

## **ABSTRACT**

VIERHOUT,CRYSTAL. Selection of Dairy Cow Families for Superior Fertility.  
(Under the direction of E. J. Eisen, R. L. McCraw, and S. P. Washburn).

The objective of this study was to determine if dairy cow families could be used to select for superior pregnancy rate. Holstein cow records in 13 southern states were obtained from Animal Improvement Laboratory of USDA. Cows were included from historical records dating back to birth year of 1981 or 1982 as the foundation cows. Historical records included cows calving and completing lactations through August, 2005. Cows from various generations were assigned to maternal family groups using dam identification within herd. Standardized milk production and pregnancy rate (based on days open) deviations were obtained within herd-year-season. A family value was calculated by averaging the first and second lactations across parity by degree of relationship to the individual (free of progeny information) for generation one through four. Each family was entered into one of three groups based on average deviations for milk production, pregnancy rate, and combining pregnancy rate and milk into a selection index. Analyses were performed on fifth generation members to determine if milk production and pregnancy rates in the fifth generation were significantly associated with historical performance of the respective cow families. Average of standardized values for milk production improved (8,542 to 10,275 kg) from generation one to five whereas 21-d pregnancy rate decreased from (26.9 to 18.3) in the same period for cow families having daughters represented in the fifth generation. After adjustments for sire predicted transmitting ability (PTA), maternal-

grandsire PTA, and family group for milk or pregnancy rate in the model the effect of maternal cow family remained significant for pregnancy rate ( $P < 0.05$ ,  $R^2=0.0759$ ) and milk production ( $P < 0.05$ ,  $R^2=0.1192$ ) when single trait selection was utilized. Utilizing a selection index with equal weights for milk and pregnancy rate, the effect of maternal cow family remained significant ( $P < 0.05$ ,  $R^2=0.0208$ ) for milk but pregnancy rate was not significant ( $P > 0.05$ ,  $R^2=0.0106$ ). These findings provide evidence that pregnancy rate and milk production are inherited through the maternal lineage. Thus, there may be potential economic value in considering maternal family history for pregnancy rate when selecting future bull dams.

Selection of Dairy Cow Families for Superior Fertility

by  
Crystal N. Vierhout

A dissertation submitted to the Graduate Faculty of  
North Carolina State University  
In partial fulfillment of the  
Requirements for the degree of  
Doctor of Philosophy

Animal Science & Poultry Science

Raleigh, NC

2008

APPROVED BY:

---

Dr. Kevin Anderson

---

Dr. David A. Dickey

---

Dr. Eugene J. Eisen  
Co-chair of Advisory Committee

---

Dr. Roger L. McCraw  
Co-chair of Advisory Committee

---

Dr. Steven P. Washburn

## **DEDICATION**

My dissertation is dedicated to my family, who always encouraged and never gave up on me.

I would also like to dedicate this to my good friend Mary Pettit. You were always there when I needed a shoulder to cry on and reminded me not to grieve over what was lost but you helped me to rejoice in what was to come!

## **BIOGRAPHY**

Crystal Nadine Vierhout, daughter of William and Charlotte Vierhout, was born in Iowa. She graduated from Sibley-Ocheyedan High School in Sibley, Iowa in May 1986. From 1967 to 1986 she was living and working with her family on a commercial Holstein farm in Melvin, Iowa. The author was involved with 4-H, FFA, music, and sports during this time period.

In August 1986 the author entered Iowa State University and received a Bachelor of Science degree in Dairy Science and Agriculture Business in December 1990. From January 1990 through September 1994, the author worked for Genex Cooperative. In August 1995, she enrolled at Virginia Polytechnic Institute and State University and received a Master of Science degree in Dairy Science in May 1998.

From 1998 to 2008 the author was employed by North Carolina State University and pursued a Ph.D. She received her Doctor of Philosophy in Animal Science and Poultry Science.

## ACKNOWLEDGMENTS

I would like to thank my advisors, Dr. Eisen, Dr. McCraw, and Dr. Washburn for their encouragement, support, and guidance through my degree program. I am very thankful for Dr. Washburn when I went through the year of “how many ways can I give myself a concussion?” The completion of this degree would not have been possible without Dr. Eisen, he volunteered to advise me after Dr. McDaniel retired when having another graduate student probably was the last thing on his mind. I’m grateful to Dr. McCraw for helping me through very difficult times and seeing me through to the other end.

Appreciation is expressed to Jan Wright at AIPL for sending data so quickly when I started my research over with a completely new data source. Appreciations go to Duane Norman, Paul VanRaden, and Jan Wright for the time taken out of your busy schedules to answer my questions.

I wish to thank Dr. Ingawa for all the enjoyable moments we have had together. We held down booths at ADSA and built computers for kids who could just use a break.

The road has been long and rocky but well traveled.

# TABLE OF CONTENTS

LIST OF TABLES .....	IX
LIST OF FIGURES .....	XI
LIST OF ABBREVIATIONS .....	X
GENERAL INTRODUCTION.....	1
CHAPTER 1 .....	4
REVIEW OF LITERATURE .....	4
Defining Reproduction.....	5
Traits Affecting Reproduction .....	5
Direct.....	5
Traditional .....	6
Endocrine .....	7
Correlated Traits .....	9
Rump Angle .....	10
Locomotion .....	10
Dystocia.....	11
Body Condition Score .....	13
Somatic Cell.....	15
Lactation Persistency .....	17
Production .....	18
Factors Affecting Reproduction.....	19
Maternal Lineage Effects .....	19
Genomic Imprinting.....	23
Inbreeding .....	25
Within-Herd Variance.....	26

Genetic Evaluations Associated with Female Reproduction .....	29
National Cow Reproduction Evaluation .....	30
Productive Life.....	31
Net Merit.....	32
Data Quality Issues Affecting Reproduction .....	34
Pedigree Identification .....	34
Misidentification .....	35
Missing Identification .....	37
Preferential treatment.....	38
Bull Dams .....	38
Daughters .....	39
Registration Status .....	40
Bovine Somatotropin .....	41
Conclusion .....	42
References.....	44
CHAPTER 2 .....	57
OVERVIEW OF THE USE OF DATA FROM AIPL AND DRPC IN DAIRY CATTLE RESEARCH .....	57
Introduction.....	57
Materials and Methods.....	59
Results and Discussion .....	60
Data Available .....	60
Data Quality .....	61
Retrospectively Recomputed Lactation Values .....	63
Lactation Records Calculated with Test Interval or Best Prediction Method .....	63



Adjustment Factors for Lactation Data.....	67
Previous Days Open Utilized in Standardized Record Adjustment.....	63
Consistent Standardization.....	69
Conclusions.....	69
References.....	71
CHAPTER 3 .....	75
FIELD DATA HERD EVALUATION OF THE EFFECTS OF MATERNAL LINEAGE ON PRODUCTION AND FERTILITY TRAITS OF HOLSTEIN CATTLE .....	75
Introduction.....	75
Materials and Methods.....	78
Daughter-Dam Estimates of Heritability .....	79
Management Group .....	79
Family Structure.....	80
Procedure 1 .....	81
Model .....	83
Procedure 2 .....	83
Model .....	84
Procedure 3 .....	85
Model .....	85
Results .....	86
Procedure 1 .....	89
Procedure 2 .....	90
Procedure 3 .....	92
Discussion.....	95

Conclusions.....	98
References.....	101
GENERAL CONCLUSIONS.....	119
APPENDICES .....	121
APPENDIX A.....	121
DATA SOURCES EDITS .....	122
Sources of data.....	122
Data Source 1: Lactation data from DRPC.....	122
Data Source 2: Genetic data for bulls from AIPL.....	127
Data Source 3: Genetic data for cows from AIPL .....	128
Data Source 4: Lactation data from AIPL .....	129
Combination of Data Sources .....	130
Combination 1: Source 3 x Source 4 .....	130
Combination 2: (Combination 1 + Remain 1) x Source 1 .....	131
Combination 3: Remain 2 x Source 3.....	132
Combination 4: Combination 3 x Source 2.....	134
Combination 5: Combination 4 (Dam) x Combination 4 (Daughter).....	135
Data Quality Check.....	136
References.....	138
APPENDIX B .....	142
DRMS DATA CLEANSING WITH AIPL ERROR CODES.....	142
APPENDIX C .....	148
MODEL SELECTION .....	148

References.....151

# LIST OF TABLES

CHAPTER 1 .....	4
Table 1-1 Economic values for each unit of predicted transmitting ability (PTA) and relative economic values of traits will be implemented with August 2006 evaluations .....	34
CHAPTER 2 .....	57
Table 2-1 Change in proportions of cows in Traditional versus A.M.-P.M. testing plans over time for herds processing through Dairy Records Management Systems (DRMS) .....	73
CHAPTER 3 .....	75
Table 3-1 Number of cows and daughters per generation across herds that had family for 5 generations.....	105
Table 3-2 Heritability estimates calculated using daughter-dam regression .....	106
Table 3-3 Means across families' by parities and generation for average Milk, DPR, and SCS .....	107
Table 3-4 Deviation of family average from herd merit for Milk, DPR, and SCS ..	108
Table 3-5 Means across families by generation for PTA milk, DPR, and SCS for November 2007 genetic information.....	109
Table 3-6 Means across families by generation for productive life, net merit dollars, and percent inbreeding coefficient .....	110
Table 3-7 Number of families per generation when comparison to contemporary herdmates when independent culling selection levels for milk and pregnancy rate are used .....	111
Table 3-8 Number of families per generation when comparison to contemporary herdmates when index selection for milk and pregnancy rate is used .....	112
Table 3-9 Comparison of daughters in the 5th generation using first four generations of maternal family members weighted by degree of relationship when families are grouped by independent culling levels for milk and pregnancy rate.....	113

Table 3-10 Comparison of daughters in the 5th generation using first four generations of maternal family members weighted by degree of relationship when families are grouped by index selection for milk and pregnancy rate .....	114
Table 3-11 Comparison of daughters in the 5th generation using first four generations of maternal family members weighted by degree of relationship when families are grouped by milk deviated from herd mates .....	115
Table 3-12 Comparison of daughters in the 5th generation using first four generations of maternal family members weighted by degree of relationship when families are grouped by daughter pregnancy rate deviated from herd mates.....	116
Table 3-13 Comparison of daughters in the 5th generation using first four generations of maternal family members weighted by degree of relationship when families are grouped by somatic cell score. ....	117
Table 3-14 Comparison of daughters in the 5th generation using first four generations of maternal family members weighted by degree of relationship when families are grouped by PTA productive life .....	118
APPENDIX A.....	121
Table A-1 Data Source 1: Data cleansing of lactation data from DPRC using AIPL error codes.....	124
Table A-2 Data Source 1: Data removal of data from DPRC using AIPL error codes .....	126
Table A-3 Data Source 2: Data cleansing of bull genetic evaluation for yield traits.....	127
Table A-4 Data Source 3: Data cleansing of cow genetic evaluation for yield traits.....	128
Table A-5 Data Source 4: Data cleansing of lactation data from AIPL .....	129
Table A-6 Combination 1: Record match of combining data Source 3 with 4 .....	131
Table A-7 Combination 2: Record match of combining Combination 1 and Remaining 1 with Source 1.....	132

Table A-8	Combination 3a: Record match of combining Remain 2 with Source 3 .....	133
Table A-9	Combination 3: Record counts from combining sources of data .....	134
Table A-10	Combination 4: Record match of combining Combination 3 with Source 2 .....	135
Table A-11	Combination 5: Record match of dam with daughter with Combination 4.....	136

# LIST OF FIGURES

CHAPTER 2.....	57
Figure 2-1. Differences between 305-d standardized milk values for best prediction method (BPM) <sup>1</sup> minus test-interval method (TIM) overall or for cows completing lactations longer or shorter than 305 d.....	74
APPENDIX A.....	122
Figure A-1 United States Holstein population on DHIA test, cows in the research region of study, and cows in research study for standardized milk (kg).....	139
Figure A-2 United States Holstein population on DHIA test, cows in the research region of study, and cows in research study for daughter pregnancy rating.....	140
Figure A-3 United States Holstein population on DHIA test, cows in the research region of study, and cows in research study for somatic cell score.....	141

## **LIST OF ABBREVIATIONS**

AI – artificial insemination

AIPL – Animal Improvement Program Laboratory

BPM – best prediction method

bST – bovine somatotropin

DHIA – Dairy Herd Improvement Association

DPR – daughter pregnancy rating

DRMS – Dairy Records Management Systems

DRPC – Dairy Records Processing Center

DIM – days in milk

HYS – herd-year-season

MCC – modified contemporary comparison

ML – maternal lineage

NAAB – National Association of Animal Breeders

PDO – previous days open

PTA – predicted transmitting ability

SAS – statistical analysis system

SCS – somatic cell score

TIM – test interval method

USDA – United States Department of Agriculture



## **GENERAL INTRODUCTION**

Over the years, the dairy industry in the United States has been in the process of changing from smaller family farms to larger enterprise machines. A dairy producer is striving to make the farm profitable in every aspect. We have witnessed dramatic steady increases in production per cow due partly to improved management, enhanced nutrition, and concentrated genetic selection.

The intense swing toward more productive cows and larger herds has been associated with a decrease in cow fertility. Although the heritability of reproductive traits in both males and females is very low, there is evidence of sufficient variance to indicate that selection progress is possible. Genetic progress for reproductive traits will be slow, but the dairy breeding industry needs to try to improve reproductive performance through genetic means.

