasibility of the two models. The harmonic means of likelihood values evaluated at the posterior draws were estimated from the marginal density of the data under each of the models:

$$\hat{p}(\mathbf{y} \mid M_i) = \left\{ \frac{1}{N} \sum_{j=1}^{N} \left[p(\mathbf{y} \mid \theta^{(j)}, M_i) \right]^{-1} \right\}^{-1}$$

where $\mathbf{y}$ is the vector of observed responses and $\theta^{(j)}$ is the Gibbs sampling sample j of parameters under model M_i. The estimated Bayes factor between models M_i and M_j is:

$$BF_{Mi,Mj} = \frac{\hat{p}(y \mid M_i)}{\hat{p}(y \mid M_j)}$$

Deviance Information Criterion. The Deviance Information Criterion (DIC), as defined by Spiegelhalter et al. (2002), was used to compare models:

$$DIC = \bar{D}(\theta) + p_D$$

where

$\bar{D}(\theta) = E_{\theta|y}[D(\theta)]$ is posterior expectation of the Bayesian deviance $D(\theta) = -2\log p(y \mid \theta)$

and $p_D = \overline{D(\theta)} - D(\bar{\theta})$ is the effective number of parameters.

It can be seen that DIC consists of a measure of goodness of fit $D(\bar{\theta})$, with a penalty (p_D) for increasing model complexity. The DIC is easy to estimate within a MCMC implementation, as the Bayesian deviance, $D(\theta)$, is computed at each iteration. At the end of iterations, the mean of the Bayesian deviance, $\overline{D(\theta)}$, and the mean value of the model parameters, $\bar{\theta}$, are calculated, leading directly to the DIC.

Results and Discussion

For all analyses, convergence was assessed using methodology presented by Raftery and Lewis (1992). The required length of the burn-in period was always less then 2,500 iterations for all parameters. Thus, 75,000 iterations of the sampler were run with a conservative 20,000 iterations discarded as burn-in; all remaining 55,000 iterations were retained without thinning for post-Gibbs analysis.

Summaries of the posterior distributions of genetic parameters for DC under different censoring scenarios are presented in Table 3.2. Posterior means of the additive variance under both methods (DCPEN and DCSIM) for both levels of censoring (12 and 20%) were similar to the true value (σ^2_u=25). However, there was a tendency of under estimation of true genetic variance, especially for the 20% censoring case. As censoring rate increased, the posterior mean of the additive variance decreased slightly; 23.23 and 23.65 vs. 21.04 and 21.93 for DCPEN and DCSIM at 12 and 20% censoring, respectively. Guo et al. (2001) assessed the influence of higher censoring rates on parameters for performance and prolificacy traits in swine using a similar approach to the present study. The authors observed decreasing sire variances for higher levels of censoring, similar to the trend in additive variances at higher levels of censoring in this study. In all cases, the true value of the genetic variance was well within the high posterior density interval HPD (95%). These results indicate that the method of handling censored fertility records did not have a significant impact on estimation of additive variance.

Posterior means of the residual variance based on five replicates under DCSIM were 370.97 and 365.40 for 12% and 20% censoring rate, respectively. These point estimates were very similar to the true value of the parameter (σ^2_e=370). Estimates of this parameter under the DCPEN method, however, were significantly higher than the true value of the parameter and beyond the expected Monte Carlo error. In fact, the true value of the residual variance was outside the HPD (95%) interval for both censoring rates, indicating a systematic bias inferring this parameter. These results imply that the DCSIM method provides a better fit to the data when censored records are present. As censoring rate increased, the posterior mean of the residual variance increased markedly under DCPEN (415.69 vs. 439.96 for 12 and 20% censoring, respectively). Guo et al. (2001) observed a similar trend in residual variances under higher levels of censoring in their study. The higher estimates of residual variance using DCPEN indicates that adding a constant number of days to all censored records within a contemporary group may be significantly overestimating the censored DC records, especially at higher levels of censoring. The slightly smaller estimate of residual variance at 20% censoring under DCSIM may imply that this method is slightly underestimating the records at higher levels of censoring. However, those differences are within Monte Carlo error, implying the absence of bias.

The posterior means of heritability were similar to the true value of the parameter (h^2=0.06) under DCSIM for both levels of censoring. As a result of the over-estimation of the residual variance under DCPEN, the posterior means of heritability for both censoring levels were slightly smaller than the true value of the parameter. However, in both cases, the true value of the parameter was well within the HPD (95%).

Pearson correlations between true breeding values and posterior means of animal effects for sires are given in Table 3.3. As expected, these correlations increased with the number of progeny. Correlations based on the complete data (non censoring) are slightly higher than those for DCPEN and DCSIM for both levels of censoring. Correlations for DCPEN and DCSIM across both levels of censoring were very similar, implying that the level of censoring appeared to have little effect on ranking of sires. There was little difference observed in correlations between DCPEN and DCSIM as the number of progeny with records increased. These results indicate that DCSIM may be slightly more accurate than DCPEN in predicting the true genetic values of these sires; however, no major re-ranking of sires will be expected between the two methods of handling censored records.

The percentage of sires in the top and bottom 10% ranked on the basis of true and predicted posterior means of animal effects for sires with daughters with records are given in Table 3.4. Due to the fact that lower values of DC are more desirable, sires with lower posterior means of animal effects are ranked higher then sires with higher posterior means of animal effects. Concordance with the true evaluation for the top 10% of sires was similar across methods of handling censored records, and across levels of censoring. However, the range in concordance with true evaluation for the bottom 10% of sires was slightly larger, especially at higher levels of censoring; 25% to 29% for DCPEN and DCSIM at 20% censoring, respectively. These results show that although the predicted ranking for the majority of sires in the top 10% was similar to the true ranking under both methods, there were differences observed between the true and predicted rankings for the bottom 10% of sires for both DCPEN and DCSIM. The lower values of concordance between rankings under DCPEN and true rankings at both levels of censoring suggest that the DCSIM method is doing a better job of identifying sires of inferior genetic merit.

Table 3.5 presents the BF values, as well as the DIC values and its components for the two methods. The DCSIM method had a better fit to the data, as illustrated by the lower value of $\bar{D}(\theta)$, but received a larger penalty value (p_D) for having a more complex model. Despite this higher penalty value, however, DCSIM had a lower DIC value, implying that it is the superior model when both goodness of fit and model complexity are considered. The superiority of the DCSIM method was further confirmed by the estimated BF between models (3563.5), showing that DCSIM was a significantly better model than DCPEN.

Implications

There were little differences observed in the estimation of additive variance of the trait or ranking of sire breeding values between the two methods of handling censored data. However, the overestimation of the residual variance and, consequently, the under estimation of heritability under the penalty method, in conjunction with the model comparison criteria, indicate that the simulation approach provides a better method for handling censored records in beef fertility data, especially at higher levels of censoring. The penalty method does appear to be significantly overestimating the censored records;

however, the lack of significant differences in the genetic ranking of sires between the two methods suggests that either censored data handling technique can be successfully used in a genetic evaluation for days to calving. Further research employing both methods to predict days to calving in beef cattle field data should be undertaken to verify the results of this simulation study.

Acknowledgements

Appreciation is extended to the Angus Society of Australia for providing the data; Meat and Livestock Australia for the research scholarship provided to the first author, and Dr. David Johnston and Carel Teseling for their contributions.

Literature Cited

Donoghue, K. A., R. Rekaya, J. K. Bertrand, D. J. Johnston and C. Teseling. 2003. Comparison of methods for handling missing fertility records in beef cattle data. J. Anim. Sci. 81(Suppl. 1): (Abstr.) (Accepted).

Guo, S-F., D. Gianola, R. Rekaya and T. Short. 2001. Bayesian analysis of lifetime performance and prolificacy in Landrace sows using a linear mixed model with censoring. Livest. Prod. Sci. 72: 243-252.

Johnston, D. J. and K. L. Bunter. 1996. Days to calving in Angus cattle: Genetic and environmental effects, and covariances with other traits. Livest. Prod. Sci. 45:13-22.

Melton, B. E. 1995. Conception to consumption: The economics of genetic improvement. In: Proc. Beef Improvement Federation 27th Research Symposium and Annual Meeting, Sheridan, Wyoming. pp 40-47.

Newton, M. A. and Raftery, A. E. 1994. Approximate Bayesian inference with the weighted likelihood bootstrap. J. R. Statist. Soc. 56:3-48.

Raftery, A. E. and Lewis, S. 1992. How many iterations in the Gibbs sampler? Bayesian Statistics 4, pp. 763-773 (ed. J. M. Bernando, J. O. Berger, A. P. Dawid and A. F. M. Smith, Oxford University Press).

Spiegelhalter, D. J., N. G. Best, B. P. Carlin and A. van der Linde. 2002. Bayesian measures of model complexity and fit. J. Roy. Statist. Soc. 64:583-639.

Tanner, M. A. 1996. Tools for Statistical Inference. Springer-Verlag.

Table 3.1. Summary of real and simulated data

Data	Records	Animals	Sires	Herds
Real	33176	62857	4238	470
Simulated				
- No censoring	33176	62857	4238	470
- 12% censoring	29067	62857	4238	470
- 20% censoring	26480	62857	4238	470

Table 3.2. Posterior means, standard deviations (SD) and bounds of highest probability density intervals (95%) for simulated data at different levels of censoring

Rate	Parameter	Mean[a]	SD[a]	HPD (95%)[a]	
No censoring	σ^2_u	23.70	3.57	16.91	30.58
	σ^2_e	371.23	4.30	362.70	379.44
	h^2	0.06	0.009	0.04	0.08
12% censoring					
- DCPEN	σ^2_u	23.23	3.72	16.22	30.31
	σ^2_e	415.69	4.67	406.34	424.52
	h^2	0.05	0.009	0.04	0.07
- DCSIM	σ^2_u	23.65	3.30	17.15	29.81
	σ^2_e	370.97	4.31	362.34	379.18
	h^2	0.06	0.008	0.04	0.07
20% censoring					
- DCPEN	σ^2_u	21.04	3.64	14.31	27.87
	σ^2_e	439.96	4.79	430.36	449.02
	h^2	0.05	0.008	0.04	0.06
- DCSIM	σ^2_u	21.93	3.42	15.57	28.92
	σ^2_e	365.40	4.43	356.60	373.84
	h^2	0.06	0.009	0.04	0.08

[a] Averaged over five replicates

Table 3.3. Pearson correlations between true breeding values and posterior means of animal effects for sires with progeny with records

Rate	All sires[a]	> 10 progeny[b]	> 20 progeny[c]
No censoring	0.45	0.56	0.62
12% censoring			
- DCPEN	0.42	0.53	0.58
- DCSIM	0.43	0.54	0.59
20% censoring			
- DCPEN	0.41	0.52	0.58
- DCSIM	0.42	0.54	0.59

[a] All sires with progeny with records (n=4238)

[b] Sires with more than 10 progeny with records (n=773)

[c] Sires with more than 20 progeny with records (n=322)

Table 3.4. Percentage of sires in top and bottom 10% for true and predicted rankings using posterior means of animal effects for sires with progeny with records

Rate	Top 10%[a]	Bottom 10%[b]
No censoring	31%	29%
12% censoring		
- DCPEN	30%	28%
- DCSIM	30%	29%
20% censoring		
- DCPEN	30%	25%
- DCSIM	31%	29%

[a] Percentage of sires in common with true rankings for the top 10% top sires (n=424)

[b] Percentage of sires in common with true rankings for the bottom 10% top sires (n=424)

Table 3.5. Expectation of the Bayesian deviance, $D(\theta)$, Bayesian deviance evaluated at the posterior mean of model parameters, $D(\bar{\theta})$, effective number of parameters (p_D), deviance information criterion (DIC) and Bayes Factor (BF) for DCSIM and DCPEN using 20% censoring

Criterion	DCSIM	DCPEN
$\bar{D}(\theta)$	293035.7	300873.2
$D(\bar{\theta})$	287502.0	296724.6
p_D	5533.7	4148.6
DIC	298569.4	305021.8
$BF_{DCSIM,DCPEN}$	3563.45	

CHAPTER 4

COMPARISON OF METHODS FOR HANDLING CENSORED RECORDS IN BEEF FERTILITY DATA: FIELD DATA[1]

---

[1] Donoghue, K. A., R. Rekaya and J. K. Bertrand. Submitted to *Journal of Animal Science*, 5/2/2003.