The new age of agriculture is upon us where crop farms do not own livestock and most livestock producers do not grow crops. As the need to expand their enterprises has become necessary for survival, farmers have made a transition to specialization in one area of expertise. An income teeter-totter effect occurs between crop and livestock producers, as crop prices increase so does the cost to feed animals and a shortage of income for the livestock producer results. Instead of sitting in a center of a teeter-totter balancing their incomes between crop and livestock, many dairy producers find themselves on one end feeling the effects of the ups and downs of the crop market.

In an age of high fuel costs, the needs for alternative fuels have spurred the increased interest in and use of biofuels. With ethanol plants finding their place in the Corn Belt skyline, land and crop prices are on their way to record-high prices. It is too early to determine if ethanol will become a major player in fueling automobiles in the future but the economic pressure has forced dairy producers to increase herd size and production in an effort to increase efficiency and lower costs. The knowledge, and quality of that knowledge, needed by dairy producers to make efficient management decisions is enormous and it has become critical to their success.

Given these economic challenges, dairy cattle fertility has received much attention by researchers and the entire dairy industry due to its economic impact. Reproductive costs are considered to be both direct and indirect expenses to a dairy producer. Direct expenses result from increases in breeding and veterinary costs associated with increased breedings per conception. Indirect costs result from decreased milk sales associated with longer calving intervals and the inability to increase selection pressure because of greater involuntary culling.

A clearer understanding of factors influencing reproduction in dairy cattle will allow researchers a better understanding of ways to help the dairy producer become more efficient.

The objectives of this study were:

- To explain differences in data available to dairy cattle researchers from AIPL and DRPC and to provide researchers a better understanding of the data's uses

- To determine the effect maternal lineage has on production and fertility traits in Holstein cattle and to make recommendations on maternal lineage utilization in dairy cattle selection

# CHAPTER 1

## REVIEW OF LITERATURE

For many years, cow fertility data, such as calving interval and days open, have been available from Dairy Herd Improvement Associations. Due to the low heritability (~0.04) of fertility traits, they have not been evaluated routinely (VanRaden et al., 2006). Substantial increases of 3,500 kg of milk, 130 kg of fat, and 100 kg of protein per cow per lactation have resulted from improvements in genetics, nutrition, and management during the past 20 years. At the same time, the interval from calving to conception increased (unfavorable) by 24 days. In a review, Shook (2006) estimated that genetics has accounted for about 55% of gains in the yield traits and about one-third of the change in interval to conception.

Over 20 years ago, Iowa State geneticist Freeman (1986) predicted, “continued successful selection for production may depress reproduction to where selection on reproduction may be necessary” and made a challenge “...will reproductive physiologists develop new techniques to enhance reproductive performance so that selection will not be necessary?” His challenge was answered to the extent that new reproductive management tools such as estrous synchronization were developed, but genetic selection for fertility is now needed since cow fertility has continued to decline (Lucy, 2001).

## **Defining Reproduction**

Fertility in dairy cows is the “ability of the animal to conceive and maintain pregnancy if served at the appropriate time in relation to ovulation” (Darwash et al., 1997). Several factors could lead to unsuccessful pregnancy, including failure to show or detect estrus, failure to ovulate, inappropriate patterns of ovarian cyclicity and embryo or fetal loss (Royal et al., 2000b). Fertility measures calculated from calving and service dates can be divided into two groups, fertility scores and interval traits. Fertility scores include non-return to first service, which is defined by whether another service follows within a pre-determined number of days (56 or 90 days), and conception at first service, which is determined through pregnancy diagnosis or subsequent calving (Pryce et al., 2004). Interval measurements include days from calving to first service or heat, days open and calving interval (Pryce et al., 2004).

Pregnancy rate measures how rapidly a cow becomes pregnant. Pregnancy rate is defined as the percentage of cows that become pregnant during each 21-day period, since each estrus cycle represents one chance for a cow to become pregnant (VanRaden et al., 2004).

## **Traits Affecting Reproduction**

### **Direct**

Attempts to improve reproductive efficiency in dairy cattle through breeding and selection have been slowed substantially by the lack of a sufficiently heritable reproductive

measure on which to base selection for improved fertility. The traditional fertility measures of interval to first service, service per conception, days open, and calving intervals are greatly influenced by management decisions and have low heritabilities, thus limiting the possibility for meaningful genetic gain through selection for these traits (Darwash et al., 1999). Another approach is to use endocrine and physiological factors affecting reproduction as a result of gene expression at the hypothalamic, pituitary, ovarian, or uterine level (Darwash et al., 1999). Estimates for endocrine measurements are less open to bias since they are not open to confounding effects of management decisions (Royal et al., 2002).

### **Traditional**

In a United Kingdom study, Wall et al. (2005a) reported correlations of calving interval with days to first service and numbers of inseminations per conception that were strong and favorable, thus suggesting that improving one fertility trait would result in a correlated improvement in other fertility traits. The positive correlation of days to first service and calving interval was similar to other studies (Dechow et al., 2001; Veerkamp et al., 2001). Wall et al. (2005a), using the United Kingdom dairy population, further reported the average calving interval was 388 days, the average number of recorded services per conception was 1.66, and results showed that 65% of the cows did not return to service within 56 days.

In a study of United States Holstein cattle born from 1992 through 1994, heritability of days open in first lactation, calculated by calving interval, was 0.037 (VanRaden et al.,

2004). This heritability estimate was similar to the United Kingdom study for calving interval (0.033) and days open to first service (0.037). The heritability for calving interval was low, which was consistent with previous estimates (Hoekstra et al., 1994; Pryce et al., 1997; Veerkamp et al., 2001; Haile-Mariam et al., 2003; Kadarmideen et al., 2003; Muir et al., 2004). This trait can be highly influenced by the length of the voluntary waiting period. Use of synchronization products will inflate environmental variance (Muir et al., 2004).

Based on the data used to calculate the genetic evaluation of fertility reported by VanRaden et al. (2004), heritabilities for cow fertility traits in Holsteins were  $0.066 \pm 0.003$  for days to first breeding,  $0.040 \pm 0.002$  for days to last breeding,  $0.018 \pm 0.001$  for number of inseminations,  $0.010 \pm 0.001$  for 70-d nonreturn rate, and  $0.103 \pm 0.004$  for gestation length. Wall et al. (2003) estimated heritability for non-return rate after 56 days of 0.018 and number of inseminations per conception at 0.020.

VanRaden et al. (2004) concluded that these heritabilities indicate that days to first breeding is an important component of fertility (VanRaden, 2003). Even though heritabilities are quite small for fertility traits, the additive genetic variation for these traits was deemed sufficient to allow effective selection for fertility (Weller and Ron, 1992; Weigel and Rekaya, 2000).

## **Endocrine**

For a cow to conceive and maintain pregnancy, it is essential that there is synchrony among a number of physiological processes and managerial practices performed. A visible expression of estrus, timely artificial insemination by skilled technicians using semen of

fertile bulls, the shedding of an ovum capable of being fertilized and the secretion of adequate hormone levels essential for optimal tubal and uterine environments to maintain the developing embryo (Darwash et al., 1999). Failure of any of the above processes will result in low fertility.

One approach to identifying possible causes of reproductive failure is to examine them from a genetic viewpoint. Such measures will be most useful for the following reasons: 1) they are quantitative in nature since they generally provide more information than qualitative measures; 2) they are objective rather than relying on subjective judgment; and 3) they require minimum input from the producer and are buffered from the management decisions of the producer (Darwash et al., 1999).

Milk progesterone levels have been used effectively to estimate intervals to postpartum ovulation and to characterize ovarian activity as a series of luteal and inter-luteal intervals (Darwash et al., 1997). Darwash et al. (1997) reported that the commencement of postpartum luteal activity has a heritability of 0.21. Similar findings were reported by Veerkamp et al. (1998) with a heritability estimate of 0.17. With this higher heritability, the use of milk progesterone measurements has the possibility of identifying sub-fertile cows earlier or for evaluating sires based on their daughters' endocrine characteristics.



## **Correlated Traits**

In the research that is to follow, the question of how fertility traits relate to other traits will become enormously important. Heritability is the extent to which genetics influences a trait or characteristic. The higher the heritability, the more genetic control for a specific trait and the more rapidly we can achieve genetic progress. Overall, yield and all type traits tend to be moderately heritable; fat and protein percentage, stature, and size have higher heritabilities; and traits of reproductive efficiency have lower heritabilities except for progesterone levels (Schutz, 1994).

Genetic correlations indicate the extent to which two traits are influenced by the same genes (Schutz, 1994). The possible range of values for correlations is 0 to 1 with the larger the value the greater the relationship between traits. For example, the genetic correlation for milk and protein yield is high at 0.9 (Wilcox, 1992). Therefore, many of the same genes that influence milk yield also influence protein yield. Nevertheless, the genetic correlation for milk yield and fat percentage is -0.3 (Wilcox, 1992); therefore, daughters with high milk yield often have low fat percentage. One of the keys to selecting for improved fertility includes the knowledge of the genetic correlations between these production traits and fertility. Because fertility is the main emphasis of this dissertation, the following traits will be emphasized because a correlation with fertility has been established.

## **Rump Angle**

Many artificial insemination organizations train their evaluators to predict that cows with high pin bones will have poorer fertility. The basis for this argument is that high pin bones tilt the vaginal canal, therefore, causing it to lie at an angle rather than lying flat. As a result, the rump angle may affect reproduction, pregnancy, and parturition, because at an upward angle the reproductive tract is more prone to infection as the vagina is unable to drain effectively (Astiz et al., 2002). Wall et al. (2005a) showed an unfavorable genetic correlation ( $-0.16$ ) between calving interval and rump angle, suggesting animals with high pin bones would have a longer calving interval. However, no significant genetic or phenotypic correlation between days to first service and nonreturn rate and rump angle was observed. No evidence of a relationship, linear or quadratic, between any fertility trait and rump angle was found (Wall et al., 2005a).

## **Locomotion**

A cow that scored the highest for desirable feet and legs was \$213 per year more profitable, produced 575 kg more milk per year, and remained in the herd for 307 more functional days than the cows scoring the lowest for feet and legs (Perez-Cabal et al., 2006). A possible correlation between locomotion and fertility probability existed when taking into consideration 307 more days of functionality. Melendez et al. (2003) in an observational study related poor legs and feet (or lameness) with having a negative impact on fertility because the cow may be less inclined to display standing heat as a sign of estrus. Cows that are diagnosed as lame within 30 days after calving were 2.63 times more likely to develop

an ovarian cyst before first breeding compared to normal non-lame cows (Melendez et al., 2003). The overall decreased fertility in lame cows has been well-documented (Lucey et al., 1986). Pryce et al. (2000) reported a negative correlation (-0.20) of calving interval with foot angle and a positive correlation (0.19) with rear legs side view. Somewhat different outcome resulted from a study by Perez-Cabal et al. (2006) where calving interval had higher correlations with foot angle and rear leg set (0.17 and 0.12), respectively.

### **Dystocia**

Dystocia is defined as delayed and difficult calving, which results in a negative economic impact on the dairy cattle industry (Thompson et al., 1983; Djemali et al., 1987). The economic costs of dystocia include loss of calf, veterinary fees, farmer labor costs, increased risk of subsequent health and fertility problems, increased culling, and reduced production (Meijering, 1984). Dairy producers assign scores (1 being no problem to 5 being extreme difficulty) to rate calving difficulty. Dematawena and Berger (1997) estimated costs for scores 1 to 5 were \$0.00, \$50.45, \$96.48, \$159.82, and \$379.61 respectively as calving difficulty increased. Martinez et al. (1983) reported calf mortality by 48 hours after calving to be 6.65%, but for the most difficult births (score 5) 57% of all calves died. For calving with a lower score of 4, calf death loss was lower at 27%.

Dystocia is one of the most economically significant secondary traits. The heritability estimates are low to moderate at 0.03 to 0.20 (Meijering, 1984; Djemali et al., 1987; Dematawena and Berger, 1997). In the United States, only a fraction of herds participates in reporting dystocia. From the years of May 1981 to January 1984 only thirty-

four percent of all Holstein cows in the DHI program reported dystocia (Djemali et al., 1987). Studies have shown that calving difficulty affected reproductive performance and was associated with increased days open and services per pregnancy (Thompson et al., 1983; Mangurkar et al., 1984).

Djemali et al. (1987) examined dystocia records among 555,562 Holsteins from Mid-States Dairy Records Processing Center from May 1981 through January 1984. Only 34% of the 555,562 Holstein cows in the DHI program had dystocia scores reported, with frequency of reporting increasing over time. Differences between first lactation cows scoring a 5 versus 1 were 465 kg milk. Days open by parity were 14 days, 26 days, and 19 days for parities 1, 2, and 3 and greater, respectively. Days open were greater for births scoring 5 versus 1; however, only 4% of second parity and greater calvings in the data scored a 5.

Dematawena and Berger (1997) who included the data from the Djemali et al. (1987) study added an additional 122,715 yield records with dystocia information from 1981 to 1991. The difference between first lactation cows scoring a 5 versus 1 was a 683.75 kg decrease in milk production and 34 d increase in days open. Over all parities, about \$380 worth of economic loss (opportunity loss of income) can be expected by a producer when a cow has extreme difficulty calving compared to a cow having no difficulty.

## **Body Condition Score**

Fertility in dairy cattle is strongly influenced by the extent and duration of negative energy balance (Butler et al., 1981; Butler and Smith, 1989). An estimated 80% of cows experience negative energy balance in early lactation if the energy demand for milk yield is not met by the diet (Butler and Smith, 1989; Nebel and McGilliard, 1993). The relationship between nutrition and fertility is largely driven by the extent to which the diet meets energy and protein requirements. The major link between nutrition and fertility is that cows in severe negative energy balance during early lactation have lower conception rates (Pryce et al., 2004).

A cow's body condition score changes throughout the lactation of the cow corresponding to changes in her energy balance. As milk yield peaks and demand for energy exceeds intake, the cow mobilizes her lipid reserves and she loses body weight. This process is related to the daily milk yield curve, which is almost exactly opposite to the energy balance and body condition score curves (Coffey et al., 2003). The genetic correlation between milk yield and body condition score has been estimated to be -0.37 (Veerkamp and Brotherstone, 1997).

Veerkamp (1998) suggested that successive measures of body condition score might be a useful indicator of energy balance because the shortfall in energy obtained from food is believed to come from mobilization of body tissue reserves. The excessive mobilization of body reserves may have adverse effects on cow fertility and health (Coffey et al., 2003). There is evidence that cows of high genetic merit mobilize more body tissue in early lactation than cows of average genetic merit (Pryce et al., 2000).

Pryce et al. (2000) reported a negative genetic correlation of -0.22 after adjusting for phenotypic milk yield and genetic correlation between calving interval and body condition. When phenotypic milk was not adjusted, genetic correlations between body condition score and calving interval (-0.36), days to first heat (-0.41), and days to first service (0.54) were higher. Cows with low body condition score have longer calving intervals; therefore, mobilization of body tissue plays a role in the genetic control of fertility.

The genetic correlation between body condition score and fertility indicates that body condition score is a moderate to strong indicator of fertility; body condition score in mid to late lactation expressed the strongest relationship with fertility (Berry et al., 2003b). Research indicates that body condition score is not genetically the same trait in various stages of lactation (Collard et al., 2000; de Vries and Veerkamp, 2000; Banos and Brotherstone, 2004). Berry et al. (2003b) reported the heritability estimates for body condition score at different stages of lactation were smallest at day 0 of lactation (0.39) and largest at day 105 of lactation (0.51).

Berry et al. (2003b) suggest that increasing body condition score levels will reduce the interval to first service and number of services. Lopez-Gatius et al. (2003) reported that among animals showing good body condition at parturition, the number of days open was significantly lower by 5.8 and 11.7 days, respectively, compared to animals with intermediate or low body condition. As a result, body condition score can serve as a predictor for the estimated breeding value of fertility (Berry et al., 2003b). This may suggest that genes associated with body tissue mobilization may have pleiotropic effects or be closely linked to genes controlling fertility (Berry et al., 2003a). With the trend of

changing correlations of body condition score with days in milk, this would suggest that selection for higher milk yield in early lactation will have a more evident deleterious effect on pregnancy rates than selection for higher milk yield in late lactation (Berry et al., 2003b).

### **Somatic Cell**

The most frequent and costly disease affecting dairy cattle is mastitis, and economic losses can be attributed to both clinical and sub-clinical disease (Caraviello et al., 2005). Approximately 50% of the total health care costs are attributed to mammary function, and a majority of mammary health care costs is due to mastitis (Hansen et al., 1979; Shanks et al., 1982). In a Danish study, Neerhof et al. (2000) used survival analysis methodology to investigate the impact on mastitis on longevity and reported the relative risk of culling of cows with clinical mastitis was 1.69 times that of an unaffected cow. Several studies (Jasper et al., 1982; Young, 1992) have estimated the economic loss from mastitis ranging from \$100 to \$200 per cow per year. This is a projected annual cost to the United States dairy industry in excess of \$2 billion (Jasper et al., 1982). High somatic cell counts in milk are also detrimental for the dairy processor, as it reduces shelf life of dairy products (Barbano et al., 1991). Even a modest increase in an individual somatic cell count has been shown to reduce cheese yield (Barbano et al., 1991).

Associated with elevated somatic cell count (SCC) in milk, mastitis is an inflammation of the mammary gland. Because high somatic cell count in milk is a response to presence of microbes in the mammary gland, somatic cell count can be used as an

indicator of mastitis and a measure of response to infection (Heringstad et al., 2006). Direct selection for mastitis resistance has been inefficient because the heritability of clinical mastitis is low 0.03 to 0.06 (Shook, 2006). The genetic correlation between clinical mastitis and somatic cell count has been estimated to be reasonably high at 0.73 (Koivula et al., 2005), therefore indicating a strong genetic relationship between these two traits. Furthermore, somatic cell count has been used as an indirect indicator of mastitis.

The genetic relationship of somatic cell score and milk yield has a mean genetic correlation of 0.28 for first parity, -0.15 for second parity, and 0.05 for later parities (Kennedy et al., 1982; Monardes and Hayes, 1985; Schutz et al., 1990). Culling of first parity cows with severe mastitis may have contributed to lower correlation estimates for later parities (Schutz, 1994). Clinical or subclinical mastitis may also limit the potential for milk yield during later parities (Schutz, 1994).

The correlation between fertility and somatic cell was researched by Kadarmideen (2004), reported genetic correlations, of non-return rates and days to first service with lactation somatic cell score of -0.02 and 0.13, respectively indicating that a high somatic cell score may interfere with a cow's ability to cycle.

Calus et al. (2005) reported that herds with lower average somatic cell score had 7 shorter days to first service than herds with higher average somatic cell scores. Lower herd averages for somatic cell score and calving interval was associated by Calus et al. (2005) with more desirable average phenotypic values with many traits. Calus et al. (2005) agreed with the hypothesis of Rougoor et al. (1999) that herds with lower somatic cell scores and shorter average calving interval have, on average, better management, resulting in better



overall fertility and health. Calus et al. (2005) reported the relative importance of some fertility traits doubles compared with yield traits across environments. Therefore, suggesting that environment-specific breeding values should be estimated to customize selection indices.

### **Lactation Persistency**

Persistency is the ability to maintain the level of milk production after peak milk yield (Muir et al., 2004). Genetic change of the shape of the lactation curve is of great interest for the dairy cattle industry for its technical and economic implications. The rates of climb of milk yield to the lactation peak and the slope of the curve in the second part of lactation have been widely investigated in dairy cattle. Shape of the lactation curve is determined by genetic and environmental factors (Gengler, 1996). Persistency is a trait with direct economic effects because of its relationship with reproduction, health, and feed cost (Dekkers et al., 1998). Persistency of lactation has an economic value of about 3.4% of that for the total lactation yield according to Dekkers et al. (1998).

Bar-Anan et al. (1985) defined persistency as the average daily yield divided by peak yield and reported that cows with higher lactation persistency had better reproductive performance. Muir et al. (2004) studied genetic relationships between lactation persistency and reproductive performance in first parity. Canadian Holsteins with first-parity reproductive, persistency, and production information were evaluated. Reproductive performance traits included age at first insemination, nonreturn rate at 56 days after insemination, and calving interval. Estimated genetic correlations between reproduction

traits and persistency ranged from 0.17 to 0.43. This result was consistent with other studies (Beechinor and Kelly, 1987; Lean et al., 1989) in finding that cows with higher lactation persistency had better reproductive performance. The genetic correlation between calving interval and persistency was estimated to be  $0.17 \pm 0.09$ . The estimated genetic correlation between calving difficulty and persistency was moderate ( $0.43 \pm 0.08$ ); therefore, first parity production was more persistent if the heifer had difficulty calving. Heifers that had a difficult first calving tended to have more persistent first lactations and would tend to have lower peak yields, possibly causing the antagonistic relationship between calving difficulty and persistency (Muir et al., 2004).

## **Production**

Incredible advancement in milk production has occurred over a 30-year period. From 1974 to 2004, annual milk yield per cow has increased from 7,565 to 12,252 kg (AIPL. USDA., 2006a). However, the correlation between production and fertility is unfavorable in dairy cows (Pryce and Veerkamp, 2001). Pregnancy rate allows herd managers to measure how quickly their cows become pregnant again after having a calf and is defined as the percentage of nonpregnant cows that become pregnant during each 21-day period (AIPL. USDA., 2006b). With increased milk production, there has been a 6% decline in pregnancy rate since 1980 (Shook, 2006), which is equivalent to an increase of 24 days open. The rate of decline in pregnancy rate is around 0.5% per year in the United States (Beam and Butler, 1998). Genetic factors have accounted for over 55% of the

phenotypic gains in yield traits and about one-third of the decline in pregnancy rate (Shook, 2006).

Washburn et al. (2002) reported that average days open increased by more than 40 days between 1982 and 1999 in Holsteins in a study of herds with continuous reporting from 1976 to 1999 in 10 southeastern states. Conception rates decreased in this study from around 50% in the early 1980s to close to 34% in 1999.

Because of the absence of direct selection pressure on fertility, there has been a downward genetic trend in fertility associated with selection for yield (Royal et al., 2000a). Management changes can partially address poor on-farm fertility but require continuous input and, as a result, increased cost. Ignoring the genetic component of poor fertility, has masked the severity of the problem and has led to continuing downward genetic trend for fertility. Failing to modify selection practices to include the improvement of fertility will lead to a continuing downward genetic trend.

## **Factors Affecting Reproduction**

### **Maternal Lineage Effects**

For many years, there has been a belief among many dairy cattle breeders that certain cow families produced daughters with higher genetic merit than sons. It has been the practice of the dairy sire analyst to research the potential bull-dam pedigrees for not only a high genetic evaluation for type but also a maternal lineage with several successive generations of outstanding females. Suggesting that some form of maternal effect exists, research (Van Vleck and Hart, 1966; Thompson and Loganathan, 1968; Seykora and

McDaniel, 1983) has consistently reported higher heritability for milk production traits from daughter-dam regression than from paternal half-sib analysis. Gibson et al. (1997) suggested that a possible cause could be correlated environmental effects within lineages; cows born into outstanding lineages receive preferential treatment in proportion to the perceived quality of the family.

Spielman and Jones (1939) reported reproductive efficiencies of groups of dairy cows descended from different foundation cows. The study included individual breeding records of 368 cows in four major dairy breeds from the Oregon State dairy herd over a 24-year period. Each foundation cow and her female descendents composed a cow group, with the groups including four to eleven generations containing 11 to 62 cows with breeding records. This study reported a noticeable difference in reproductive efficiency among cow groups. A correlation of 0.55 was reported between the reproductive efficiency of the foundation cows and the mean reproductive efficiency of their female descendents. This study suggested that reproductive efficiency of the foundation cow was an important factor in determining mean reproductive efficiency of her offspring.

Maternal genetic effects are defined as the genetic effects of the dam in addition to her direct additive genetic contribution (Southwood et al., 1989). Even though maternal effects may arise through several pathways, only two paths are frequently examined in maternal lineage research, additive effects and cytoplasmic effects. Additive maternal effects are inherited in a Mendelian manner. Maternal effects can be defined as any influence from a dam on its offspring, excluding the effects of directly transmitted genes that affect performance of the offspring (Legates, 1972).

Biological mechanisms to explain maternal effects include cytoplasmic inheritance, intrauterine and postpartum nutrition provided by the dam, antibodies and pathogens transmitted from dam to offspring, and maternal behavior (Hohenboken, 1985).

Cytoplasmic genes are transferred directly from dam to offspring. With no mutation or heterogeneity of the cytoplasmic components within a dam, all offspring of a dam will receive an identical copy of cytoplasmic genes (Southwood et al., 1989).

The early studies of maternal effects centered mostly on mothering ability. Because dairy dams do not nurse their young for an extended period, little attention has been given to maternal effects in dairy cattle. Even though additive maternal effects do not appear appreciably to influence milk production (Van Vleck and Hart, 1966), other maternal effects, such as cytoplasmic inheritance, possibly could contribute to difference in performance (Bell et al., 1985).

Bell et al. (1985) analyzed 4,461 first lactation Holstein cows for production and reproductive traits. The research traced back 25 generations between 1949 and 1980 in experimental and control herds in North Carolina. Cytoplasmic source effects accounted for 2.0, 1.8, and 3.5% of the variation for milk yield, fat yield, and fat percentage, respectively. Reproductive performance results indicated a possible maternal or cytoplasmic effect. Using data from herds included in the Bell et al. (1985) study, Seykora and McDaniel (1983) obtained higher heritability for days open from daughter-dam regression (0.13) than from paternal half-sisters (0.05). This result indicates possible cytoplasmic effects on days open. Higher production results were reported in study of 290 cows from a Dutch experimental station in explaining cytoplasmic components.

Cytoplasmic effects explained 6, 10, and 13 percent of the variation in milk yield, combined fat, and protein yields (Huizinga et al., 1986). Huizinga et al. (1986) found that 8 to 10% of the variation in reproductive performance was controlled cytoplasmically.

In associating cytoplasmic inheritance with maternal lineages, Schutz et al. (1994) estimated maternal lineage ranges of 2934 kg for milk, 154 kg for fat yield, and 0.907% for fat percentage in records from the Iowa State University herd. With data pooled from Iowa State University and North Carolina State University herds, Boettcher et al. (1996) estimated contributions of maternal lineage to total variance of 0.38%, 0.71%, and 2.90% for milk yield, fat yield, and fat percent, respectively.

The impact of cytoplasmic effects on all aspects of genetic evaluations has been addressed. In a simulation study, ignoring cytoplasmic effects caused overestimations of additive genetic variance Southwood et al. (1989). With the reason that upward biased estimates of heritability would overestimate selection differential, ignoring cytoplasmic effects would cause inaccurate genetic evaluations.

Researchers disagree on whether selection on cytoplasmic effects in genetic evaluations appears to be feasible. Gibson et al. (1997) state that because selection intensities in this pathway are very low, dealing with cytoplasmic effects are not an important issue now. Not accounting for cytoplasmic effects will cause only a small bias in the estimation of cow genetic merit because estimates of additive genetic effects of cows will be biased toward the cytoplasmic effect when cytoplasmic effect is in the model. In the current availability of pedigree information and current attitude, we may not be able to

estimate maternal lineage variance, but we cannot discard its effects (Roughsedge et al., 2001).