ABSTRACT

The purpose of this study was to compare methods for handling censored days to calving records in beef cattle data, and verify results of an earlier simulation study. Data were records from natural service matings of 33,176 first-calf females in Australian Angus herds. Three methods for handling censored records were evaluated. Censored records were assigned penalty values on a within contemporary group basis under the first method (DCPEN). Under the second method (DCSIM), censored records were drawn from their respective predictive truncated normal distributions, while censored records were deleted under the third method (DCMISS). Data were analyzed using a mixed linear model that included the fixed effects of contemporary group and sex of calf, linear and quadratic covariates for age at mating, and random effects of animal and residual error. Bayesian approach via Gibbs sampling was used to estimate variance components and predict breeding values. Posterior means (PM) (SD) of additive genetic variance for DCPEN, DCSIM and DCMISS were 22.56 (4.24), 26.10 (3.61), and 13.51 (2.89), respectively. PM (SD) of residual variance for DCPEN, DCSIM and DCMISS were 431.4 (5.04), 371.4 (4.49), and 262.2 (3.41), respectively. PM (SD) of heritability for DCPEN, DCSIM and DCMISS were 0.05 (0.01), 0.07 (0.01), and 0.05 (0.01), respectively. Simulating trait records for non-calving females resulted in similar heritability to the penalty method, but lower residual variance. Pearson correlations between posterior means of animal effects for sires with more than 20 daughters with records between DCPEN-DCSIM, DCPEN-DCMISS and DCSIM-DCMISS were 0.99, 0.77 and 0.81, respectively. Of the 424 sires ranked in the top 10% and bottom 10% of sires in DCPEN, 91% and 89%, respectively, were also ranked in the top 10% and bottom

10% in DCSIM. Little difference was observed between DCPEN and DCSIM for correlations between posterior means of animal effects for sires, indicating that no major re-ranking of sires would be expected. This suggests that either censored data handling technique can be successfully used in a genetic evaluation for days to calving.

Key words: Beef cattle, Fertility, Censored records

Introduction

The economic value of fertility to the cow-calf operation has been reported to be at least twice as important as production traits (Melton, 1995). However, relatively little attention has been given to these traits in genetic evaluations as they are difficult to measure, report, evaluate and interpret. In order to make the best use of the data available for reproductive performance, information for open cows must be included in the evaluation. Notter (1988) suggested using threshold theory to assign a predicted value to censored females, and this approach was used in several studies (Buddenberg et al., 1990 and Meyer et al., 1991). As noted by Meyer et al. (1991), using threshold theory to generate missing records provided a simple, ad hoc approach to account for the censored nature of the trait of days to calving, where days to calving was defined as the days between the time a bull was turned out in the pasture and the subsequent parturition. Other more sophisticated procedures did exist at the time of these studies (for example Carriquiry et al., 1987), but were more complex to implement.

Using field data, Johnston and Bunter (1996) generated censored days to calving records by adding a 21-day penalty to the largest days to calving record within a contemporary group. Donoghue et al. (2003) suggested an alternative approach to handling censored records; a truncated normal distribution is assumed for the uncensored records, and random draws from this truncated distribution generate records for censored females on a contemporary group basis. Simulated data were used to compare this approach with the approach used by Johnston and Bunter (1996). Model comparison criteria indicated the superiority of the generating censored records using the truncated normal distribution compared to the penalty method, however, no major re-ranking of sires was observed.

The objective of this study was to verify the results of the simulation study using a field data set. Three methods for handling censored records were evaluated and compared; assigning a penalty value on a within contemporary group basis; simulating records for these animals from their respective predictive distributions and deleting censored records from the data set. Correlations between breeding values for sires under all methods were used to investigate reasons for rank changes.

Materials and Methods

Data Characteristics

The data set consisted of days to calving (DC) records from natural service matings of first-calf females in Australian Angus herds. The trait was defined as the number of days between the time a bull is turned out in the pasture and the subsequent parturition of the female. After editing, records from 33,176 females born between 1987 and 2000 were available for analysis. Edits performed included removal of 1) animals with incomplete records; 2) mating records resulting in multiple births; 3) single record contemporary groups and 4) contemporary groups consisting of only non-calvers. Only animals having their first mating record between 270 and 625 days of age were included in the analysis. Contemporary group was defined to include animals from the same herd who were mated in the same month and year to the same sire. The sex of calf effect was arbitrarily set to male for all animals with censored records. A summary of the data set can be found in Table 4.1.

Data Analyses

Penalty method (DCPEN). Penalty values were assigned to each censored record on a within contemporary group basis. As suggested by Johnston and Bunter (1996), the highest DC record within each contemporary group was identified, and a constant number of days (21) were added to this record to generate the projected value for all censored records within that group.

Simulation method (DCSIM). Using the approach of Donoghue et al. (2003), trait values for censored records were simulated from their respective predictive distributions (truncated normal distributions). For all animals in the same contemporary group, the truncation point was the largest observed DC record. The predicted DC for a censored record was between the truncation point and positive infinity. Thus, an animal with a censored record could not receive a simulated record that was smaller than a non-censored record within her contemporary group. The number of days added to this truncation point for each of the censored records was determined by drawing samples at random from the truncated distribution, and depended upon the fixed effects in the model, as well as her relationships with other animals.

Missing method (DCMISS). In order to evaluate the effect of ignoring censored records, a third method was included. Under DCMISS, records for non-calvers were deleted from data set.

Data analysis. A single trait mixed linear model was used for analysis of DC. In matrix notation, the following model was adopted:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e}$$

where **y** is a vector of observations; **b** is the vector of fixed effects, **u** is the vector of additive genetic values of all animals and **e** is the vector of residual terms. **X** and **Z** are known incidence matrices. The vector **b** included contemporary group effects, sex of calf effects, and linear and quadratic covariates for age at mating.

The vector **y** includes uncensored data points $\mathbf{y_0}$ (m x 1) and censored records $\mathbf{y_c}$ (n – m x 1), where n is the total number of observations, such that $\mathbf{y'} = [\mathbf{y'_o}, \mathbf{y'_c}]$. If we let $\mu_i = \mathbf{x'_i b} + \mathbf{z'_i u}$, (i = 1,…n), then the density of the conditional distribution of the uncensored and censored records, given the parameters, can be written as:

$$f(\mathbf{y} \mid \mathbf{b}, \mathbf{u}, \sigma_e^2) = (2\pi\sigma_e^2)^{-m/2} \exp\left[\frac{-1}{2\sigma_e^2}\sum_{i=1}^{m}(y_{oi} - \mu_i)^2\right] \times \prod_{i=m+1}^{n}\left[1 - \Phi\left(\frac{c_i - \mu_i}{\sigma_e}\right)\right]$$

where y_{ci} is the unobserved DC, c_i is the value of the trait at censoring time, and Φ(.) is the standard normal cumulative distribution function.

Augmenting the posterior distribution with the unobserved calving dates corresponding to the censored observations (Tanner, 1996) simplifies the procedure. Let $\mathbf{w} = \{w_j\}$ with $w_j > c_j$, j = m+1, m+2, …n and let the augmented data vector be $\mathbf{y'_A} = [\mathbf{y'_o}, \mathbf{w'}]$. The joint density of calving dates, both observed and unobserved, given the parameters, is:

$$f(\mathbf{y_A} \mid \mathbf{b}, \mathbf{u}, \sigma_e^2) \propto (\sigma_e^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma_e^2}\left[\sum_{i=1}^{m}(y_{oi} - \mu_i)^2 + \sum_{i=m+1}^{n}(w_i - \mu_i)^2\right]\right\} \qquad [1]$$

where $w_i \geq c_i$ for all appropriate i values. Taking into account the augmented data **w,** the new vector of unknown parameters of the model will be $\boldsymbol{\theta}'_A = (\mathbf{b'}, \mathbf{u'}, \mathbf{w'}, \sigma^2_u, \sigma^2_e)$. To complete the Bayesian formulation, the following prior distributions were assumed to the unknown parameters:

$p(\mathbf{b}) \sim N(0,10^6)$ [2]

$p(\mathbf{u}|\sigma^2_u) \sim N(0,A\,\sigma^2_u)$ [3]

where **A** is the matrix of additive relationships between animals and σ^2_u is the additive variance

$p(\sigma^2_e) \sim U[0,10000]$ [4]

$p(\sigma^2_u) \sim U[0,1000]$ [5]

where U[.] is a uniform distribution.

The augmented joint posterior density is easily obtained as the product of distributions in [1] to [5]. Full conditional distributions needed for the implementation of Gibbs sampling can be found in Donoghue et al. (2003). These distributions are normal for the position parameters, truncated normal for the missing records and scaled inverted chi-square for the dispersion parameters.

Parameters were drawn from the conditional posterior distributions using Gibbs sampling. Pearson correlations between posterior means of animal effects for sires were calculated to evaluate the effect of each method on sire ranking. As an additional measure of effect on sire ranking, the percentage of sires in common in the top and bottom 10% of sires ranked with posterior means under different methods was calculated.

Results and Discussion

For all analyses, convergence was assessed using methodology presented by Raftery and Lewis (1992). The required length of the burn-in period was always less then 2,500 iterations for all parameters. Thus, 60,000 iterations of the sampler were run with a conservative 20,000 iterations discarded as burn-in; all remaining 40,000 iterations were retained without thinning for post-Gibbs analysis.

The mean and standard deviation of DC for uncensored females was 302±20 days for DCPEN and DCSIM, and 301±19 days for DCMISS. The corresponding statistics for DC for censored females were 354±25 days and 345±23 days for DCPEN and DCSIM, respectively. Under the DCSIM method, the number of days added to the largest observed DC record within a contemporary group ranged from 5 to 24 days. The majority of censored animals (62%) receiving a record ranging from 9 to 15 days greater than the largest observed DC record within their contemporary group, while relatively few animals (2%) received a censored record equal to the constant number of days (21) assigned under the DCPEN method.