## **Genomic Imprinting**

Mendelian genetics has the fundamental assumption that behavior of an allele is identical whether it arrives through the paternal or maternal germline pathway (Ruvinsky, 1999). The gametic imprinting phenomenon discovered in mammals showed limitations to the classical view of Mendelian genetics in special cases. Genomic imprinting is defined as the process by which the expression of certain genes depends on the sex of the parent from which they are inherited (Monk, 1988; Reik, 1989; Sapienza, 1990). During mammalian development, parental allele-specific expression patterns are thought to result from the acquisition at certain genetic loci of different epigenetic modification (imprints) in the paternal germlines (Moore and Reik, 1996). About 20 to 25 of all transgene loci studied demonstrate similarities with imprinted genes (Peterson and Sapienza, 1993). Methylation of some of these transgenes is dependent of paternal gametic pathway and reversible in the next generation (Ruvinsky, 1999).

Consequences of gametic imprinting include 1) a deviation from the Mendelian form of inheritance; 2) although genetically identical, the input of maternal and paternal alleles in development is quite different in some cases; and 3) the expression pattern of maternal and paternal alleles may cause significant changes in ontogenesis (Ruvinsky, 1999).

Growth and development traits are affected by gametic imprinting (Ruvinsky, 1999). In cattle, growth abnormalities resulting from the manipulation of preimplantation embryos in vitro (large calf syndrome) could reflect shifts of the balance between maternal and paternal contributions (Moore and Reik, 1996). Gametic imprinting is a mammalian innovation and there are differences in imprinting patterns between species (Ruvinsky, 1999). Proposed involvement of imprinted genes in the control of fetal growth and fetal-maternal interactions, therefore, provides a balance between conflicting fetal and maternal requirements (Ruvinsky, 1999).

In Australian Friesian heifers, Kaiser et al. (1998) determined paternal imprinting accounted for 1% of the variance in adjusted milk yield, while maternal imprinting accounted for 0.5%. Heritabilities were estimated without gametic imprinting in the model and with sire or maternal imprinting were 0.283, 0.262, and 0.276 respectively. Neither additive nor residual variances were greatly affected by including gametic imprinting in the model. Therefore, the results suggest that gametic imprinting does not cause significant variation in milk production (Kaiser et al., 1998).

Essl and Voith (2002) investigated the effect of imprinting on milk yield, fat and protein content, persistency, days open and herd life in the Austrian Simmental population. For a majority of the investigated traits, no substantial imprinting effects were detected. Significant evidence of genomic imprinting was found for protein content and days open.



## **Inbreeding**

Inbreeding results from the mating of closely related individuals and results in reduced heterozygosity (Falconer, 1989). One consequence of inbreeding is inbreeding depression, which by definition is the reduction of the mean phenotypic value, particularly for traits connected to reproduction and fitness that are affected by dominance (Falconer, 1989).

The use of reproductive technology has allowed an extensive increase in the selection intensity through worldwide use of semen and the practice of increasing the reproductive rate of elite females have increased the potential for inbreeding. Thompson et al. (2000a) reported that inbreeding depression has been shown to decrease milk production by approximately 9 to 26 kg of milk per lactation for each 1% increase of inbreeding.

The estimated effect of inbreeding on nonproductive traits include decreased reproductive efficiency (Hermas et al., 1987; Smith et al., 1998), decreased longevity (Huizinga et al., 1986; Smith et al., 1998), and a slight increase in somatic cell score (Miglior et al., 1995). Other research has documented somatic cell score not to be affected by inbreeding (Thompson et al., 2000a). A decrease of 13.1 days in length of productive life for each 1% increase in inbreeding has been reported by Smith et al. (1998). Inbreeding levels greater than 10% decreased lactation lengths by 2 to 8 days with survival being reduced for all levels of inbreeding (Thompson et al., 2000a). A decrease of 3.7 days per 1% increase in inbreeding in age at first calving has been reported (Hermas et al., 1987).

The effects of inbreeding have been shown to be more severe as the level of inbreeding increases (Wall et al., 2005b). The average inbreeding coefficient for the United States Holstein population is approximately 5.1% with a current average annual increase of 0.2% (AIPL. USDA., 2006a). At this current rate of inbreeding, more animals will soon be in a higher inbreeding class, where the effect of inbreeding will become more distinct.

### **Within-Herd Variance**

In genetic evaluations, within herd variance has a notable effect on sire evaluations and a potentially more serious effect on cow evaluations. Initial studies of within-herd variance focused on the association of herd variances with herd averages. Most studies, with two exceptions (Legates, 1962; Lofgren et al., 1985), reported that environmental variances increased with herd averages (Burnside and Rennie, 1961; McDaniel and Corley, 1967; Hill et al., 1983; Powell et al., 1983; De Veer and Van Vleck, 1987; Boldman and Freeman, 1990; Dong and Mao, 1990).

For three independent analyses, Hill et al. (1983) obtained records on first lactation daughters of British Friesian sires. Herds were first separated on mean production into high and low levels of average production. The groups of low and high milk yield herds had heritabilities of 0.24 and 0.30, respectively. The data were also split into two groups by within-herd variance, and again by within year-season coefficient of variation using variances and means computed from the previous splits. An increase in milk yield per herd from low to high in yield both within and between sires were reported. Higher heritability

was found in herds with larger coefficients of variation. Therefore, daughters were evaluated more accurately in herds with higher mean or variance for production. Hill et al. (1983) concluded that for British herds with higher production levels there was a corresponding increase in heritability for milk yield, fat, and protein. Powell and Norman (1983) obtained similar results with United States data from 1979 herd-average milk yields. They reported that heritability estimated for groups at 5 herd average levels ranged from 0.26 in the lowest to 0.35 for the highest producing group of herds.

Two studies (Legates, 1962; Lofgren et al., 1985) did not report increases in heritability as production levels increased. In 1962, Legates (1962) used Guernsey, Jersey, and Holstein daughter-dam pairs to estimate heritability for fat yield and its relationship with herd-average levels. Increases in phenotypic and genetic variance were present as milk yield increased but the coefficient of variation declined. Consequently, no significant relationship was reported between herd production levels and heritability estimates. The author reported no need to use different heritability estimates for herds at different production levels.

Lofgren et al. (1985) did not find an increase in heritability as production levels increased, using data from Holstein and Jersey cows born after 1963. Three groups of herd means and variances were formed using modified contemporary average and heritability was estimated in each group. The lowest heritability was the middle group with higher estimates in the extreme groups. Lofgren et al. (1985) concluded a possible reason for the contrast was that older data were used in the previous study as well as a different method of analysis.

Within-herd variance has potentially serious effects on cow evaluations if variation remains constant in the same herd over years (Brotherstone and Hill, 1986). Approximately half of the bias in a dam's index is reflected in the index of her daughter. Furthermore, with most daughter-dam pairs in the same herd, within-herd variation has a cumulative effect over several generations.

Powell et al. (1983) used the modified contemporary comparison to illustrate those herds with higher milk production per cow tended to use bulls with higher predicted differences and to have higher cow indexes. A larger proportion of cows in these herds will attain elite status for production than in herds with lower production. A cow's elite status should reflect genetic superiority and not environmental opportunity. The work concluded that genetically superior cows with higher yields seem to result from herds with higher heritability. Higher heritability would also be associated with larger within herd variance. McDaniel and Corley (1967) reported that daughters in herds with higher milk production also had increased variation among their evaluations.

Hill (1984) concluded that using the same intensity of selection among members of two groups of equal size with the same mean but different variances would cause three-quarters of the selected animals to come from the more variable group. Everett et al. (1982) using 7,398 Holstein herds with equal genetic merit reported 10% of the cows qualified for elite status in herds with larger variance and none qualified in the lower variance herds. Large within-herd variation causes more cows to reach elite status.

The improvement in evaluation procedures has accomplished changes in production technology and breeding structure through test-day models. Environmental effects such as

age and lactation stage have been reported to be major sources of variation among herds (Kachman and Everett, 1989; Stanton et al., 1992). The improved accounting for environmental effects, an advantage of test-day models over genetic evaluations based on lactation records, is accounting for differences in the shape of the lactation curve (Bormann et al., 2003). Everett et al. (1994) investigated the percentage of total variation removed in a model considering age and stage within the herd. In the herd effects model, 60% of the total daily yield variation was removed.

## **Genetic Evaluations Associated with Female Reproduction**

A persistent increase in the complexity of genetic evaluations has significantly influenced the history of dairy cattle breeding. Information used to evaluate animals has been expanded from the physical appearance of the cow, to records of yields, to the further inclusion of correlated traits and information on relatives (Henderson, 1973).

Henderson (1949) initiated the theory required for genetic evaluation using an animal model. His theory referred to an animal model as a method of cow and sire evaluation in which the performance of an animal is used to estimate the breeding value of that individual (Cassell, 1988). Only recently, because of computational advances, has the impact of Henderson's theory been realized.

Fertility in dairy cattle breeding may be considered as having two components: 1) a measure of time, such as onset of estrus, and 2) a measure of the ability of the cow to conceive at first or second insemination (Pryce et al., 2000).

## National Cow Reproduction Evaluation

Cow fertility is negatively correlated with yield but is a major component of how long a cow stays in the herd. Because of the importance of cow fertility, a national fertility evaluation was implemented in 2004 (VanRaden et al., 2004). The national fertility evaluation was developed based on pregnancy rate, which measures the percentage of nonpregnant cows becoming pregnant within each 21-day opportunity period. Pregnancy rate calculations were used in this evaluation instead of days open because increasing positive values of pregnancy rate are easier to follow by producer than an inverse relationship with fewer days open being more desirable (VanRaden et al., 2004). For each increase of 1% in predicted transmitting ability pregnancy rate, there is a decrease of 4 d in predicted transmitting ability days open (VanRaden et al., 2004) with a 0.99 genetic correlation between days open and pregnancy rate.

$$\text{Pregnancy rate} = 21 / (\text{days open} - \text{voluntary waiting period} + 11)$$

VanRaden (2003)

Furthermore, the voluntary waiting period is the initial phase of lactation during which no insemination occurs. The voluntary waiting period is selected by the producer and may vary by herd, state and season.

The above formula for pregnancy rate could be analyzed using separate binary variables within a 21-day cycle. A cow may require more than one cycle to become pregnant, which would result in multiple observations per lactation. VanRaden (2003) reported a simpler analysis of a single fertility record per lactation, which is used for routine

evaluations. The simpler analysis of a single fertility record was used due to higher heritability per observation over the above formula. The author has also selected this simpler formula in calculating pregnancy rate.

$$\text{Pregnancy rate} = 0.25 \times (233 - \text{days open}) \text{ VanRaden (2003)}$$

National records were adjusted for region, year of calving, season of calving effects and herd variance (VanRaden, 2003). Only records having 250 days in milk or greater were included for evaluation of pregnancy rate. The 250 days was used to avoid a potential bias that could be created by using only cows that were bred early in lactation (Kuhn et al., 2004). Since daughter pregnancy rating is particularly important for early progeny test results, a longer waiting period of 250 days was required. In recent work, Kuhn et al. (2004) investigated use of early lactation days open records on genetic evaluations of cow fertility and reported days open can be adequately predicted for single-trait genetic evaluation by 130 days in milk using a predictor, fit within days in milk group.

## **Productive Life**

Cows leave herds for multiple reasons, including low yield, mastitis, reproductive failure, sales for dairy purposes, and death (Dentine et al., 1987). The addition of a genetic evaluation for productive life in 1995 was the first genetic evaluation that was directly associated with longevity. Productive life was defined as the total months in milk through 85 months of age. Productive life had a limit of 305 days in milk per lactation (VanRaden and Klaaskate, 1993). VanRaden and Wiggans in 1995 reported that a single-trait

evaluation for productive life was more valuable in prediction of herd life than the collective predicted transmitting ability of 14 type traits. However, they did indicate that type traits could add to the accuracy of productive life evaluations when both sources are combined in a multi-trait evaluation.

Since no revisions were proposed for productive life calculations for 10 years, in August 2006 an economic definition was introduced to replace the 1995 version. The 2006 revision was important because more than 50% of the cows had lactations longer than 305 days and weight was given beyond the 305-day period. With the 2006 revision, weights were based on standard lactation curves, with the highest weight at the peak of lactation and diminishing credits across the remainder of the lactation (VanRaden, 2006). Also included in the 2006 version, are slightly more weighting for later lactations, therefore ensuring that cows with multiple lactations get more total weight than cows with a single lengthy lactation. Changes were also made in productive life to consider increased costs associated with raising replacement heifers. Heritability for productive life improved from 0.017 to 0.070 with the addition of censoring being increased from 36 to 96 months (VanRaden et al., 2006).

## **Net Merit**

“A key priority in research and education should be to identify those traits that really affect cost of producing milk and concentrate selection on them.” Since McDaniel



(1976) made that statement in 1976, it took almost 30 years for net merit to become the selection index closely fitting his goal.

In 1994, VanRaden (2004) introduced “net merit dollar,” combining productive life, somatic cell score, and yield traits. These traits were weighted 10:4:-1 for yield, productive life, and somatic cell score, respectively.

In August 2000, VanRaden (2006) revised net merit dollar to include linear conformation composites using a lifetime profit function. By including linear conformation in 2000, net merit dollar became an early predictor of direct selection rather than productive life. Three years later in 2003, net merit went through another revision to include service sire calving ease, daughter calving ease, and daughter pregnancy rating into the calculation (VanRaden, 2004). The 2003 net merit dollar index had a correlation of 0.98 with the 2000 net merit dollar index. Because more traits were directly affecting profit, accuracy of selection has increased over the 2000 net merit index. Net merit dollar in 2003 resulted in an expected 2% increase in genetic progress. This 2% increase has an estimated worth of \$5 million per year nationally (VanRaden, 2004).

VanRaden (2006) again revised net merit in 2006 to include a newly revised definition of productive life and new genetic evaluation for service sire and daughter stillbirth. A calving ability index, which combines the correlated traits of calving ease and stillbirth, was included in net merit dollar index. The 2006 net merit dollar index has a correlation of 0.975 with the 2003 version. Half the changes in the 2006 version on net merit dollar are caused by revision of productive life and the rest from addition of stillbirth

and changes in trait economic values. An estimated rate of genetic progress worth \$6 million per year is expected on a national basis from the 2003 version of net merit dollar.

Table 1-1 Economic values for each unit of predicted transmitting ability (PTA) and relative economic values of traits will be implemented with August 2006 evaluations<sup>1</sup>.

PTA Traits	Correlation		Expected Genetic Progress	
	with PTA Index		from Net Merit \$	
	2003	2006	PTA	Breed Value
	NM\$	NM\$	Change/Year	Change/Decade
Protein	.74	.62	2.6	52
Fat	.67	.66	3.8	76
Milk	.58	.54	86	1720
Productive Life	.58	.67	.30	6.0
Somatic Cell Score	-.38	-.37	-.017	-.34
Udder	.22	.17	.04	.80
Feet/legs	.16	.13	.03	.60
Body size	-.10	-.17	-.04	-.80
Daughter Pregnancy Rate	.15	.27	.07	1.4
Calving Ability	.23	.34	1.3	25

<sup>1</sup><http://aipl.arsusda.gov/reference/nmcalc.htm>

## Data Quality Issues Affecting Reproduction

### Pedigree Identification

With introduction of the “animal model” for national genetic evaluation of dairy cattle over a decade ago, correct pedigree information is vital in the successful breeding program. A simple pedigree error in the animal model will affect the predicted transmitting ability of that cow and all of her relatives. Sanders et al. (2006) concluded that the combined impact of misidentification and missing sire information on genetic gain was

relatively large, especially for sires having low numbers of progeny and traits of low heritability. The impact of misidentification on the efficiency is about 1.4 times more harmful than the impact of missing identification.

### **Misidentification**

With the reported rate of paternity misidentification in various cattle populations around the world approaching 25%, its impact and possible solutions are being investigated. Misidentification rates by country are estimated as follows: around 5% in Israel (Ron et al., 1996), around 10% in the United Kingdom (Visscher et al., 2002), around 12% in the Netherlands (Bovenhuis and van Arendonk, 1991), 5 to 15% in Denmark (Christensen et al., 1982), 8 to 20% in Ireland (Beechinor and Kelly, 1987), and 4 to 23% in Germany (Geldermann et al., 1986).

Van Vleck (1970a) raised questions about valid identification in a study of differences in heritability estimates from parent offspring compared to paternal half-sib methods. He found lower heritability estimates when using paternal half sib methods. Heritability was estimated to be 64% of its true value when only 80% of cows were identified correctly by sire. Therefore, misidentification of daughters leads to substantial underestimation of heritability (Van Vleck, 1970b).

Geldermann et al. (1986) evaluated the effect of misidentification on genetic gain on the sire-to-daughter path of inheritance for milk and fat yields. Genetic gains with an incorrect paternity identification rate of 15% were 8.7% lower for heritability of 0.5 and 16.9% lower for heritability of 0.2, as compared with completely correct paternity

determination. In a later stochastic simulation study, Israel and Weller (2000) investigated the effect of a 10% incorrect paternity on genetic progress over a period of 20 year in a large dairy cattle population undergoing selection. They estimated that the annual rate of response was reduced by 3 to 4% due to misidentification. This study reflects both the impact of incorrect identification of both sexes and the accumulated effect of pedigree errors over multiple generations.

Misidentification is expected to reduce the differences between the estimated breeding values. Cows misrepresented as daughters of an exceptionally high bull would be over evaluated, as their true sire would likely be of lower genetic merit. Misidentification reduces apparent genetic variation. The impact of misidentification on genetic gain is twofold. First, the reliability of the bull proofs are reduced because of a lower correlation between the bull proof and the true breeding value, which causes a lower genetic gain than can be achieved when there are no pedigree errors. Second, the expected genetic gain is lower than that achieved because of the downward bias in the heritability estimate (Visscher et al., 2002).

Christensen et al. (1982) gave the following reasons for errors in paternity recording: 1) mistakes by AI institutes in labeling semen; 2) AI technicians incorrectly identifying semen samples; 3) insemination of cows already pregnant by a previous insemination; 4) clerical errors when the bull's herdbook number or name is entered into the insemination record; 5) use of natural-service bulls leading to pregnancies of previously inseminated cows which were assumed to be pregnant from the AI bull; 6) mistakes in sire identification when a cow enters the milking herd in schemes where pedigree information

on milk recorded cows is obtained through the milk recording program; and 7) interchange of calves at the farm.

Weller et al. (2004) genotyped 104 microsatellites in the progeny of 11 sires and reported an 11.7% frequency of rejected paternity determination. Of the 7 causes listed by Christensen et al. (1982) for paternity mistakes, it appeared that inseminator recording mistakes, and possibly mistakes with respect to semen labeling at the artificial insemination institutes, are most significant. Incorrect paternity recording due to multiple inseminations by different sires could explain at most 20% of the paternity mistakes. Implementing a system of quality control, at the level of the inseminator, should reduce paternity errors to no more than 8%, and increase genetic progress by at least 1%.

### **Missing Identification**

The second source of pedigree errors is missing information. Until recently there has been little information on the extent and impact of this source of error. Harder et al. (2005) noted that missing identification influenced substantially the variance of estimated sire breeding values and reduced the response to selection.

Missing identification is the primary reason why lactation records are excluded from genetic evaluations. Genetic evaluations during 1984 were computed with 48% of records from approved test plans being eliminated (Wiggans and Waite, 1985). Slightly over 25% of records from usable plans did not include sire identification. During January 1992, 35% of records were eliminated; 31% of records submitted for genetic evaluations were excluded because of missing sire identification (Norman et al., 1994).

Meinert et al. (1997) reported only a small increase in usable official records from 1968 to 1978 but a steeper linear increase from 1979 to 1988. This increase in correct identifications could be a result of better heifer record systems and improved education on the values of identification. From 1988 to 1990, usability of official records changed little, possibly because of larger herd sizes and a high proportion of grade cows. Sire identification was less complete in the grade population (Ehlers et al., 1975).

### **Preferential treatment**

Preferential treatment is any management practice that increases production and is applied to one or several cows, but not to their equally worthy contemporaries (Kuhn et al., 1994). Preferential treatment may involve providing better or more feed, different housing, or even breeding a cow one more time where another cow would not receive that one last chance.

### **Bull Dams**

Financial gain and name recognition of having a bull enter AI create a temptation for giving a potential bull mother preferential treatment. A dairy producer may not intentionally give an individual cow preferential treatment or may not realize favoritism at all. Unfortunately, intentional preferential treatment may also occur so that a cow will be chosen as a bull dam and, thus, sell at a higher price.

Overestimation of genetic merit of bull dams contributes to overestimation of parent average. A cow's inflated evaluation can remain until information from her progeny,

particularly her sons, causes her evaluation to reflect her actual genetic merit more accurately (Powell et al., 1993).

Kuhn et al. (1994), using simulated data, reported potential for substantial bias from 15 kg to 893 kg in female predicted transmitting abilities from cows that received preferential treatment. Weigel et al. (1994) reported similar results with simulated data.

Accurate first-crop daughter information is a necessity for genetic progress when the 10 to 15% of young sires that graduate to the proven lineup will receive very heavy use. Overestimation of a bull's genetic merit will result in heavier use than true genetic merit justifies. Underestimation of a bull may result in the bull's being culled from the AI organization with his true genetic merit left unknown. Both cases are of great interest to dairy producers and AI organization since they reduce potential genetic progress and can erode producers' confidence.

### **Daughters**

Preferential treatment of bull dams is one way that bias in predicting transmitting ability may occur. Another source of bias in a bull's proof is preferential treatment of his daughters. In a simulation study, Kuhn and Freeman (1995) reported no more than 5 to 6% of a bull's daughters, when distributed across herds, can receive preferential treatment without causing the bull's evaluation to be inflated by 5%. If a sire's dam received preferential treatment equivalent to increased milk yield less than 2,268 kg, then 20 to 30 daughters without preferential treatment were necessary to obtain an approximately

unbiased evaluation of the bull. Bias in a sire's predicted transmitting ability when daughters receive preferential treatment can be substantial.

### **Registration Status**

Animals recorded by the Holstein Association with both parent lines tracing back to original importation into North America from Europe are considered registered Holsteins. A grade individual has a less complete record of ancestry. The percentage of usable records for genetic evaluations from registered cows declined from 64% in 1969 to 42% in 1990 with the balance of records contributed by grade cows (Meinert et al., 1997).

Powell and Norman (1986) studied preferential treatment according to registry status using cows with calving dates from 1960 to 1984 that were enrolled in official DHIA test programs and that passed edits for use in MCC genetic evaluations. Consequently, sire identification was present on all individuals but missing dam identification did not exclude individuals. Only about one-third of grade cows on official test were identified by sire. Within the same herd and year, average milk yields of registered cows generally were higher than grade cows but as cows aged, production of grade cattle surpassed that of registered cows (Dentine et al., 1987). Grade cows had a higher culling rate than registered cows. More sire-identified grades than registered 2-year-old cows are included in the data but by the age of 5 year, only 8% more grades than registered cows were still milking. Dentine et al. (1987) using July 1983 summary information on a random sample of cows that calved for the first time between 1965 and 1980, reported registered 2-year-old cows out-produced grades in the same herd by 99 kg milk. However, increased intensity of



culling among grade cows resulted in registered cows out-producing grades by only 21 kg over all lactations.

Herds with more grade cows tended to have a higher average sire predicted difference, but within a typical herd, registered cows were sired by bulls with higher predicted difference (Powell and Norman, 1986). Powell and Norman (1986) reported that registered cows received preferential treatment within herds with a mix of registered and grade cows. Approximately 19% of herds had at least one registered and one grade cow (Dentine et al., 1987). Preferential treatment of registered over grade cattle was about 90 kg milk per lactation, which resulted in about 3 kg in cow index across herds (Powell and Norman, 1986).

### **Bovine Somatotropin**

Preferential treatment can be defined in general as a hidden environmental effect that cannot be accounted for in the model of analysis (Swalve, 1991). Recombinant bovine somatotropin consistently increases milk yield by 10 to 15% per cow (Dentine et al., 1987). Supplementation of bovine somatotropin not only causes increases in milk production, but also increases persistency of production. Therefore, calving interval for optimal economic returns may be substantially increased when bovine somatotropin is administered (Bauman, 1992).

Santos et al. (2004) measured conception to first postpartum artificial insemination and reported that both bST treatment and artificial insemination protocol affect fertility. Cows inseminated at detected estrus had higher conception rates than those inseminated at

fixed-time following the Ovsynch protocol, and bST improved conception rates in cyclic cows. Insemination protocol had no impact on pregnancy loss, but treatment with bST tended to reduce pregnancy loss in all cows and improved pregnancy maintenance in cyclic dairy cows.

Given the reported increase in milk yield and fertility and the possibility that its use may be limited to certain cows, bST has become an instrument of preferential treatment. Research has been limited on preferential treatment of bST due to the limited reporting of usage by producers. More research is needed in this area.

## **Conclusion**

Over the years, the dairy industry in the United States has been in the process of changing from smaller family farms to larger and more business-oriented ones. Dairy producers are striving to make their farms profitable in every aspect. We have witnessed dramatic steady increases in production per cow due partly to improved management, enhanced nutrition, and concentrated genetic selection.

The intense swing toward more productive cows and larger herds has been associated with decreases in cow fertility. Although heritability of reproductive traits in both males and females is very low, there is evidence of sufficient variation to indicate that selection progress is possible. Genetic progress for reproductive traits will be slow, but the dairy breeding industry needs to try to improve reproductive performance through genetic means. The proposed research should assist in determining if a difference in fertility of

cow families currently exists and whether those differences may be incorporated into an improved system for improving dairy cow fertility.

## References

- AIPL. USDA.** 2006a. AIPL inbreeding coefficients for Holstein cows (Calculated February, 2006) Online: <http://aipl.arsusda.gov/dynamic/inbrd/current/Hot.html>. Accessed August 14, 2006.
- AIPL. USDA.** 2006b. Daughter pregnancy rate evaluation of cow fertility. Online: [http://aipl.arsusda.gov/reference/fertility/DPR\\_rpt.htm](http://aipl.arsusda.gov/reference/fertility/DPR_rpt.htm). Accessed November 7, 2006.
- Astiz, B. S., M. J. V. Gonzalez, G. L. Ayala, and V. A. Monge.** 2002. The influence of the pelvic conformation on the incidence of urovagina – an epidemiological study. Proc. XXII World Buiatrics Congress, Hanover, Germany. Page 362-365.
- Banos, G. and S. Brotherstone.** 2004. Evaluation of body condition score measured throughout lactation as an indicator of fertility in dairy cattle. *J. Dairy Sci.* 87:2669-2676.
- Bar-Anan, R., M. Ron, and G. R. Wiggans.** 1985. Associations among milk yield, yield persistency, conception, and culling of Israeli Holstein dairy cattle. *J. Dairy Sci.* 68:381-386.
- Barbano, D. M., R. R. Rasmussen, and J. M. Lynch.** 1991. Influence of milk somatic cell count and milk age on cheese yield. *J. Dairy Sci.* 74:369-388.
- Bauman, D. E.** 1992. Bovine somatotropin: review of an emerging animal technology. *J. Dairy Sci.* 75:3432-3451.
- Beam, S. W. and W. R. Butler.** 1998. Energy balance, metabolic hormones, and early postpartum follicular development in dairy cows fed prilled lipid. *J. Dairy Sci.* 81:121-131.
- Beechinor, J. G. and E. P. Kelly.** 1987. Errors of identification amongst cattle presented as progeny of some bulls used in the artificial insemination service in Ireland. *Ir. Vet. J.* 41:348-353.
- Bell, B. R., B. T. McDaniel, and O. W. Robinson.** 1985. Effects of cytoplasmic inheritance on production traits of dairy cattle. *J. Dairy Sci.* 68:2038-2051.
- Berry, D. P., F. Buckley, P. Dillon, R. D. Evans, M. Rath, and R. F. Veekamp.** 2003a. Genetic relationships among body condition score, body weight, milk yield, and fertility in dairy cows. *J. Dairy Sci.* 86:2193-2204.