Summaries of the posterior distributions of genetic parameters for DC under the different methods of handling censored records are presented in Table 4.2. Posterior means of the additive variance under DCPEN and DCSIM were similar, and both estimates were within the high posterior density interval HPD (95%) interval of the other method. The corresponding estimate for DCMISS, however, was significantly lower than these estimates, and was outside the HPD (95%) interval for both DCPEN and DCSIM. Since the DCMISS method does not consider censored females, thus ignoring an important source of genetic variation, it was expected that this method would produce lower estimates for additive variance. These results show that using either DCPEN or DCSIM to handle censored fertility records had little impact on estimation of additive variance, while ignoring censored records (DCMISS) produced significantly lower estimates of this parameter. In the previous simulation study, Donoghue et al. (2003)

also found that the method of handling censored records had little effect on additive variance estimation.

Posterior means of the residual variance under all methods were significantly different. In fact, the posterior mean for each method was outside the HPD (95%) interval for the other two methods. The estimate under DCMISS was the smallest of the values, while the estimate under DCPEN was the highest value. These results imply that the DCSIM method provides a better fit to the data than DCPEN when censored records are present. These findings follow the same trend evident for this parameter in the simulation study by Donoghue et al. (2003), from which the authors concluded that the higher estimates of residual variance using DCPEN indicated that adding a constant number of days to all censored records within a contemporary group was significantly overestimating the censored DC records, especially at higher levels of censoring. This suggests that even larger differences between DCPEN and DCSIM in the estimation of residual variance in field data would have occurred had the level of censoring been greater than the 12% observed in the current study.

Posterior means of the heritability under all methods were similar, and all estimates were within the HPD (95%) interval of the other two methods. As a result of the lower estimate of the additive variance under DCMISS, the posterior mean of heritability was slightly smaller than the other methods. The point estimates of heritability in this study for DCPEN and DCSIM are smaller than estimates reported in literature for DC or the equivalent trait of calving date for analyses when censored records are included (Johnston and Bunter, 1996; Morris et al., 2000; Johnston et al., 2001). However, our estimates are within the range of standard error associated with these literature estimates. The point estimate of heritability for DCMISS was similar and smaller, respectively, to the estimates obtained by Azzam and Nielsen (1987) and Meacham and Notter (1987) for calving date with censored records ignored.

Pearson correlations between posterior means of animal effects for sires based on different methods are given in Table 4.3. As expected, these correlations increased with the number of progeny across all methods. The correlation for all sires with daughters with records between DCPEN-DCSIM was 0.99, indicating that no major re-ranking of sires would be expected to occur across these methods. These findings confirm the results reported for ranking of sires in the simulation study by Donoghue et al. (2003). Lower correlations were observed between DCPEN/DCSIM and DCMISS for all sires with daughters with records (0.77-0.81), implying that some re-ranking of sires is occurring when censored records are ignored. Notter (1988) noted that inclusion of censored records is required to accurately estimate true sire differences in daughters' fertility, especially for sires that differ markedly in the frequency of censored daughters; this result confirms this belief. A similar trend to that observed for all sires was observed between DCPEN-DCSIM, and DCPEN/DCSIM with DCMISS for higher accuracy sires (sires with greater then 20 daughters with records). The lack of significant differences in the genetic ranking of sires between DCPEN and DCSIM suggests that either censored data handling technique can be successfully used in a genetic evaluation for days to calving.

The percentage of sires in common in the top and bottom 10% ranked on the basis of posterior means of animal effects for sires with daughters with records are given in Table 4.3. The concordance between evaluations for the top 10% of sires ranged between

60% (DCPEN-DCMISS) and 91% (DCPEN-DCSIM). Concordance between evaluations for the bottom 10% of sires was slightly lower, and ranged from 52% (DCPEN-DCMISS) to 89% (DCPEN-DCSIM). These results show that although most sires in the top or bottom 10% ranked similarly under DCPEN-DCSIM, there were exceptions. Further investigation of the sires who ranked outside the top or bottom 10% for the other method, as well as the biggest overall rank changes for all sires, revealed a general trend; a large proportion of their daughters belonged to contemporary groups which had a high percentage of censored records. Larger rank changes were observed for sires in the bottom 10% than for the top 10%, suggesting that perhaps the DCSIM method is doing a better job of identifying sires of inferior genetic merit.

Implications

Results for estimation of genetic parameters and ranking of breeding values for sires verify the findings of a previous simulation study. There were little differences observed in the heritability of the trait or ranking of sire breeding values between the penalty or simulation method. However, the larger estimate of the residual variance under the penalty method suggests that the simulation approach provides a better method for handling censored records in beef fertility data. The lower estimate of additive variance when censored records are ignored, in conjunction with the expected re-ranking of sires, highlight the unsuitability of this method for evaluation. The lack of significant differences in the genetic ranking of sires between the penalty and simulation methods suggests that either censored data handling technique can be successfully used in a genetic evaluation for days to calving.

Acknowledgements

Appreciation is extended to the Angus Society of Australia for providing the data; Meat and Livestock Australia for the research scholarship provided to the first author, and Dr. David Johnston and Carel Teseling for their contributions.

Literature Cited

Azzam, S. M. and M. K. Nielsen. 1987. Genetic parameters for gestation length, birth date and first breeding date in beef cattle. J. Anim. Sci. 64:348-356.

Buddenberg, B. J., C. J. Brown and A. H. Brown. 1990. Heritability estimates of calving date in Hereford cattle maintained on range under natural mating. J. Anim. Sci. 68:70-74.

Carriquiry, A. L., D. Gianola and R. L. Fernando. 1987. Mixed-Model analysis of a censored normal distribution with reference to animal breeding. Biometrics. 43:929-939.

Donoghue, K. A., R. Rekaya and J. K. Bertrand. 2003. Comparison of methods for handling censored fertility records in beef cattle data: Simulation study. J. Anim. Sci. (Submitted)

Johnston, D. J. and K. L. Bunter. 1996. Days to calving in Angus cattle: Genetic and environmental effects, and covariances with other traits. Livest. Prod. Sci. 45:13-22.

Johnston, D. J., J. M. Henshall and B. Tier. 2001. Estimate of the genetic correlation between calving success and days to calving in Angus females. Proc. Assoc. Advmt. Anim. Breed. Genet., Queenstown, New Zealand. pp 353-355.

Meacham, N. S. and D. R. Notter. 1987. Heritability estimates for calving date in Simmental cattle. J. Anim. Sci. 64:701-705.

Melton, B. E. 1995. Conception to consumption: The economics of genetic improvement. In: Proc. Beef Improvement Federation 27th Research Symposium and Annual Meeting, Sheridan, Wyoming. pp 40-47.

Meyer, K., K. Hammond, P. F. Parnell, M. J. Mackinnon and S. Sivarajasingam. 1990. Estimates of heritability and repeatability for reproductive traits in Australian beef cattle. Livest. Prod. Sci. 25:15-30.

Meyer, K., K. Hammond, M. J. Mackinnon and P. F. Parnell. 1991. Estimates of covariances between reproduction and growth in Australian beef cattle. J. Anim. Sci. 69:3533-3543.

Morris, C. A., J. A. Wilson, G. L. Bennett, N. G. Cullen, S. M. Hickey and J. C. Hunter. 2000. Genetic parameters for growth, puberty, and beef cow reproductive traits in a puberty selection experiment. New Zealand J. Agric. Res. 43:83-91.

Notter, D. R. 1988. Evaluating and reporting reproductive traits. In: Proc. Beef Improvement Federation 20th Research Symposium and Annual Meeting, Albuquerque, New Mexico. pp 21- 42.

Raftery, A. E. and Lewis, S. 1992. How many iterations in the Gibbs sampler? Bayesian Statistics 4, pp. 763-773 (ed. J. M. Bernando, J. O. Berger, A. P. Dawid and A. F. M. Smith, Oxford University Press).

Tanner, M. A. 1996. Tools for Statistical Inference. Springer-Verlag.

Table 4.1. Summary of data characteristics

Data	Records	Animals	Sires	Herds	CG
DCPEN	33176	62857	4238	470	3568
DCSIM	33176	62857	4238	470	3568
DCMISS	29020	56797	4045	459	3416

Table 4.2. Posterior means, standard deviations (SD) and bounds of highest probability density intervals (95%) under the three methods for handling censored data

Method	Parameter	Mean	SD	HPD (95%)
DCPEN	σ^2_u	22.56	4.24	15.59 – 30.28
	σ^2_e	431.39	5.04	421.33 – 440.92
	h^2	0.05	0.01	0.03 – 0.07
DCSIM	σ^2_u	26.10	3.61	19.21 – 33.75
	σ^2_e	371.36	4.49	362.2 – 380.0
	h^2	0.07	0.01	0.05 – 0.08
DCMISS	σ^2_u	13.51	2.89	8.56 – 18.69
	σ^2_e	262.21	3.41	255.6 – 268.8
	h^2	0.05	0.01	0.03 – 0.07

Table 4.3. Pearson correlations and percentage of top and bottom 10% sires in common using posterior means of animal effects for sires with progeny with records

	All sires[a]	>20 progeny[b]	Top 10%[c]	Bottom 10%[d]
DCPEN-DCSIM	0.99	0.99	91	89
DCPEN-DCMISS	0.77	0.80	60	52
DCSIM-DCMISS	0.81	0.84	62	56

[a] All sires with progeny with records (n=4238 for DCPEN-DCSIM and n=4045 for DCPEN/DCSIM-DCMISS)

[b] Sires with more than 20 progeny with records (n=322 for DCPEN-DCSIM and n=260 for DCPEN/DCSIM-DCMISS)

[c] Percentage of sires in common in the top 10% top sires

[d] Percentage of sires in common in the bottom 10% top sires

CHAPTER 5

GENETIC EVALUATION OF CALVING TO FIRST INSEMINATION USING NATURAL AND ARTIFICIAL INSEMINATION MATING DATA[1]

---

[1] Donoghue, K. A., R. Rekaya, J. K. Bertrand and I. Misztal. Submitted to *Journal of Animal Science*, 5/26/2003.

ABSTRACT

Mating and calving records for 51,084 first calf-heifers in Australian Angus herds were used to examine the relationship between probability of calving to first insemination (CFI) in artificial insemination (AI) and natural service (NS) mating data. Calving to first insemination was defined as a binary trait for both sources of data. Two Bayesian models were employed; 1) a bivariate threshold model with CFI in AI data regarded as a separate trait to CFI in NS data, and 2) a univariate threshold model with CFI regarded as the same trait for both sources of data. Posterior means (SD) of additive variance in the bivariate analysis were similar; 0.049 (0.013) and 0.075 (0.021) for AI and NS data, respectively; indicating lack of heterogeneity for this parameter. A similar trend was observed for heritability in the bivariate analysis, with posterior means (SD) of 0.025 (0.007) and 0.048 (0.012) for AI and NS data, respectively. The posterior means (SD) of the additive covariance and corresponding genetic correlation between the traits were 0.048 (0.006) and 0.821 (0.138), respectively. Some differences were observed between posterior means for herd-year variance; 0.843 vs. 0.280 for AI and NS data, respectively; which may be a reflection of the higher incidence of extreme category problem in the AI data. Parameter estimates under the univariate model were close to the weighted average of the corresponding parameters under the bivariate model. Posterior means (SD) for additive, herd-year and service sire variance and heritability under the univariate model were 0.063 (0.007), 0.56 (0.029), 0.131 (0.013) and 0.036 (0.007), respectively. These results indicate that, genetically, cows with a higher probability of CFI when mated using AI also have a high probability of CFI when mated via NS. The high correlation between the two traits, along with the lack of heterogeneity for the additive variance, implies that an analysis of CFI with a common additive variance for AI and NS data could be undertaken. The low estimates of heritability indicate that response to selection for probability of calving to first insemination would be expected to be low.