- Berry, D. P., F. Buckley, P. Dillon, R. D. Evans, M. Rath, and R. F. Veerkamp.** 2003b. Genetic parameters for body condition score, body weight, milk yield, and fertility estimated using random regression models. *J. Dairy Sci.* 86:3704-3717.
- Boettcher, P. J., D. W. B. Steverink, D. C. Beitz, A. E. Freeman, and B. T. McDaniel.** 1996. Multiple herd evaluation of the effects of maternal lineage on yield traits of Holstein cattle. *J. Dairy Sci.* 79:655-662.
- Boldman, K. G. and A. E. Freeman.** 1990. Adjustment of heterogeneity of variances by herd production level in dairy cow and sire evaluation. *J. Dairy Sci.* 73:503-512.
- Bormann, J., G. R. Wiggans, T. Druet, and N. Gengler.** 2003. Within-herd effects of age at test day and lactation stage on test-day yields. *J. Dairy Sci.* 86:3765-3774.
- Bovenhuis, H. and J. A. M. Van Arendonk.** 1991. Estimation of milk protein gene frequencies in crossbred cattle by maximum likelihood. *J. Dairy Sci.* 74:2728-2736.
- Brotherstone, S. and W. G. Hill.** 1986. Heterogeneity of variance amongst herds for milk production. *Anim. Prod.* 42:297-303.
- Burnside, E. B. and J. C. Rennie.** 1961. The heritability of milk yield at different levels of production and the effect of production differences on dairy sire appraisals. *J. Dairy Sci.* 44(Suppl.1):2669-2676.
- Butler, W. R., R. W. Everett, and C. E. Coppock.** 1981. The relationships between energy balance, milk production and ovulation in postpartum Holstein cows. *J. Anim. Sci.* 53:742-748.
- Butler, W. R. and R. D. Smith.** 1989. Interrelationships between energy balance and postpartum reproductive function in dairy cattle. *J. Dairy Sci.* 72:767-783.
- Calus, M. P. L., J. J. Windig, and R. F. Veerkamp.** 2005. Associations Among Descriptors of Herd Management and Phenotypic and Genetic Levels of Health and Fertility. *J. Dairy Sci.* 88:2178-2189.
- Caraviello, D. Z., P. L. Ruegg, G. E. Shook, and K. A. Weigel.** 2005. Assessment of the impact of somatic cell count on functional longevity in Holstein and Jersey cattle using survival analysis methodology. *J. Dairy Sci.* 88:804-811.
- Cassell, B. G.** 1988. What extension workers need to tell dairy farmers. *J. Dairy Sci.* 71(Suppl. 2):85-90(Abstr.).

- Christensen, L. G., P. Madsen, and J. Petersen.** 1982. The influence of incorrect sire identification on the estimates of genetic parameters and breeding values. 2nd World Congress on Genetics applied to Livestock Production, 4th-8th October 1982. 7. Symposia (1). 200.
- Coffey, M. P., G. Simm, W. G. Hill, and S. Brotherstone.** 2003. Genetic evaluations of dairy bulls for daughter energy balance profiles using linear type scores and body condition score analyzed using random regression. *J. Dairy Sci.* 86:2205-2212.
- Collard, B. L., P. J. Boettcher, J. C. M. Dekkers, D. Petitclerc, and L. R. Schaeffer.** 2000. Relationships between energy balance and health traits of dairy cattle in early lactation. *J. Dairy Sci.* 83:2683-2690.
- Darwash, A. O., G. E. Lamming, and J. A. Woolliams.** 1999. The potential for identifying heritable endocrine parameters associated with fertility in post-partum dairy cows. *Anim. Sci.* 68:333-347.
- Darwash, A. O., G. E. Lamming, and J. A. Woolliams.** 1997. Estimation of genetic variation in the interval from calving to postpartum ovulation of dairy cows. *J. Dairy Sci.* 80:1227-1234.
- De Veer, J. C. and L. D. Van Vleck.** 1987. Genetic parameters for first lactation milk yields at three levels of herd production. *J. Dairy Sci.* 70:1434-1441.
- de Vries, M. J. and R. F. Veerkamp.** 2000. Energy Balance of Dairy Cattle in Relation to Milk Production Variables and Fertility. *J. Dairy Sci.* 83:62-69.
- Dechow, C. D., G. W. Rogers, and J. S. Clay.** 2001. Heritabilities and correlations among body condition scores, production traits, and reproductive performance. *J. Dairy Sci.* 84:266-274.
- Dekkers, J. C. M., J. H. TenHag, and A. Weersink.** 1998. Economic aspects of persistency of lactation in dairy cattle. *Livest. Prod. Sci.* 53:237-252.
- Dematawena, C. M. B. and P. J. Berger.** 1997. Effect of dystocia on yield, fertility, and cow losses and an economic evaluation of dystocia scores for Holsteins. *J. Dairy Sci.* 80:754-761.
- Dentine, M. R., B. T. McDaniel, and H. D. Norman.** 1987. Comparison of culling rates, reasons for disposal, and yields for registered and grade Holstein cattle. *J. Dairy Sci.* 70:2616-2622.

- Djemali, M., A. E. Freeman, and P. J. Berger.** 1987. Reporting of dystocia scores and effects of dystocia on production, days open, and days dry from dairy herd improvement data. *J. Dairy Sci.* 70:2127-2131.
- Dong, M. C. and I. L. Mao.** 1990. Heterogeneity of (co)variance and heritability in different levels of intraherd milk production variance and of herd average. *J. Dairy Sci.* 73:843-851.
- Ehlers, M. H., B. F. Kelso, and R. W. Anderson.** 1975. Sire identification as related to size of holstein herds and proportion of grade cows. *J. Dairy Sci.* 28:754(Abstr.).
- Essl, A. and K. Voith.** 2002. Genomic imprinting effects on dairy and fitness-related traits in cattle. *J. Anim. Breed. Genet.* 119:182-189.
- Everett, R. W., J. F. Kewon, and J. F. Taylor.** 1982. The problem of heterogeneous within herd variances when identifying elite cows. *J. Dairy Sci.* 65(Suppl. 1):100(Abstr.).
- Everett, R. W., F. Schmitz, and L. H. Wadell.** 1994. A test-day model for monitoring management and genetics in dairy cattle. *J. Dairy Sci.* 77(Suppl. 1):267(Abstr.).
- Falconer, D. S.** 1989. *Introductions to Quantitative Genetics.* 3rd rev. ed. ed. John Wiley & Sons, Inc, New York, NY.
- Freeman, A. E.** 1986. Genetic control of reproduction and lactation in dairy cattle. 3rd World Congress on Genetics applied to Livestock Production, Lincoln, Nebraska, USA, July 16-22, 1986. XI. Genetics of reproduction, lactation, growth, adaptation, disease, and parasite resistance. 3.
- Geldermann, H., U. Pieper, and W. E. Weber.** 1986. Effect of misidentification on the estimation of breeding value and heritability in cattle. *J. Anim. Sci.* 63:1759-1768.
- Gengler, N.** 1996. Persistency of lactation yields: A review. *Interbull Bull.* 12:87-96.
- Gibson, J. P., A. E. Freeman, and P. J. Boettcher.** 1997. Cytoplasmic and mitochondrial inheritance of economic traits in cattle. *Livest. Prod. Sci.* 47:115-124.
- Haile-Mariam, M., P. J. Bowman, and M. E. Goddard.** 2003. Genetic and environmental relationship among calving interval, survival, persistency of milk yield and somatic cell count in dairy cattle. *Livest. Prod. Sci.* 80:189-200.

- Hansen, L. B., C. W. Young, K. P. Miller, and R. W. Touchberry.** 1979. Health care requirements of dairy cattle. I. Response to milk yield selection. *J. Dairy Sci.* 62:1922-1931.
- Harder, B., J. Bennewitz, N. Reinsch, M. Mayer, and E. Kalm.** 2005. Effect of missing sire information on genetic evaluation. *Arch. Tierz.* 48:219-232.
- Henderson, C. R.** 1973. Sire evaluation and genetic trends. Page 10 in Proceedings of the Animal Breeding and Genetics Symposium in honor of Dr. Jay L. Lush, held July 29, 1972, at Virginia Polytechnic Institute and State University, Blacksburg, Virginia, .
- Henderson, C. R.** 1949. Estimation of changes in herd environment. *J. Dairy Sci.* 32:706-715.
- Heringstad, B., J. Ødegård, G. Klemetsdal, D. Gianola, and Y. M. Chang.** 2006. Genetic associations between clinical mastitis and somatic cell score in early first-lactation cows. *J. Dairy Sci.* 89:2236-2244.
- Hermas, S. A., C. W. Young, and J. W. Rust.** 1987. Effects of mild inbreeding on productive and reproductive performance of Guernsey cattle. *J. Dairy Sci.* 70:712-715.
- Hill, W. G.** 1984. On selection among groups with heterogeneous variance. *Anim. Prod.* 39:473-477.
- Hill, W. G., M. R. Edwards, M. K. A. Ahmed, and R. Thompson.** 1983. Heritability of milk yield and composition at different levels and variability of production. *Anim. Prod.* 36:59-68.
- Hoekstra, J., A. W. van der Lugt, J. H. J. van der Werf, and W. Ouweltjes.** 1994. Genetic and phenotypic parameters for milk production and fertility traits in upgraded dairy cattle. *Livest. Prod. Sci.* 40:225-232.
- Hohenboken, W. D.** 1985. Maternal Effects, Page 135-150 in *General and Quantitative Genetics*. A.B Chapman. Elsevier Science Pub. Co., Amsterdam ; New York.
- Huizinga, H. A., S. Korver, B. T. McDaniel, and R. D. Politiek.** 1986. Maternal effects due to cytoplasmic inheritance in dairy cattle. Influence on milk production and reproduction traits. *Livest. Prod. Sci.* 15:11-26.
- Israel, C. and J. I. Weller.** 2000. Effect of misidentification on genetic gain and estimation of breeding value in dairy cattle populations. *J. Dairy Sci.* 83:181-187.



- Jasper, D. E., J. S. McDonald, R. D. Mocherie, W. N. Philpo, R. J. Famswolth, and S. B. Spencer.** 1982. Bovine mastitis research needs, funding, and sources of suppon. Page 184 in Roc. 21st Annu. Mtg. Natl. Mastitis Counc., Lexington, KY. atl. Mastitis Counc., Arlington, VA.
- Kachman, S. D. and R. W. Everett.** 1989. Test day data model with individual herd correction factors. *J. Dairy Sci.* 72(Suppl. 1):60(Abstr.).
- Kadarmideen, H. N.** 2004. Genetic correlations among body condition score, somatic cell score, milk production, fertility and conformation traits in dairy cows. *Anim. Sci.* 79:191-201.
- Kadarmideen, H. N., R. Thompson, M. P. Coffey, and M. A. Kossaibati.** 2003. Genetic parameters and evaluations from single- and multiple-trait analysis of dairy cow fertility and milk production. *Livest. Prod. Sci.* 81:183-195.
- Kaiser, C. J., M. E. Goddard, and A. Reverter.** 1998. Analysis of gametic imprinting effects for test day milk yield in Australian Holstein cattle. in Proc. 6th World Congr. Genet. Appl. Livestock Prod. 23:355-358.
- Kennedy, B. W., M. S. Sethar, J. E. Moxley, and B. R. Downey.** 1982. Heritability of somatic cell count and its relationship with milk yield and composition in Holsteins. *J. Dairy Sci.* 65:843-847.
- Koivula, M., T. Serenius, E. Negussie, and E. A. Mäntysaari.** 2005. Genetic and phenotypic relationships among milk yield and somatic cell count before and after clinical mastitis. *J. Dairy Sci.* 88:827-833.
- Kuhn, M. T., P. J. Boettcher, and A. E. Freeman.** 1994. Potential biases in predicted transmitting abilities of females from preferential treatment. *J. Dairy Sci.* 77:2428-2437.
- Kuhn, M. T. and A. E. Freeman.** 1995. Biases in predicted transmitting abilities of sires when daughters receive preferential treatment. *J. Dairy Sci.* 78:2067-2072.
- Kuhn, M. T., P. M. VanRaden, and J. L. Hutchison.** 2004. Use of early lactation days open records for genetic evaluation of cow fertility. *J. Dairy Sci.* 87:2277-2284.
- Lean, I. J., J. C. Galland, and J. L. Scott.** 1989. Relationships between fertility, peak milk yields and lactational persistency in dairy cows. *Theriogenology.* 31:1093-1103.
- Legates, J. E.** 1962. Heritability of fat yields in herds with different production levels. *J. Dairy Sci.* 45:990-996.

- Legates, J. E.** 1972. The role of maternal effects in animal breeding. IV. Maternal effects in laboratory species. *J. Anim. Sci.* 35:1294-1302.
- Lofgren, D. L., W. E. Vinson, R. E. Pearson, and R. L. Powell.** 1985. Heritability of milk yield at different herd means and variance for production. *J. Dairy Sci.* 68:2737-2739.
- Lopez-Gatius, F., J. Yaniz, and D. Madriles-Helm.** 2003. Effects of body condition score and score change on the reproductive performance of dairy cows: a meta-analysis. *Theriogenology.* 59:801-812.
- Lucey, S., G. J. Rowlands, and A. M. Russell.** 1986. The association between lameness and fertility in dairy cows. *Vet. Rec.* 118:628-631.
- Lucy, M. C.** 2001. Reproductive loss in high-producing dairy cattle: where will it end? *J. Dairy Sci.* 84:1277-1293.
- Mangurkar, B. R., J. F. Hayes, and J. E. Moxley.** 1984. Effects of calving ease-calf survival on production and reproduction in Holsteins. *J. Dairy Sci.* 67:1496-1509.
- Martinez, M. L., A. E. Freeman, and P. J. Berger.** 1983. Genetic relationship between calf livability and calving difficulty of Holsteins. *J. Dairy Sci.* 66:1494-1502.
- McDaniel, B. T.** 1976. Selection goals for dairy cattle. *Proc. Natl. Workshop Genet. Improv. Dairy Cattle.* St Louis, MO.
- McDaniel, B. T. and E. L. Corley.** 1967. Relationships between sire evaluations at different herdmate levels. *J. Dairy Sci.* 50:735-741.
- Meijering, A.** 1984. Dystocia and stillbirth in cattle -- A review of causes, relations and implications. *Livest. Prod. Sci.* 11:143-177.
- Meinert, T. R., H. D. Norman, J. M. Mattison, and C. G. Sattler.** 1997. Usability for genetic evaluations of records from herds participating in progeny test programs of artificial insemination organizations. *J. Dairy Sci.* 80:2599-2605.
- Melendez, P., J. Bartolome, L. F. Archbald, and A. Donovan.** 2003. The association between lameness, ovarian cysts and fertility in lactating dairy cows. *Theriogenology.* 59:927-937.
- Miglior, F., E. B. Burnside and J. C. M. Dekkers.** 1995. Nonadditive genetic effects and inbreeding depression for somatic cell counts of Holstein cattle. *J. Dairy Sci.* 78:1168-1173.

- Monardes, H. G. and J. F. Hayes.** 1985. Genetic and phenotypic relationships between lactation cell counts and milk yield and composition of Holstein cows. *J. Dairy Sci.* 68:1250-1256.
- Monk, M.** 1988. Genomic imprinting. *Genes Dev.* 2:921-925.
- Moore, T. and W. Reik.** 1996. Genetic conflict in early development: parental imprinting in normal and abnormal growth. *Rev. Reprod.* 1:73-77.
- Muir, B. L., L. R. Schaeffer, and J. Fatehi.** 2004. Genetic relationships between persistency and reproductive performance in first-lactation Canadian Holsteins. *J. Dairy Sci.* 87:3029-3037.
- Nebel, R. L. and M. L. McGilliard.** 1993. Interactions of High Milk Yield and Reproductive Performance in Dairy Cows. *J. Dairy Sci.* 76:3257-3268.
- Neerhof, H. J., P. Madsen, V. P. Ducrocq, A. R. Vollema, J. Jensen, and I. R. Korsgaard.** 2000. Relationships between mastitis and functional longevity in danish black and white dairy cattle estimated using survival analysis. *J. Dairy Sci.* 83:1064-1071.
- Norman, H. D., L. G. Waite, G. R. Wiggans, and L. M. Walton.** 1994. Improving accuracy of the United States genetics database with a new editing system for dairy records. *J. Dairy Sci.* 77:3198-3208.
- Perez-Cabal, M. A., C. Garcia, O. Gonzalez-Recio, and R. Alenda.** 2006. Genetic and phenotypic relationships among locomotion type traits, profit, production, longevity, and fertility in Spanish dairy cows. *J. Dairy Sci.* 89:1776-1783.
- Peterson, K. and C. Sapienza.** 1993. Imprinting the genome: imprinted genes, imprinting genes, and a hypothesis for their interaction. *Annu. Rev. Genet.* 27:7-31.
- Powell, R. L. and H. D. Norman.** 1986. Genetic and environmental differences between registered and grade Holstein cows. *J. Dairy Sci.* 69:2897-2907.
- Powell, R. L., H. D. Norman, and B. T. Weinland.** 1983. Cow evaluation at different milk yields of herds. *J. Dairy Sci.* 66:148-154.
- Powell, R. L. and H. D. Norman.** 1983. Heritabilities of milk and fat yield according to herd-average yield. *J. Dairy Sci.* 66(Suppl. 1):123(Abstr.).
- Powell, R. L., G. R. Wiggans, and H. D. Norman.** 1993. Utilization of genetically elite Holstein cows. *J. Dairy Sci.* 76:1412-1420.

- Pryce, J. E., M. P. Coffey, and S. Brotherstone.** 2000. The genetic relationship between calving interval, body condition score and linear type and management traits in registered Holsteins. *J. Dairy Sci.* 83:2664-2671.
- Pryce, J. E., M. D. Royal, P. C. Garnsworthy, and I. L. Mao.** 2004. Fertility in the high-producing dairy cow. *Livest. Prod. Sci.* 86:125-135.
- Pryce, J. E., R. F. Veerkamp, R. Thompson, W. G. Hill, and G. Simm.** 1997. Genetic aspects of common health disorders and measures of fertility in Holstein Friesian dairy cattle. *Anim. Sci.* 65:353-360.
- Pryce, J. E. and R. F. Veerkamp.** 2001. The incorporation of fertility indices in genetic improvement programmes. Pages 223–236 in *Fertility in the High-Producing Dairy Cow*. M. Diskin, ed. British Society of Animal Science Occasional Publication No. 26., Edinburgh, Scotland.
- Reik, W.** 1989. Genomic imprinting and genetic disorders in man. *Trends Genet.* 5:331-336.
- Ron, M., Y. Blanc, M. Band, E. Ezra, and J. I. Weller.** 1996. Misidentification rate in the Israeli dairy cattle population and its implications for genetic improvement. *J. Dairy Sci.* 79:676-681.
- Roughsedge, T., S. Brotherstone, and P. M. Visscher.** 2001. Bias and power in the estimation of a maternal family variance component in the presence of incomplete and incorrect pedigrees information. *J. Dairy Sci.* 84:944-950.
- Rougoor, C. W., W. J. A. Hanekamp, A. A. Dijkhuizen, M. Nielen, and J. B. M. Wilmink.** 1999. Relationships between dairy cow mastitis and fertility management and farm performance. *Prev. Vet. Med.* 39:247-264.
- Royal, M. D., A. P. F. Flint, and J. A. Woolliams.** 2002. Genetic and phenotypic relationships among endocrine and traditional fertility traits and production traits in Holstein-Friesian dairy cows. *J. Dairy Sci.* 85:958-967.
- Royal, M. D., A. O. Darwash, A. P. F. Flint, R. Webb, J. A. Woolliams, and G. E. Lamming.** 2000a. Declining fertility in dairy cattle: changes in traditional and endocrine parameters of fertility. *Anim. Sci.* 70:487-501.
- Royal, M. D., G. E. Mann and A. P. F. Flint.** 2000b. Strategies for reversing the trend towards subfertility in dairy cattle. *Vet. J.* 160:53-60.

- Ruvinsky, A.** 1999. Basics of gametic imprinting. *J. Anim. Sci.* 77(Suppl. 2):228-237(Abstr.).
- Sanders, K., E. Kalm, and J. Bennewitz.** 2006. Wrong and missing sire information affects genetic gain in the Angeln dairy cattle population. *J. Dairy Sci.* 89:315-321.
- Santos, J. E. P., S. O. Juchem, R. L. A. Cerri, K. N. Galvao, R. C. Chebel, W. W. Thatcher, C. S. Dei, and C. R. Bilby.** 2004. Effect of bST and reproductive management on reproductive performance of Holstein dairy cows. *J. Dairy Sci.* 87:868-881.
- Sapienza, C.** 1990. Parental imprinting of genes. *Sci. Am.* 263:52-60.
- Schutz, M. M.** 1994. Genetic evaluation of somatic cell scores for United States dairy cattle. *J. Dairy Sci.* 77:2113-2129.
- Schutz, M. M., A. E. Freeman, G. L. Lindberg, C. M. Koehler, and D. C. Beitz.** 1994. The effect of mitochondrial DNA on milk production and health of dairy cattle. *Livest. Prod. Sci.* 37:283-295.
- Schutz, M. M., L. B. Hansen, G. R. Steuernagel, J. K. Reneau, and A. L. Kuck.** 1990. Genetic parameters for somatic cells, protein, and fat in milk of Holsteins. *J. Dairy Sci.* 73:494-502.
- Seykora, A. J. and B. T. McDaniel.** 1983. Heritabilities and correlations of lactation yields and fertility for Holsteins. *J. Dairy Sci.* 66:1486-1493.
- Shanks, R. D., P. J. Berger, A. E. Freeman, D. H. Kelley, and F. N. Dickinson.** 1982. Projecting health cost from research herds. *J. Dairy Sci.* 65:644-652.
- Shook, G. E.** 2006. Major Advances in Determining Appropriate Selection Goals. *J. Dairy Sci.* 89:1349-1361.
- Smith, L. A., B. G. Cassell, and R. E. Pearson.** 1998. The effects of inbreeding on the lifetime performance of dairy cattle. *J. Dairy Sci.* 81:2729-2737.
- Southwood, O. I., B. W. Kennedy, K. Meyer, and J. P. Gibson.** 1989. Estimation of additive maternal and cytoplasmic genetic variances in animal models. *J. Dairy Sci.* 72:3006-3012.
- Spielman, A. and L. R. Jones.** 1939. The reproductive efficiency of dairy cattle. *J. Dairy Sci.* 22:329-334.

- Stanton, T. L., L. R. Jones, R. W. Everett, and S. D. Kachman.** 1992. Estimating milk, fat, and protein lactation curves with a test day model. *J. Dairy Sci.* 75:1691-1700.
- Swalve, H. H.** 1991. Detection of bovine somatotropin treatment in dairy cattle performance records. *J. Dairy Sci.* 74:1690-1699.
- Thompson, J. R., R. W. Everett, and N. L. Hammerschmidt.** 2000a. Effects of inbreeding on production and survival in Holsteins. *J. Dairy Sci.* 83:1856-1864.
- Thompson, J. R., R. W. Everett, and C. W. Wolfe.** 2000b. Effects of inbreeding on production and survival in Jerseys. *J. Dairy Sci.* 83:2131-2138.
- Thompson, J. R., E. J. Pollak, and C. L. Pelissier.** 1983. Interrelationships of parturition problems, production of subsequent lactation, reproduction, and age at first calving Dairy cows. *J. Dairy Sci.* 66:1119-1127.
- Thompson, N. R. and S. Loganathan.** 1968. Composition of cow's milk. II. Genetic influences. *J. Dairy Sci.* 51:1933-1935.
- Van Vleck, L. D.** 1970a. Misidentification and sire evaluation. *J. Dairy Sci.* 1697-1702.
- Van Vleck, L. D.** 1970b. Misidentification in estimating the paternal sib correlation. *J. Dairy Sci.* 1469-1474.
- Van Vleck, L. D. and C. L. Hart.** 1966. Covariances among first-lactation milk records of cousins. *J. Dairy Sci.* 49:41-44.
- VanRaden, P. M.** 2006. Net merit as a measure of lifetime profit: 2006 version. AIPL Res. Rep. NMS1 (11-000). Online: <http://aipl.arsusda.gov/reference/nmcalc.htm>. Accessed September 4, 2006.
- VanRaden, P. M.** 2004. Invited review: selection on net merit to improve lifetime profit. *J. Dairy Sci.* 87:3125-3131.
- VanRaden, P. M.** 2003. Longevity and fertility trait definitions compared in theory and simulation. *Interbull Bull.* 30:43-46.
- VanRaden, P. M. and E. J. H. Klaaskate.** 1993. Genetic evaluation of length of productive life including predicted longevity of live cows. *J. Dairy Sci.* 76:2758-2764.
- VanRaden, P. M., A. H. Sanders, M. E. Tooker, R. H. Miller, H. D. Norman, M. T. Kuhn, and G. R. Wiggans.** 2004. Development of a national genetic evaluation for cow fertility. *J. Dairy Sci.* 87:2285-2292.