Key words: Beef cattle, Fertility

Introduction

Although fertility traits have high economic value to the beef industry (Melton, 1995), relatively little emphasis has been placed on these traits in genetic evaluations as they are difficult to measure, report, evaluate and interpret. Previous work investigating reproductive performance under natural service matings has focused on the continuous traits of calving date and days to calving (e.g. Buddenberg et al., 1990; Meyer et al., 1990; Johnston and Bunter, 1996). These studies found both traits to be heritable, as well as having a clear economic interpretation. Studies have also used information from artificial insemination mating data to evaluate reproductive performance (e.g. Evans et al., 1999; Doyle et al., 2000). Traits that describe female fertility measures arising from artificial insemination data conform well to the threshold model, and also exhibit genetic variation.

Little attention, however, has been given to genetic evaluations of fertility that incorporate both natural service and artificial insemination information. In order to develop such an evaluation, traits must be defined such that information from both natural and artificial matings can be fairly compared. A binary trait that evaluates the probability that an animal will produce a live calf from her first insemination could fulfill this criteria. Under this trait definition, an animal that is pregnant and produces a calf after her first opportunity to do so in artificial insemination data could be comparable to an animal that produces a calf as a result of becoming pregnant within the first 21 days of the breeding season in natural mating data.

The objective of this study was to develop a genetic evaluation of fertility in which information from both natural and artificial insemination matings could be included as a single trait. Firstly, a bivariate analysis was conducted, where probability of calving to first insemination was defined as a separate binary trait under the two different types of mating. Genetic variances were examined for homogeneity, and the genetic correlation between the trait under the types of mating was estimated in order to determine whether the sources of information could be combined. Secondly, a single trait analysis was conducted such that information from both types of matings was combined into a single trait.

Materials and Methods

Data

Bivariate Analysis. The data consisted of mating and calving records for first-calf females from Angus herds in Australia. Animals had either an artificial insemination (AI) or a natural service (NS) record, but not both types of records. Only animals having their first mating record between 270 and 625 days of age were included in the analysis. After editing, records from 51,084 females born between 1987 and 2000 were available for analysis. Edits performed included removal of 1) animals with incomplete records; 2) mating records resulting in multiple births and 3) single record contemporary groups. The final AI and NS data sets consisted of 16,358 and 34,726 records, respectively. A total of 2,239 and 4,465 sires were represented in the AI and NS data sets, respectively, with 1,658 sires having progeny in both data sets. The structure of the data sets are shown in Table 5.1.

For AI records, gestation length (GL) was computed as the difference between the insemination date and subsequent calving date, and averaged by sex of calf. Mating records whose GL was more than two standard deviations lower than the mean (by sex of calf) were considered outliers and removed from the data set (Kadarmideen and Coffey, 2001). Calving to first insemination for AI records (AICFI) was defined as a binary trait; GL records within two standard deviations (SD) of the mean (by sex of calf) were coded as 1, while records more than two SD above the mean, including animals who failed to calve, were coded as 0. An approximate date of insemination for natural service (NS) records was calculated by adjusting days to calving (DC) records, defined as the number of days between the time a bull is turned out in the pasture and the subsequent calving date, by mean GL (by sex of calf) obtained from the AI records. Mating records whose approximate date of insemination led to a GL that was more than two standard deviations lower than the mean (by sex of calf) were removed from the data set. Calving to first insemination for NS records (NSCFI) was defined as a binary trait; mating records whose approximate date of insemination was within the first 21 days of the breeding cycle were coded as 1, while records outside this range, including animals who failed to calve, were coded as 0.

Single trait Analysis. Records from the AI and NS data sets were combined into one data set for a single trait analysis of calving to first insemination (CFI). Descriptive statistics for the combined data set are found in Table 5.1.

Model

A threshold-liability model (Gianola and Foulley, 1983) was used for analysis of the binary data. The threshold model postulates an underlying continuous random variable or liability (λ), such that an observed binary response takes the value 1 if λ is larger than a fixed threshold (τ), and 0 otherwise.

Bivariate Analysis. A bivariate threshold model was used, with both AICFI and NSCFI as binary traits possessing an underlying bivariate normal distribution. These models have been implemented in the animal breeding field using marginal maximum likelihood or empirical Bayes methods by several researchers (Foulley et al., 1983; Hoeschele et al., 1986), and are an extension of the univariate threshold model (Gianola and Foulley, 1983).

A mixed linear model was used for analyses of the underlying liability for both AICFI and NSCFI. In matrix notation the model can be written as:

$$\boldsymbol{\lambda} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z_s s} + \mathbf{Z_u u} + \mathbf{e}$$

where $\boldsymbol{\lambda}$ is a vector of unobserved liabilities; $\boldsymbol{\beta}$ is the vector of fixed effects (herd × year effects; month of mating effects; linear and quadratic covariates for age at mating); $\mathbf{s}$ is the vector of service sire effects; $\mathbf{u}$ is the vector of additive genetic values and $\mathbf{e}$ is the vector of residual terms. $\mathbf{X}$, $\mathbf{Z_s}$ and $\mathbf{Z_u}$ are known incidence matrices with the appropriate dimensions.

Conditionally on the model parameters, it was assumed that the sampling distribution of liabilities was:

$$p(\boldsymbol{\lambda}|\boldsymbol{\beta},\mathbf{s},\mathbf{u},\mathbf{R_0}) \sim N(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z_s s} + \mathbf{Z_u u}, \mathbf{I} \otimes \mathbf{R_0})$$

where $\mathbf{R_0}$ is a 2×2 variance-covariance matrix with the following structure:

$$\mathbf{R_0} = \begin{bmatrix} \sigma^2_{e1} & \sigma_{e12} \\ \sigma_{e21} & \sigma^2_{e2} \end{bmatrix}$$

Given the well-known non-definability problem of the threshold models, at least two restrictions are needed. In this study, the threshold and the residual variances were arbitrarily set to zero and 1, respectively. Further, all animals had either an AI or a NS record; none had both traits measured. Consequently, the residual covariance cannot be inferred and it was set to zero leading to:

$$\mathbf{R_0} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

The vector of systematic effects was partitioned into sub-vectors $\beta = (\beta'_{H}, \beta'_{R})'$, where β_H is a vector of herd × year effects and β_R is a vector containing month of mating effects and linear and quadratic covariates for age at mating. The average number of records per herd × year class was small, and many herd × year classes were either all 0s or all 1s for any of the two binary traits. This causes the well-known problem of extreme case categories in analysis of threshold models. Hence, herd × year effects were assigned a normal prior with unknown mean and variance. This is based on the results of Rekaya et al. (2000), who found that this prior distribution alleviated the "extreme category" problems in such data. The prior distribution for the vector β_H was:

$$p(\beta_H|\eta,\mathbf{H}) \sim N(\eta, \mathbf{I} \otimes \mathbf{H})$$

where $\eta = (\mathbf{1}'\eta_1, \mathbf{1}'\eta_2)'$ is a vector of mean herd × year effects, with η_1= mean herd × year effect for AICFI and η_2= mean herd × year effect for NSCFI. Further,

$$\mathbf{H} = \begin{bmatrix} \sigma^2_{h1} & 0 \\ 0 & \sigma^2_{h2} \end{bmatrix}$$

is a 2 × 2 (co)variance matrix and **I** is the identity matrix. Both η_1 and η_2 and the elements of **H** were assumed unknown and, hence, priors were specified as follows:

$$p(\eta_i) \sim U[-10, 10], i = 1,2.$$

The diagonal elements of **H** were assigned the following uniform prior distributions:

$$p(\sigma^2_{hi}) \sim U[0,1] \text{ for } i = 1,2.$$

where U [.] is the uniform distribution.

The prior distribution for vector $\boldsymbol{\beta}_R$ was:

$$p(\boldsymbol{\beta}_R|\sigma^2_R) \sim N(\mathbf{0},\mathbf{I}\sigma^2_R)$$

with $\sigma^2_R = 10^5$. Since σ^2_R is large relative to the residual variance, this prior distribution conveys vague prior knowledge about each of the elements of $\boldsymbol{\beta}_R$.

A normal distribution was used as prior for the effect of service sire:

$$p(\mathbf{s}|\mathbf{S}) \sim N(0, \mathbf{I} \otimes \mathbf{S})$$

where $\mathbf{S} = \begin{bmatrix} \sigma^2_{s1} & 0 \\ 0 & \sigma^2_{s2} \end{bmatrix}$ is the service sire (co)variance matrix. The following uniform bounded priors were assigned to the diagonal elements of matrix **S**:

$$p(\sigma^2_{si}) \sim U[0,1] \text{ for } i = 1,2.$$

A multivariate normal distribution was used as prior for the animal effects:

$$p(\mathbf{u} |\mathbf{A}, \mathbf{G}_0) \sim N(\mathbf{0}, \mathbf{A} \otimes \mathbf{G}_0)$$

where $\mathbf{G_0} = \begin{bmatrix} \sigma^2_{u1} & \sigma_{u12} \\ \sigma_{u21} & \sigma^2_{u2} \end{bmatrix}$ is the additive (co)variance matrix, and **A** is the additive relationship matrix between animals. The following uniform bounded priors were assigned to the elements of matrix $\mathbf{G}_0$:

$$p(\sigma^2_{ui}) \sim U[0,1] \text{ for } i = 1,2,$$

and

$$p(\sigma_{u12}) \sim U\left[-\sqrt{\sigma_{u1}^2\sigma_{u2}^2}, \sqrt{\sigma_{u1}^2\sigma_{u2}^2}\right]$$

The joint posterior density is proportional to the product of the density of the conditional distribution of the observation times the joint prior density. Draws from the conditional posterior distribution of all the parameters were obtained using a Gibbs sampler with data augmentation (Sorensen et al., 1995). The joint posterior was augmented with the liabilities for each female. After augmentation, all the fully conditional posterior distributions of all model parameters can be derived as described by Albert and Chib (1993) and Sorensen et al. (1995). These distributions are normal for the location parameters; truncated normal for the liabilities; and scaled-inverted Wishart distributions for the dispersion parameters. Liabilities were sampled from their truncated normal distribution using inverse cumulative distribution function technique (Devroye, 1986).

Single trait Analysis. A univariate version of the above bivariate model was implemented. Similar priors (after adjustment of dimensions) were assumed to the model parameters. A full description of the derivation and implementation using a Bayesian approach via MCMC methods can be found Heringstad et al. (2001).

Results and Discussion

For all analyses, convergence was assessed using methodology presented by Raftery and Lewis (1992). The required length of the burn-in period was always less then 2,500 iterations for all parameters. Thus, 75,000 iterations of the sampler were run with a conservative 25,000 iterations discarded as burn-in; all remaining 50,000 iterations were retained without thinning for post-Gibbs analysis.

Summaries of the posterior distributions of (co)variance components, heritabilities and genetic correlation from the bivariate analysis are presented in Table 5.2. The posterior mean of the additive variance in the AI data (0.049) tended to be slightly lower than in NS data (0.075). This could be due to the nature of an AI program; a poor AI technician could reduce an animal's opportunity to express her genetic potential for fertility, thus leading to a lower estimate of additive variance. Despite this trend, the posterior mean of additive variance in the AI data was within the high posterior density interval HPD (95%) interval for the NS data. The posterior mean (SD) of the additive covariance between AICFI and NSCFI was 0.048 (0.006), and the corresponding genetic correlation was 0.821 (0.138). These results suggest a high, positive genetic correlation between CFI for AI matings and the same trait under NS matings. A large positive genetic correlation means that, genetically, cows with a higher probability of calving as a result of the first insemination when mated using AI also have a high probability of calving as a result of the first insemination when mated via natural service. These results indicate that there was no major statistical evidence of heterogeneity of variance for the additive effect in these two data sets. Together with the high positive correlation between both traits, it appears that a single trait analysis with homogenous additive variances is a reasonable assumption. Thus, combining AI and NS data could be undertaken for the trait of calving to first insemination, assuming that other variances in the model did not exhibit heterogeneity.

The posterior means (SD) of heritability (h^2) for CFI in both traits were low, and ranged from 0.03-0.05 (0.007-0.012). These heritabilities represent the lower bounds for the trait due to the inclusion of herd-year variance in the total phenotypic variance. By not including the herd-year variance in the total phenotypic variance, upper bounds of 0.044 and 0.056 could be computed for AICFI and NSCFI, respectively. As a result of the lower estimate of the additive variance in the AI data set, the posterior mean of both lower and upper bound of h^2 were slightly smaller than the estimate in the NS data. Early studies of fertility reported heritabilities expressed on the observed scale for traits such as pregnancy, conception and calving rates, and found heifer fertility to be lowly heritable (Koots et al., 1994). Buddenberg et al. (1990) examined the h^2 of pregnancy on both the observed and underlying scales in Angus, Hereford and Poll Hereford females. For first-calf females, these estimates ranged from 0.04 to 0.17 and 0.08 to 0.34 on the observed and underlying scale, respectively. Evans et al. (1999) used threshold analytical procedures and Method R to estimate variance components for heifer pregnancy, and obtained h^2 estimates of 0.138 ± 0.09, with a 90% confidence interval (CI) of 0.015 to 0.347. The point estimates of h^2 in our study are lower than the estimates reported in these studies, however, they are both within the 90% CI reported by Evans et al. (1999). The low estimates of h^2 observed in this study could result from more appropriate analytical procedures for the data; Bayesian approach vs. Method R for a small data set; or perhaps are a reflection of the slight differences in trait definitions.

The posterior mean of herd-year variance was higher in the AI data in comparison to the NS data (0.843 vs. 0.280). This could be due, in part, to the fact that insemination technician is nested within herd for AI data, and variation in heat detection and insemination techniques may have been partitioned to the herd-year variance in this data. However, this result is most likely caused by a higher incidence of extreme category problem (ECP) in the AI data than in the NS data; in the AI data, 50% of herd-year classes contained observations that fell into the same category, compared to 22% in the NS data. This bias in the estimate for herd-year variance is most likely positive in magnitude, as the majority of ECPs in the AI data were herd-year classes where all observations had a trait value equal to 1. The higher incidence of ECP in the AI data could be a sign of incomplete data recording for AI matings; that is, only successful AI matings are being reported in some herds. It could also be a reflection of different management levels under AI and NS mating~~s~~; due to the cost and higher level of management required to implement an AI program, these breeders may be employing more intensive techniques in order to maximize pregnancy rates. The different estimates for herd-year variance in AI and NS data sets indicate that some heterogeneity of variance may be occurring. However, as this is most likely a consequence of higher incidence of ECP in the AI data set, rather than real differences between the two data sets, this is not expected to be of major concern. Further research should be undertaken to investigate whether ECP is the cause of this heterogeneity, and if so, methods to reduce the impact of ECP on the variances should be explored.

The posterior means of service sire variance were lower in magnitude than the herd-year variance, but higher than the additive variance for both AI and NS data. As well, the posterior mean was higher in the AI data in comparison to the NS data (0.058 vs. 0.195). This difference is most likely a reflection of the smaller number of service sires in the AI data set, compared to the NS data set (687 vs. 3,504). However, this result may also reflect greater sire differences in NS data due to variation in libido, reduction in fertility between breeding soundness exam and actual use, and differences due to inexperience or bull aggressiveness. Further investigation is required to determine if service sire variances are really heterogeneous, and if so, the nature of these differences.

Summaries of the posterior distributions of the variance components and heritability from the single-trait analysis are presented in Table 5.3. The posterior mean of the additive variance for CFI was very close to the weighted mean of the point estimates for this parameter in the bivariate analysis. As well, both point estimates from the bivariate analysis were within the HPD (95%) for the parameter in the single-trait analysis. Pearson correlations between posterior means of animal effects for sires with progeny with records were 0.962, 0.970 and 0.968 for AICFI-NSCFI, AICFI-CFI and NSCFI-CFI, respectively. These correlations indicate that no major re-ranking of sires would be expected when AI and NS data is combined. These results, in conjunction with the lack of heterogeneity for additive variance observed for the bivariate analysis, indicate that fertility data from AI and NS matings can be combined using a common additive variance for the trait of CFI.

The same trend was observed for heritability for CFI, where the posterior mean was close to the weighted mean of the point estimates for this parameter in the bivariate analysis, and both point estimates from the bivariate analysis were within the HPD (95%) for the parameter in the single-trait analysis. As discussed previously, this heritability represents the lower bound for the trait due to the inclusion of herd-year variance in the total phenotypic variance. By not including the herd-year variance in the total phenotypic variance, an upper bound of 0.05 could be computed for the heritability of probability of of CFI. Both lower and upper bound estimates of h^2 for the probability of CFI were lower than previously reported estimates of heifer fertility (Evans et al., 1999), but within range of standard error associated with these literature estimates. These results indicate that the trait of probability of CFI, with AI and NS data combined for a single-trait analysis, is lowly heritable, and response to selection on this trait is expected to be small.

Posterior means for herd-year and service sire variance were also close to the weighted mean of the point estimates for these parameters in the bivariate analysis. However, for both parameters, point estimates from the bivariate analysis were outside the HPD (95%) for the corresponding parameter in the single-trait analysis. These results are most likely a reflection of the data structure for the AI data, in particular the higher incidence of ECP, as discussed previously. Further research is needed to determine the effect of data structure; specifically ECP; on inferences, and to decide the best method of handling data from each source. If heterogeneous variances do exist for herd-year and service sire, these need to be accounted for in the model.

Implications

The lack of heterogeneity observed for the additive variance in the bivariate analysis, in conjunction with the high, positive genetic correlation of calving to first insemination under artificial insemination and natural service matings, indicates that fertility data from these sources can be combined for a single-trait analysis using a common additive variance. The trait of probability of calving to first insemination had a low heritability under both bivariate and single-trait analyses, indicating that response to selection is expected to be low. Some differences were observed between the sources of data for herd-year and service sire variances, which are most likely a reflection of the poor structure for the artificial insemination data. Further research to investigate the presence of heterogeneity for these parameters based on model comparison is needed, as well as exploration of methods to better handle extreme case problem data.

Acknowledgements

Appreciation is extended to the Angus Society of Australia for providing the data; Meat and Livestock Australia for the research scholarship provided to the first author, and Dr. David Johnston and Carel Teseling for their contributions.

Literature Cited

Albert, J. and S. Chib. 1993. Bayesian analysis of binary and polychotomous response data. J. Am. Statist. Assoc. 88:669-679.

Buddenberg, B. J., C. J. Brown and A. H. Brown. 1990. Heritability estimates of calving date in Hereford cattle maintained on range under natural mating. J. Anim. Sci. 68:70-74.

Devroye, L. 1986. Non-Uniform Random Variate Generation. Springer Verlag.

Doyle, S. P., B. L. Golden, R. D. Green and J. S. Brinks. 2000. Additive genetic parameter estimates for heifer pregnancy and subsequent reproduction in Angus females. J. Anim. Sci. 78:2091-2098.

Evans, J. L., B. L. Golden, R. M. Bourdon and K. L. Long. 1999. Additive genetic relationships between heifer pregnancy and scrotal circumference in Hereford cattle. J. Anim. Sci. 77:2621-2628.

Foulley, J. L., D. Gianola and R. Thompson. 1983. Prediction of genetic merit from data on binary and quantitative variates with an application to calving difficulty, birth weight, and pelvic opening. Genet. Sel. Evol. 15:401-424.

Gianola, D. and J. L. Foulley. 1983. Sire evaluation for ordered categorical data with a threshold model. Genet. Sel. Evol. 15:201-223.

Heringstad, B. R., R. Rekaya, D. Gianola, G. Klemetsdal and K. A. Weigel. 2001. Bayesian analysis of liability to clinical mastitis in Norwegian cattle with a threshold model; effects of data sampling method and model specification. J Dairy Sci. 84:2337-2346.

Hoeschele, I., J. L. Foulley, J. J. Colleau and D. Gianola. 1986. Genetic evaluation for multiple binary responses. Genet. Sel. Evol. 18:299-320.

Johnston, D. J. and K. L. Bunter. 1996. Days to calving in Angus cattle: Genetic and environmental effects, and covariances with other traits. Livest. Prod. Sci. 45:13-22.

Kadarmideen, H. N. and M. P. Coffey. 2001. Quality and validation of insemination data for national genetic evaluations for dairy cow fertility in the United Kingdom. Proc. Interbull Meet. Budapest, Hungary, August 30-31 2001. *Interbull Bulletin 27.*

Koots, K. R., J. P. Gibson, C. Smith and J. W. Wilton. 1994a. Analyses of published genetic parameter estimates for beef production traits. 1. Heritability. Anim. Breed. Abstr. 62:309-338.

Melton, B. E. 1995. Conception to consumption: The economics of genetic improvement. In: Proc. Beef Improvement Federation 27th Research Symposium and Annual Meeting, Sheridan, Wyoming. pp 40-47.

Meyer, K., K. Hammond, P. F. Parnell, M. J. Mackinnon and S. Sivarajasingam. 1990. Estimates of heritability and repeatability for reproductive traits in Australian beef cattle. Livest. Prod. Sci. 25:15-30.

Raftery, A. E. and Lewis, S. 1992. How many iterations in the Gibbs sampler? Bayesian Statistics 4, pp. 763-773 (ed. J. M. Bernando, J. O. Berger, A. P. Dawid and A. F. M. Smith, Oxford University Press).

Rekaya, R., K. A. Weigel, D. Gianola, B. Heringstad and G. Klemetsdal. 2000. Methods for attenuating bias of variance component estimates in threshold models when herds are small. J. Dairy Sci. 83(Suppl. 1):56-57.

Sorensen, D. A., S. Andersen, D. Gianola and I. Korsgaard. 1995. Bayesian inference in threshold using Gibbs sampling. Genet. Sel. Evol. 27:229-249.

Table 5.1. Descriptive statistics of the datasets for the bivariate and single trait analyses

	AI	NS	Combined
Number of records	16358	34726	51084
Number of sires	2239	4465	5047
Number of herd×year classes	1136	2014	3150
Number of service sires	687	3504	3969
CFI[1] frequency, %	79.3	53.0	61.5

[1] Calving to first insemination

Table 5.2. Means, standard deviations (SD) and bounds of highest probability density (HPD) intervals (95%) of the posterior distribution of (co)variance components, heritabilities and correlations for bivariate analysis.

Trait	Parameter[1,2,3]	Mean	SD	HPD (95%)	
AICFI[4]	σ^2_h	0.843	0.095	0.685	0.999
	σ^2_s	0.058	0.014	0.034	0.085
	σ^2_u	0.049	0.013	0.032	0.074
	h^2	0.025	0.007	0.016	0.038
NSCFI[5]	σ^2_h	0.280	0.023	0.235	0.324
	σ^2_s	0.195	0.018	0.161	0.231
	σ^2_u	0.075	0.021	0.043	0.120
	h^2	0.048	0.012	0.031	0.077
AICFI-NSCFI	σ_{u12}	0.048	0.006	0.032	0.057
	r_g	0.821	0.138	0.535	0.999

[1] σ^2_h = herd-year variance; σ^2_s = service sire variance; σ^2_u = additive variance

[2] $h^2 = \sigma^2_u / (\sigma^2_h + \sigma^2_s + \sigma^2_u + 1)$

[3] σ_{u12} is the additive covariance between the two traits, respectively, and r_g is the corresponding correlation

[4] Probability of calving to first insemination in artificial insemination mating data

[5] Probability of calving to first insemination in natural service mating data

Table 5.3. Means, standard deviations (SD) and bounds of highest probability density intervals (95%) of herd-year variance (σ^2_h), service sire variance (σ^2_s), additive variance (σ^2_u), and of heritability[1] of probability of calving to first insemination for single trait analysis.

Parameter	Mean	SD	HPD (95%)	
σ^2_h	0.560	0.029	0.505	0.618
σ^2_s	0.131	0.013	0.107	0.157
σ^2_u	0.063	0.012	0.042	0.086
h^2	0.036	0.007	0.025	0.048

[1] $h^2 = \sigma^2_u / (\sigma^2_u + \sigma^2_h + \sigma^2_s + 1)$

CHAPTER 6

THRESHOLD-LINEAR ANALYSIS OF MEASURES OF FERTILITY IN ARTIFICIAL INSEMINATION DATA AND DAYS TO CALVING IN BEEF CATTLE[1]

---

[1] Donoghue, K. A., R. Rekaya, J. K. Bertrand and I. Misztal. Submitted to *Journal of Animal Science*, 7/3/03.

ABSTRACT

Mating and calving records for 47,533 first calf-heifers in Australian Angus herds were used to examine the relationship between days to calving (DC) and two measures of fertility in artificial insemination data; calving to first insemination (CFI) and calving success (CS). Calving to first insemination and calving success were defined as binary traits. A threshold-linear Bayesian model was employed for both analyses; 1) DC and CFI and 2) DC and CS. Posterior means (SD) of additive covariance and corresponding genetic correlation between the DC and CFI were –0.623 (0.193) and –0.662 (0.116), respectively. The corresponding point estimates between the DC and CS were –0.703 (0.142) and –0.728 (0.057), respectively. These genetic correlations indicate a strong, negative relationship between DC and both measure of fertility in AI data. Selecting for animals with shorter DC intervals, genetically, will lead to correlated increases in both CS and CFI. Posterior means (SD) for additive and residual variance and heritability for DC for the DC-CFI analysis were 23.5 (4.104), 363.17 (4.806) and 0.061 (0.01), respectively. The corresponding parameter estimates for the DC-CS analysis were very similar. Posterior means (SD) for additive, herd-year and service sire variance and heritability for CFI were 0.038 (0.013), 0.062 (0.064), 0.135 (0.158) and 0.031 (0.009), respectively. Posterior means (SD) for additive, herd-year and service sire variance and heritability for CS were 0.035 (0.012), 0.065 (0.068), 0.138 (0.155) and 0.03 (0.009), respectively. The similarity of the parameter estimates for CFI and CS suggest that either trait could be used as a measure of fertility in AI data. However, the definition of CFI allows the identification of animals that not only record a calving event, but calve to their first insemination, and the value of this trait would be even greater in a more complete

data set than used in this study. The magnitude of the correlations between DC and CS/CFI suggest that it may be possible to use a multi-trait approach to the evaluation of artificial insemination and natural service data, and report one genetic value that could be used for selection purposes.

Key words: Beef cattle, Heifer fertility, Genetic evaluation

Introduction

Little attention has been given to genetic evaluations of fertility that incorporate both natural service (NS) and artificial insemination (AI) information in one analysis. Previous work investigating reproductive performance under NS matings has focused on the traits of calving date and days to calving (DC) (e.g. Buddenberg et al., 1990; Johnston and Bunter, 1996), while studies using information from AI mating data have primarily evaluated binary traits, such as heifer pregnancy and calving success (CS).

Donoghue et al. (2003b) investigated calving to first insemination (CFI) as a potential trait under which information from both natural and artificial matings could be combined and fairly compared. These authors reported a genetic correlation of 0.82 between CFI under the two types of mating data, and concluded that cows with a higher probability of CFI when mated using AI also had a high probability of CFI when mated via NS.

An alternative method for combining both sources of mating information may be to fit a threshold-linear bivariate analysis to the data, and report one genetic value for fertility. Continuous measures of fertility, such as DC, have several advantages over binary traits like CS and CFI; they not only allow identification of animals more likely to conceive, but also the identification of animals that will conceive early in the breeding season. It may be desirable to report only the genetic value for DC when AI and NS mating are combined, given that reasonable genetic relationships exist with the binary trait.

The objective of this study was to investigate the genetic relationships between days to calving and measures of fertility in AI data (CS and CFI). Two bivariate analyses were conducted; DC with CFI; and DC with CS. Genetic correlations between DC and CFI/CS were used to examine the magnitude of the genetic relationships. As well, correlations between breeding values of CFI/CS and DC were compared for sires with progeny with records.

Materials and Methods

Data

The data consisted of mating and calving records for first-calf females from Angus herds in Australia. Animals had either an artificial insemination (AI) or a natural service (NS) record, but not both types of records. The majority of animals with AI records had only a single mating record, while a small number (n=2,277) had two mating records; AI mating followed by NS mating. The traits of calving success (CS) and calving to first insemination (CFI) were defined only for AI records, while days to calving (DC) was defined only for NS records. Only animals having their first mating record between 270 and 625 days of age were included in the analysis. After editing, records from 47,533 females born between 1987 and 2000 were available for analysis. Edits performed for all traits included removal of 1) animals with incomplete records and 2) mating records resulting in multiple births. In addition, single record contemporary groups were removed from the AI data, while contemporary groups with less than four records and contemporary groups consisting of only non-calvers were removed from the NS data. The final AI and NS data sets consisted of 16,358 and 31,175 records, respectively. The total number of animals including ancestors was 78,912. A total of 2,239 and 3,945 sires were represented in the AI and NS data sets, respectively, with 1,547 sires having progeny in both data sets. The measures of fertility for AI matings were recorded for the same animals, and only differed in definition; calving to first insemination by AI sire (CFI) vs. calving to any insemination (CS). The structure of the data sets are shown in Table 6.1, with summary statistics for CFI and CS combined.

Calving to first insemination (CFI) was recorded only for AI data, and was defined as a binary trait. Gestation length (GL) records within two standard deviations of the mean (by sex of calf) were coded as 1, while records more than two SD above the mean, including animals who failed to calve, were coded as 0. GL was computed as the difference between the insemination date and subsequent calving date, and averaged by sex of calf. Mating records whose GL was more than two standard deviations lower than the mean (by sex of calf) were considered outliers and removed from the data set (Kadarmideen and Coffey, 2001). Non-genetic effects were linear and quadratic covariates for age at mating, month of mating, and random effects of service sire and contemporary group. Contemporary group was defined to include animals from the same herd mated in the same year. The mean incidence of calving to first insemination was 79%.

Calving success (CS) was recorded only for AI data, and was defined as a binary trait. Females that calved, whether to an AI mating or a subsequent NS mating, were coded as 1, while animals that failed to calve were coded as 0. Non-genetic effects were linear and quadratic covariates for age at mating, month of mating, and random effects of service sire and contemporary group. Service sire was designated as the first sire to whom the female was mated; that is, the AI mating sire. Contemporary group was defined to include animals from the same herd mated in the same year. The mean incidence of calving success was 92%.

The continuous trait of days to calving (DC) was recorded only for NS data, and was defined as the number of days between the time a bull was turned out in the pasture and the subsequent parturition of the female. Non-genetic effects were linear and quadratic covariates for age at mating, sex of calf and contemporary group. Contemporary group was defined to include animals from the same herd who were mated to the same service sire in the same month and year. The sex of calf effect was arbitrarily set to female for all animals with censored records. Using the approach of Donoghue et al. (2003), trait values for censored records were simulated from their respective predictive distributions (truncated normal distributions). For all animals in the same contemporary group, the truncation point was the largest observed DC record. The predicted DC for a censored record was between the truncation point and positive infinity. Thus, an animal with a censored record could not receive a simulated record that was smaller than a non-censored record within her contemporary group. The number of days added to this truncation point for each of the censored records was determined by drawing samples at random from the truncated distribution, and depended upon the fixed effects in the model, as well as her relationships with other animals.

Model

Two analyses were undertaken; DC with CFI; and DC with CS. Both models were bivariate, with CFI or CS as binary (threshold) and DC as Gaussian. Here it was postulated that liabilities of CFI/CS and DC were jointly Gaussian. This model was developed by Foulley et al. (1983).

The following mixed linear animal model was used for both analyses of traits:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_s\mathbf{s} + \mathbf{Z}_u\mathbf{u} + \mathbf{e}$$

where $\mathbf{y}$ was a vector of DC observations or unobserved liabilities of CFI/CS; $\boldsymbol{\beta}=(\boldsymbol{\beta}'_1,\boldsymbol{\beta}'_2)'$ was the vector of systematic effects; $\mathbf{s}$ was the vector of service sire effects (for CFI and CS only); $\mathbf{u}$ was the vector of additive genetic values and $\mathbf{e}$ was the vector of residual terms. $\mathbf{X}$, $\mathbf{Z}_s$ and $\mathbf{Z}_u$ were known incidence matrices with the appropriate dimensions. For the traits of CFI and CS, $\boldsymbol{\beta}_1$ included herd × year effects; month of mating effects and linear and quadratic covariates for age at mating. For the trait of DC, $\boldsymbol{\beta}_2$ included contemporary group effects; sex of calf effects and linear and quadratic covariates for age at mating. The random effect of service sire was not included for the trait of DC. This effect was included as part of the contemporary group definition for DC, which is the current practice in the genetic evaluation program in Australia (Schneeberger et al., 1991).

Conditionally on the model parameters, it was assumed that the sampling distribution of observations was:

$$p(\mathbf{y}|\boldsymbol{\beta},\mathbf{s},\mathbf{u},\mathbf{R}_0) \sim N(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_s\mathbf{s} + \mathbf{Z}_u\mathbf{u}, \mathbf{I} \otimes \mathbf{R}_0)$$

where $\mathbf{R}_0$ is a 2 × 2 variance-covariance matrix with the following structure:

$$\mathbf{R}_0 = \begin{bmatrix} \sigma^2_{e1} & \sigma_{e12} \\ \sigma_{e21} & \sigma^2_{e2} \end{bmatrix}$$

Given the well-known non-definability problem of threshold models, at least two restrictions were needed. In this study for the traits of CFI and CS, the threshold and residual variance (σ^2_{e1}) were arbitrarily set to zero and 1, respectively. Further, all animals had either an AI or a NS record; none had both traits measured. Consequently, the residual covariance ($\sigma_{e12} = \sigma_{e21}$) cannot be inferred and was set to zero. The prior distribution for the residual variance for DC was:

$$p(\sigma^2_{e2}) \sim U(0,10000)$$

For the traits of CS and CFI, the vector of systematic effects was partitioned into sub-vectors $\boldsymbol{\beta}_1 = (\boldsymbol{\beta}'_H, \boldsymbol{\beta}'_R)'$, where $\boldsymbol{\beta}_H$ was a vector of herd × year effects and $\boldsymbol{\beta}_R$ was a vector containing month of mating effects and linear and quadratic covariates for age at mating. The average number of records per herd × year class was small, and many herd × year classes were either all 0s or all 1s for any of the two binary traits. This causes the well-known problem of extreme case categories in analysis of threshold models. Hence, herd × year effects were assigned a normal prior with unknown mean and variance. This is based on the results of Rekaya et al. (2000), who found that this prior distribution alleviated the "extreme category" problems in such data. The prior distribution for the vector $\boldsymbol{\beta}_H$ was:

$$p(\boldsymbol{\beta}_H|\eta,\sigma^2_h) \sim N(\eta, \mathbf{I}\sigma^2_h)$$

where η and σ^2_h are the mean and variance of herd × year effects, respectively. Both η and σ^2_h were assumed unknown and, hence, priors were specified as follows:

$$p(\eta) \sim U[-10, 10]$$

$$p(\sigma^2_h) \sim U[0,1]$$

where U [.] is the uniform distribution.

The prior distribution for vector $\boldsymbol{\beta}_\mathbf{R}$ was:

$$p(\boldsymbol{\beta}_R|\sigma^2_R) \sim N(\mathbf{0},\mathbf{I}\sigma^2_R)$$

with $\sigma^2_R = 10^5$. Since σ^2_R is large relative to the residual variance, this prior distribution conveys vague prior knowledge about each of the elements of $\boldsymbol{\beta}_R$.

For the trait of DC the following prior distribution was assumed for the systematic effects:

$$p(\boldsymbol{\beta}_2) \sim N(0,10^6)$$

A normal distribution was used as prior for the effect of service sire for the traits of CFI and CS:

$$p(\mathbf{s}|\sigma^2_s) \sim N(0,\sigma^2_s)$$

The following uniform bounded prior was assigned to the service sire variance for these traits:

$$p(\sigma^2_s) \sim U[0,1]$$

A multivariate normal distribution was used as prior for the animal effects:

$$p(\mathbf{u}\,|\mathbf{A}, \mathbf{G}_0) \sim N(\mathbf{0}, \mathbf{A} \otimes \mathbf{G}_0)$$

where $\mathbf{G_0} = \begin{bmatrix} \sigma^2_{u1} & \sigma_{u12} \\ \sigma_{u21} & \sigma^2_{u2} \end{bmatrix}$ was the additive (co)variance matrix, and $\mathbf{A}$ was the additive relationship matrix between animals. A conjugate proper prior was assumed. The scaling factor for the prior of $\mathbf{G}_0$, shown below, was taken from the literature (Johnston et al., 2001). The degree of belief a priori were set to 5 conveying little weight to the prior information.

$$\mathbf{G_0} = \begin{bmatrix} 25 & -0.8 \\ & 0.05 \end{bmatrix}$$

The joint prior density had the form:

$$p(\boldsymbol{\beta}_1, \boldsymbol{\beta}_R, \boldsymbol{\eta}, \sigma^2_h, \mathbf{s}, \mathbf{G}_0, \sigma^2_{e2}) =$$

$$p(\boldsymbol{\beta}_1)\ p(\boldsymbol{\beta}_R)\ p(\boldsymbol{\beta}_H|\boldsymbol{\eta},\sigma^2_h)\ p(\boldsymbol{\eta})\ p(\sigma^2_h)\ p(\mathbf{s}|\sigma^2_s)\ p(\sigma^2_s)\ p(\sigma^2_s)\ p(\mathbf{u}|\mathbf{G}_0)\ p(\mathbf{G}_0)\ p(\sigma^2_{e2})$$

The joint posterior density is proportional to the product of the density of the conditional distribution of the observation times the joint prior density. Draws from the conditional posterior distribution of all the parameters were obtained using a Gibbs sampler with data augmentation (Sorensen et al., 1995). The joint posterior was augmented with the univariate normal liabilities for CFI/CS. After augmentation, all the fully conditional posterior distributions of model parameters can be derived as described by Albert and Chib (1993) and Sorensen et al. (1995). These distributions are normal for the systematic parameters, service sire and animal effects; truncated normal for the liabilities; scaled-inverted chi square for the residual variance for DC ; and scaled-inverted Wishart distributions for the dispersion parameters. Liabilities were sampled from their truncated normal distribution using inverse cumulative distribution function technique (Devroye, 1986). The posterior distribution was augmented with the unobserved calving dates corresponding to the censored DC observations as described in Donoghue et al. (2003a).

(Co)variance components were estimated for both analyses, and genetic correlations between CFI/CS with DC were compared. Breeding values were predicted for all animals, and correlations between breeding values of CFI/CS and DC were compared for sires with progeny with records.

Results and Discussion

For all analyses, convergence was assessed using methodology presented by Raftery and Lewis (1992). The required length of the burn-in period was always less then 5,000 iterations for all parameters. Thus, 200,000 iterations of the sampler were run with a conservative 50,000 iterations discarded as burn-in; the remaining 150,000 iterations were retained with thinning for post-Gibbs analysis.

The mean and standard deviation of DC for uncensored females was 302±19, while the corresponding statistic for censored females was 354±25 days. The number of days added to the largest observed DC record within a contemporary group ranged from 5 to 26 days. The majority of censored animals (63%) received a record ranging from 10 to 16 days greater than the largest observed DC record within their contemporary group, while 21, 14 and 2% of censored females received records with 5-9, 17-21 and 21-26 days added, respectively.

Summaries of the posterior distributions of (co)variance components, heritabilities and genetic correlation from the CFI-DC bivariate analysis are presented in Table 6.2. The posterior mean (SD) of the additive covariance between CFI and DC was –0.623 (0.193), and the corresponding genetic correlation was –0.662 (0.116). These results suggest a high, negative correlation between probability of CFI and DC. A large, negative correlation indicates that, genetically, cows with a higher probability of calving to their first insemination will also record a shorter DC interval. Thus, selection for increased probability of CFI would result in correlated decreases in DC interval. This value is similar in magnitude to that reported by Johnston et al. (2001) for the genetic correlation between DC and calving success (-0.66). Morris et al. (2000) also found a similar relationship between higher pregnancy rates and early calving dates in Angus heifers, but did not report the magnitude of this association.

The posterior means (SD) for additive variance and heritability for DC were similar to previous estimates reported for this population (Donoghue et al., 2003a); furthermore, both estimates were within the high posterior density (HPD) (95%) intervals presented in the earlier study. The point estimates of heritability were smaller than estimates reported in literature for DC or the equivalent trait of calving date for analyses when censored records are included (Johnston and Bunter, 1996; Morris et al., 2000; Johnston et al., 2001). However, our estimates were within the range of standard error associated with these literature estimates. The point estimate of residual variance for DC was similar to estimates previously reported for this population (Donoghue et al., 2003a). In the current study the effect of service sire was included as part of the contemporary group definition to align with current industry practices for evaluation of DC. An additional analysis was undertaken with the random effect of service sire included in the model. The results from this analysis (unreported) were very similar to the results obtained in the current study. Fitting service sire as an additional random effect allowed the formation of larger contemporary groups, and had little effect on the estimation of additive variance, or on heritability of the trait. Thus, including service sire as a random effect may be a reasonable option to avoid small contemporary groups.

The posterior means for additive variance and heritability for CFI were similar to previous estimates reported for this population (Donoghue et al., 2003b); furthermore, both estimates were within the HPD (95%) intervals presented in the earlier study. The point estimate of h^2 in the current study was lower than estimates reported in literature for heifer fertility, however, it is within the 90% CI reported by Evans et al. (1999) for the trait of heifer pregnancy. The low estimates of h^2 observed in this study could result from more appropriate analytical procedures for data analysis; Bayesian approach vs. Method R for a small data set; or perhaps are a reflection of the slight differences in trait definitions between CFI and heifer pregnancy.

The posterior mean of herd-year variance for CFI in this study was much smaller than reported in an earlier study for the same population; 0.062 vs. 0.843 (Donoghue et al., 2003b). The variable nature of this parameter is most likely caused by a high incidence of extreme category problem (ECP) in the AI data; 50% of herd-year classes contained observations that fell into the same category for CFI. Possible reasons for this high incidence of ECP in the AI data were discussed in the earlier study, and included incomplete data recording and implementation of different management levels under AI matings. The highly fluctuative nature of this parameter suggests that caution should be used in the interpretation of the herd-year variance in the presence of ECP. Despite the large difference in this parameter observed between this study and the earlier study for CFI, additive variance and heritability point estimates were very similar in both studies, indicating that estimation of these parameters appears stable.

The posterior mean of service sire variance for CFI was larger in magnitude than both herd-year and additive variances. The point estimate in the current study was higher then the estimate reported in the earlier study for this population, and was outside the HPD (95%) interval for that study. However, examination of the standard deviation and HPD (95%) interval associated with this point estimate indicates lack of statistical evidence that it was significantly different from zero. This result was most likely a reflection of the small number of service sires (n=687) represented in the AI data.

Summaries of the posterior distributions of (co)variance components, heritabilities and genetic correlation from the CS-DC bivariate analysis are presented in Table 6.3. The posterior mean (SD) of the additive covariance between CS and DC was -0.703 (0.142), and the corresponding genetic correlation was -0.728 (0.057). These results suggest a high, negative correlation between probability of CS and DC. A large, negative correlation indicates that, genetically, females with a higher probability of calving success will also record a shorter DC interval. Thus, selection for increased probability of CS would result in correlated decreases in DC interval. This value is similar in magnitude to that reported by Johnston et al. (2001) for the genetic correlation between the same traits (-0.66), and is slightly larger than the genetic correlation between CFI and DC (-0.686) in the current study. However, the standard deviations and HPD (95%) intervals associated with the point estimates of genetic correlations between CS/CFI and DC in the current study indicate that they were not significantly different from each other.

The posterior means for additive, service sire and residual variances DC were very similar to estimates for DC under the DC-CFI bivariate; furthermore, all estimates were within the high posterior density (HPD) (95%) intervals for this analysis.

The posterior means for additive variance and heritability for CS in this study were low in magnitude. The point estimate of h^2 for CS is very similar to the estimate for the same trait reported by Johnston et al. (2001), when the categorical nature of the trait was taken into account. Higher estimates of h^2 for CS have been reported in literature when the trait was analyzed without adjustment for the categorical nature (Johnston and Bunter, 1996). The point estimates for all parameters for CS were very close to the corresponding estimates for CFI, with all estimates within the high posterior density (HPD) (95%) intervals for CFI. These results may reflect the similar definitions of the two traits, and the fact that only a small number of animals changed from a trait value of 0 under CFI to a trait value of 1 under CS (n=2,219). The level of ECP was high under both trait definitions; 50 and 74 % of herd-year classes contained observations that fell into the same category for CFI and CS, respectively. The similarity of results observed in this study between CFI and CS may not hold for a data set when complete recording is available; that is, both successful and unsuccessful AI matings are reported, as well as information regarding NS matings that may follow. The event of CFI would be expected to be lower under complete recording, with lower incidence of ECP.

There were 183 sires with more than 10 daughters with CFI/CS records, as well as more than 10 daughters with DC records. The mean (SD) DC breeding value (BV) for these sires was 0.30 days (2.57) with a range from –5.60 to +8.67 days. These BV were predicted using parameters from the DC-CFI analysis, and are not reported for DC BV under the DC-CS analysis, due to the similarity of the parameters. The mean (SD) CS BV was -0.02% (0.08), ranging from –0.28 to +0.17%. The mean (SD) CFI BV was -0.01% (0.08), ranging from –0.25 to +0.16%. The correlations between DC and CS BVs, and DC and CFI BVs for these sires were –0.993 and –0.997, respectively. These results indicate that sires whose daughters have either an increased probability of calving success or an increased probability of calving to first insemination also produce daughters with shorter DC records. The regression coefficient of CFI on DC was –0.03 % successful calving to first insemination/day; for every 1 day decrease in DC BV, there is a 0.03% increase in CFI BV, and a similar result was observed for the regression coefficient between CS and DC.

Implications

The potential of calving to first insemination as a measure of fertility in artificial insemination data has been confirmed; the parameter estimates for this trait were very close to those of calving success. The definition of calving to first insemination allows the identification of animals that not only record a calving event, but calve to their first insemination. The value of this trait would be even greater in a more complete data set than used in this study. The genetic correlations reported between days to calving and both measures of fertility in artificial insemination data (calving success and calving to first insemination) indicate a strong, negative relationship between these traits. Selecting for animals with shorter days to calving intervals, genetically, will lead to correlated increases in both calving success and calving to first insemination. The magnitude of these correlations suggest that it may be possible to use a multi-trait approach to the evaluation of artificial insemination and natural service data, but report one genetic value that could be used for selection purposes.

Acknowledgements

Appreciation is extended to the Angus Society of Australia for providing the data; Meat and Livestock Australia for the research scholarship provided to the first author, and Dr. David Johnston and Carel Teseling for their contributions.

Literature Cited

Albert, J. and S. Chib. 1993. Bayesian analysis of binary and polychotomous response data. J. Am. Statist. Assoc. 88:669-679.

Buddenberg, B. J., C. J. Brown and A. H. Brown. 1990. Heritability estimates of calving date in Hereford cattle maintained on range under natural mating. J. Anim. Sci. 68:70-74.

Devroye, L. 1986. Non-Uniform Random Variate Generation. Springer Verlag.

Donoghue, K. A., R. Rekaya and J. K. Bertrand. 2003a. Comparison of methods for handling censored records in beef fertility data: Field data. J. Anim. Sci. (Submitted).

Donoghue, K. A., R. Rekaya, J. K. Bertrand and I. Misztal. 2003b. Genetic evaluation of calving to first insemination using natural and artificial insemination mating data. J. Anim. Sci. (Submitted).

Doyle, S. P., B. L. Golden, R. D. Green and J. S. Brinks. 2000. Additive genetic parameter estimates for heifer pregnancy and subsequent reproduction in Angus females. J. Anim. Sci. 78:2091-2098.

Evans, J. L., B. L. Golden, R. M. Bourdon and K. L. Long. 1999. Additive genetic relationships between heifer pregnancy and scrotal circumference in Hereford cattle. J. Anim. Sci. 77:2621-2628.

Foulley, J. L., D. Gianola and R. Thompson. 1983. Prediction of genetic merit from data on binary and quantitative variates with an application to calving difficulty, birth weight, and pelvic opening. Genet. Sel. Evol. 15:401-424.

Johnston, D. J. and K. L. Bunter. 1996. Days to calving in Angus cattle: Genetic and environmental effects, and covariances with other traits. Livest. Prod. Sci. 45:13-22.

Johnston, D. J., J. M. Henshall and B. Tier. 2001. Estimate of the genetic correlation between calving success and days to calving in Angus females. Proc. Assoc. Advmt. Anim. Breed. Genet., Queenstown, New Zealand. pp 353-355.

Kadarmideen, H. N. and M. P. Coffey. 2001. Quality and validation of insemination data for national genetic evaluations for dairy cow fertility in the United Kingdom. Proc. Interbull Meet. Budapest, Hungary, August 30-31 2001. *Interbull Bulletin 27.*

Morris, C. A., J. A. Wilson, G. L. Bennett, N. G. Cullen, S. M. Hickey and J. C. Hunter. 2000. Genetic parameters for growth, puberty, and beef cow reproductive traits in a puberty selection experiment. New Zealand J. Agric. Res. 43:83-91.

Raftery, A. E. and Lewis, S. 1992. How many iterations in the Gibbs sampler? Bayesian Statistics 4, pp. 763-773 (ed. J. M. Bernando, J. O. Berger, A. P. Dawid and A. F. M. Smith, Oxford University Press).

Rekaya, R., K. A. Weigel, D. Gianola, B. Heringstad and G. Klemetsdal. 2000. Methods for attenuating bias of variance component estimates in threshold models when herds are small. J. Dairy Sci. 83(Suppl. 1):56-57.

Schneeberger, M., B. Tier, and K. Hammond. 1991. Introducing the third generation of BREEDPLAN and GROUP BREEDPLAN. Proceedings of the Ninth Conference of the Australia Association of Animal Breeding and Genetics, Melbourne, Australia p. 194-199.

Sorensen, D. A., S. Andersen, D. Gianola and I. Korsgaard. 1995. Bayesian inference in threshold using Gibbs sampling. Genet. Sel. Evol. 27:229-249.

Table 6.1. Descriptive statistics of the datasets for CFI/CS[1,2] and DC[3]

	CS/CFI	DC
Number of records	16,358	31,175
Number of sires	2,239	3,945
Number of contemporary groups	1,136	1,730
Number of service sires	687	2,539

[1] Calving to first insemination

[2] Calving success

[3] Days to calving

Table 6.2. Means, standard deviations (SD) and bounds of highest probability density (HPD) intervals (95%) of the posterior distribution of (co)variance components, heritabilities and correlations for CFI-DC bivariate analysis.

Trait	Parameter[1,2,3]	Mean	SD	HPD (95%)	
CFI[4]	σ^2_h	0.062	0.064	0.005	0.192
	σ^2_s	0.135	0.158	0.006	0.490
	σ^2_u	0.038	0.013	0.019	0.067
	h^2	0.031	0.009	0.014	0.049
DC[5]	σ^2_e	363.17	4.806	353.6	372.4
	σ^2_u	23.50	4.104	16.47	32.62
	h^2	0.061	0.010	0.043	0.084
CFI-DC	σ_{u12}	-0.623	0.193	-1.038	-0.312
	r_g	-0.662	0.116	-0.885	-0.438

[1] σ^2_h = herd-year variance; σ^2_s = service sire variance; σ^2_u = additive variance; σ^2_e = residual variance

[2] CFI: $h^2 = \sigma^2_u / (\sigma^2_h + \sigma^2_s + \sigma^2_u + 1)$; DC: $h^2 = \sigma^2_u / (\sigma^2_u + \sigma^2_e)$

[3] σ_{u12} is the additive covariance between the two traits, respectively, and r_g is the corresponding correlation

[4] Probability of calving to first insemination

[5] Days to calving

Table 6.3. Means, standard deviations (SD) and bounds of highest probability density (HPD) intervals (95%) of the posterior distribution of (co)variance components, heritabilities and correlations for HP-DC bivariate analysis.

Trait	Parameter[1,2,3]	Mean	SD	HPD (95%)	
CFI[4]	σ^2_h	0.065	0.068	0.063	0.188
	σ^2_s	0.138	0.155	0.004	0.442
	σ^2_u	0.035	0.012	0.013	0.062
	h^2	0.030	0.009	0.012	0.049
DC[5]	σ^2_e	363.20	4.801	353.8	372.6
	σ^2_u	23.54	4.105	15.45	31.54
	h^2	0.064	0.011	0.040	0.081
CFI-DC	σ_{u12}	-0.703	0.142	-1.064	-0.287
	r_g	-0.728	0.057	-0.912	-0.448

---

[1] σ^2_h = herd-year variance; σ^2_s = service sire variance; σ^2_u = additive variance; σ^2_e = residual variance

[2] HP: $h^2 = \sigma^2_u / (\sigma^2_h + \sigma^2_s + \sigma^2_u + 1)$; DC: $h^2 = \sigma^2_u / (\sigma^2_u + \sigma^2_e)$

[3] σ_{u12} is the additive covariance between the two traits, respectively, and r_g is the corresponding correlation

[4] Probability of calving success

[5] Days to calving

CHAPTER 7

CONCLUSIONS

The results from these studies indicate the potential to further enhance current genetic evaluations of fertility in beef cattle. The first two studies indicated the superiority of a more sophisticated method of handling censored records for the trait of days to calving. The refinement of this continuous measure of fertility is important, because it has several advantages over binary traits; it not only allows identification of animals more likely to conceive, but also identification of animals that will conceive early in the breeding season. The third study confirmed the feasibility of implementing a single trait evaluation of a new binary trait (calving to first insemination) to combine mating data from artificial insemination and natural service. Artificial insemination data conforms well to a binary trait approach; thus, incorporation of this data into an evaluation may need to focus on binary traits. The fourth study confirmed the presence of a strong, negative relationship between days to calving and calving to first insemination. This result suggests that selecting for animals with shorter days to calving intervals, genetically, will lead to correlated increases in calving to first insemination. Thus, it may be possible to use a multi-trait approach, and report one genetic value (days to calving) that could be used for selection purposes. The clearer and more meaningful economic interpretations associated with genetic values for days to calving highlight it as the preferred trait to report for use by industry.