- VanRaden, P. M., M. E. Tooker, R. E. Pearson, and C. M. B. Dematawewa.** 2006. Productive life including all lactations and longer lactations with diminishing credits. *J. Dairy Sci.* 89:3213-3220.
- VanRaden, P. M. and G. R. Wiggans.** 1995. Productive life evaluations: calculation, accuracy, and economic value. *J. Dairy Sci.* 78:631-638.
- Veerkamp, R. F.** 1998. Selection for economic efficiency of dairy cattle using information on live weight and feed intake: A Review. *J. Dairy Sci.* 81:1109-1119.
- Veerkamp, R. F. and S. Brotherstone.** 1997. Genetic correlations between linear type traits, food intake, live weight and condition score in Holstein Friesian dairy cattle. *Anim. Sci.* 64:385-392.
- Veerkamp, R. F., E. P. C. Koenen, and G. De Jong.** 2001. Genetic correlations among body condition score, yield, and fertility in first-parity cows estimated by random regression models. *J. Dairy Sci.* 84:2327-2335.
- Veerkamp, R. F., J. K. Oldenbroek, and T. Van der Lende.** 1998. The use of milk progesterone measurements for genetic improvement of fertility traits in dairy cattle; Fertility and Reproduction, GRUB, Germany, November Proceedings of the International Workshop on Genetic Improvement of Functional Traits 23-25, 1997. *Interbull.* 18:62-65.
- Visscher, P. M., J. A. Woolliams, D. Smith, and J. L. Williams.** 2002. Estimation of pedigree errors in the UK dairy population using microsatellite markers and the impact on selection. *J. Dairy Sci.* 85:2368-2375.
- Wall, E., S. Brotherstone, M. P. Coffey, and I. M. S. White.** 2005a. The relationship between fertility, rump angle, and selected type information in Holstein-Friesian cows. *J. Dairy Sci.* 88:1521-1528.
- Wall, E., S. Brotherstone, J. F. Kearney, J. A. Woolliams, and M. P. Coffey.** 2005b. Impact of nonadditive genetic effects in the estimation of breeding values for fertility and correlated traits. *J. Dairy Sci.* 88:376-385.
- Wall, E., S. Brotherstone, J. A. Woolliams, G. Banos, and M. P. Coffey.** 2003. Genetic evaluation of fertility using direct and correlated traits. *J. Dairy Sci.* 86:4093-4102.
- Washburn, S. P., W. J. Silvia, C. H. Brown, B. T. McDaniel, and A. J. McAllister.** 2002. Trends in reproductive performance in southeastern Holstein and Jersey DHI herds. *J. Dairy Sci.* 85:244-251.

- Weigel, D. J., R. E. Pearson, and I. Hoeschele.** 1994. Impact of different strategies and amounts of preferential treatment on various methods of bull-dam selection. *J. Dairy Sci.* 77:3163-3173.
- Weigel, K. A. and R. Rekaya.** 2000. Genetic parameters for reproductive traits of Holstein cattle in California and Minnesota. *J. Dairy Sci.* 83:1072-1080.
- Weller, J. I., E. Feldmesser, M. Golik, I. Tager-Cohen, R. Domochofsky, O. Alus, E. Ezra, and M. Ron.** 2004. Factors affecting incorrect paternity assignment in the Israeli Holstein population. *J. Dairy Sci.* 87:2627-2640.
- Weller, J. I. and M. Ron.** 1992. Genetic Analysis of Fertility Traits in Israeli Holsteins by Linear and Threshold Models. *J. Dairy Sci.* 75:2541-2548.
- Wiggans, G. R. and L. G. Waite.** 1985. Editing lactation records for USDA-DHIA genetic evaluations. Natl. Coop. DHI Prog. Handbook, Fact Sheet H-9. Washington, DC.
- Wilcox, C. J.** 1992. Genetics: basic concepts. Page 1 in Large Dairy Herd Management. H. H. Van Horn and C. J. Wilcox, ed. Am. Dairy Sci. Assoc., Champaign, IL.
- Young, C. W.** 1992. Breeding dairy cattle for disease resistance. Page 42 in Large Dairy Herd Management. H. H. Van Horn and C. J. Wilcox, ed. Am. Dairy Sci. Assoc., Champaign, IL.



## CHAPTER 2

# OVERVIEW OF THE USE OF DATA FROM AIPL AND DRPC IN DAIRY CATTLE RESEARCH

### Introduction

For researchers working in the area of dairy genetics in the United States, obtaining accurate data to accomplish a specific goal is essential. Possible sources of large quantities of data include artificial insemination (**AI**) organizations, breed associations, the Animal Improvement Programs Laboratory of the USDA (**AIPL**), and Dairy Records Processing Centers (**DRPC**). The goal of this section is to provide a researcher with a better understanding of “Where do I go for information and what will I receive?”

Regional organizations, referred to as Dairy Herd Improvement Associations (**DHIA**), provide milk recording, component determination, and management reports for member dairy producers (Murrill, 1985). Traditionally, field staff members, referred to as technicians with a regional DHIA, visit producers on a monthly basis to collect data and milk samples. Producers have many options in providing data including collecting the milk samples themselves, referred to as “owner-sampler.” Another option is to have a technician present once a month for all milkings (traditional) or for only one milking (a.m. or p.m.) each month. Milk samples are sent to a testing laboratory for determination of fat and protein percentages as well as somatic cell count or score. Milk weight and component information data are combined into a database at one of four DRPCs in the United States.

Reports are then provided to producers from their chosen DRPC (Wiggans, 1994). Upon combining producer information, each of the four DRPCs transmits data received from farms to AIPL in standardized computer format. The national dairy genetics database resides at AIPL.

AIPL exchanges data through standardized computer formats among all four DRPCs, breed associations, and the National Association of Animal Breeders (NAAB). NAAB is composed of A.I. organizations, each of which samples a group of bulls and transmits information on their bulls to AIPL. Breed associations have responsibility for maintaining the national herdbook for their respective breed as well as for conducting visual appraisal programs for linear type evaluations (Wiggans, 1994). These organizations also maintain information on genetic defects.

AIPL receives lactation records with pedigree information included by the various processing centers (Norman et al., 1994). Pedigree data are also provided to AIPL by breed associations and in the past by the National DHIA from Verified Identification Program which has been discontinued (Voelker, 1981). Nearly all pedigree information received by AIPL for non-registered animals is provided from the DRPC through a cow's lactation records (Majeskie, 1992).

AIPL provides industry cooperators (DRPC, breed associations, and A.I. organizations) with on-line access to accurate pedigree and lactation information (Norman et al., 1994). AIPL calculates genetic evaluations three times a year to evaluate production traits to improve efficiency of dairy production in the United States. These genetic

evaluations are made available to the entire dairy industry to help make better management and genetic selection decisions.

## **Materials and Methods**

Data for this study were obtained from the largest of the four DRPC processing centers in the United States. Data included historical lactation records dating from January 1980 to August 2005. Records were from herds in 13 states processed by Dairy Records Management Systems (**DRMS**) in Raleigh, NC and included Alabama, Georgia, Florida, Indiana, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, Virginia, and Vermont.

From those DRPC data, 7,840,594 lactation records on 2,773,097 cows were sent to AIPL to be cross-matched with standardized lactation yields. When multiple records per cow per lactation were observed, the records with the lower number of days in milk were deleted (1.18%). The AIPL dataset returned included 3,420,771 lactation records on 1,440,577 cows. Genetic information was obtained from AIPL for cow data on the November 2007 genetic evaluation, which included the first herd of the cow's first lactation and the last herd in which she completed her terminating record. To eliminate effects on records of possible transfer issues between processing centers, 118,395 cows with 225,387 (7.21%) corresponding lactations without matching first and last herd identification were eliminated from the study. Because DRMS does not produce a standardized value for cows with less than 40 days in milk, 115 additional records with less than 40 days and a

standardized milk value were removed. Individual lactation records for Holsteins that appeared in both AIPL and DRMS sources were used in this study.

## **Results and Discussion**

### **Data Available**

Researchers who wish to use data from DRMS are required to submit a letter with research objectives to DRMS to obtain authorization to use the data for research purposes. DRMS data made available for research include a current and an historical individual lactation database. The lactation database includes a cow's complete lactation history with one record per lactation. For many variables in the DRMS dataset, information is included only on the last lactation of a cow, making comparison of cows by lactation very difficult or impossible. DRMS also has a standard herd summary database available to researchers. For the current study, DRMS data included information on herds from 1980 through 2005.

Due to the pooling of data from multiple dairy industry cooperators, AIPL has the ability to provide researchers a broader amount of information on a specific cow than any one dairy industry cooperator (Norman et al., 1994). AIPL has standard formats available for pedigree, production, type appraisal, calving ease, health, and reproduction information making up the cow's complete lactation history. AIPL also has retrospectively recomputed genetic evaluations on available cow and bull information.

For research where detection of broad trends is required, a database scanning multiple regions and individual farms are required (Wiggans, 1994). Data from a specific

DRPC can be used when researching a regional concern, whereas AIPL's multi-region data will more accurately detect broader trends in the entire United States.

For research that requires generational or historical information on a herd or cow level, additional understanding of data source is required. Data from AIPL include complete historical information on a herd or cow basis. When using DRPC data, knowledge of procedures used by the processing center for specific states may be required. Over the years, consolidation has occurred such that there are now only four DRPCs in the U.S. The reduction in processing centers required herds and their respective cow data to be transferred to active DRPCs. For a complete herd transfer, National DHIA policy requires that a standard transfer file be transmitted from one processing center to another for each active cow. National DHIA recommends, but does not require, transfer of records on animals that have left the herd in the last 14 months (National Dairy Herd Information Program, 2000). Due to lack of requirements for DRPCs to transfer ancestral and historical information, when using DRPC data in research, the length of time the herd or state has been processed at the DRPC will determine the amount of historical information received.

## **Data Quality**

Validity and appropriateness of interpretations of research results are directly reliant on accuracy of the database and the appropriateness of the data editing system utilized (Norman et al., 1994). DRPC performs initial quality edits on data received from producers before forwarding data to AIPL. After leaving the DRPC, extremely sophisticated editing

for data quality occurs at AIPL (Norman et al., 1994). A cross-referencing system occurs at AIPL for animals enrolled in several identification programs. For Holsteins, research identification data could be present at Holstein-Friesian of Canada, Holstein Association of America, and NAAB, as well as ear tag numbers from the USDA uniform tag series. Multiple identifications for the same animal are detected by AIPL through verification from other pedigrees with the same sire, dam, and birth date (Norman et al., 1994). Due to cross-referencing edits occurring at AIPL, matching historical DRPC with retrospectively re-computed AIPL data can be extremely challenging.

When data are received at AIPL, they are checked immediately for inconsistencies. Inconsistencies found in pedigree, lactation, reproduction, and health status are brought to the attention of the DRPC submitting the record. If a discrepancy in a record occurs between DRPC and AIPL, DRPC is notified of the conflict (Norman et al., 1994). At Dairy Records Management Systems (DRMS), the DRPC that provided records in the current study, the producer is notified if a discrepancy occurs on cow information within a producer's herd to ensure that dairy producers have the opportunity to correct the problem before genetic evaluations are calculated (J. S. Clay, DRMS, Raleigh, NC, personal communication). Producers choose whether to amend data conflicts, thereby allowing producers some control on accuracy of their herd data at DRMS. If a change occurs at AIPL on an animal not currently active in the producer's herd, notification is sent from AIPL, but the producer is not notified nor does the correction occur at DRMS.

Just as DRPCs have a responsibility to notify the producer if a discrepancy occurs on cow information, in order to assist in maintaining highly accurate data, a researcher

should also consider this same responsibility to notify the data provider when discovering quality issues within research data. Problems can only be corrected if they are detected; thus, a meticulous and knowledgeable researcher is an asset in finding data errors (Wiggans, 1994).

## **Lactation Records Calculated with Test Interval or Best Prediction Method**

Beginning in 1969, test-interval method (TIM) was used to estimate lactation records (Sargent et al., 1968). Yield traits were estimated by simple linear interpolation between measured yields (VanRaden, 1997). Shook et al. (1980) developed a set of factors that improved estimation of daily yields before first test, after last test, and at peak yield. This set of factors was named Shook factors or projection factors.

Since implementation of TIM, many new test plans have been introduced, which can vary greatly in test intervals compared to traditional monthly tests. A reduction in cost of obtaining lactation records has occurred due to these new test plans considerations of incomplete data on test day, reduced supervision, and electronic records. With this reduction in cost of lactation records, the resulting accuracy has been lowered in some cases but increased in others (VanRaden, 1997). Because of those changes, VanRaden (1997) developed a procedure to estimate accuracy of records for current test plans by considering number and distribution of tests. The best prediction method (BPM) provides a prediction of yield for any day of lactation or a lactation total for any length through 305 days. Correlation between 305-d actual milk yield and BPM was 0.97 (Norman et al., 1999).

When comparing test plans (monthly, a.m.-p.m., tri-monthly) with TIM and BPM, BPM had a reduction in estimation error compared to TIM of 4% for monthly, 6% for a.m.-p.m., and 10% for tri-monthly test plans. As frequency of testing decreases, advantages of BPM increase (Norman et al., 1999). For activity year 2005, 70% of cows processed at DRMS were tested under the a.m.-p.m. test plan (AIPL. USDA., 2005). Table 2-1 presents changes in test methods over time resulting from DRMS data. From 1980 to 2000, the a.m.-p.m. test method increased by 52% for herds processed through DRMS. DRMS data were calculated using TIM method whereas a reduction in estimation error by 6% could have been realized by using BPM.

demonstrates the difference between AIPL using BPM after 1995 and DRMS using TIM for their respective calculations. Across all cow records for 2003, subtracting the BPM value from the TIM, increased milk per record by 112 kg. When comparing cows completing lactations in less than 305 days in milk (DIM), a difference of 575 kg of milk resulted between methods of calculation with BPM being higher. A similar resulting 553 kg difference between TIM and BPM was reported by AIPL between 1996 and 1997 in the national standardized Holstein breed average (Powell and Sanders, 1999). In contrast, if a cow completed her lactation with DIM in greater than 305 days, TIM method was actually higher than BPM by 128 kg of milk in 2003. An initial increase in difference between AIPL and DRMS occurred around 1987. Since AIPL was reporting DRMS standardized record calculations prior to 1997, further research is required to understand this initial increase.



Replacement of TIM with BPM for computing 305-d records occurred in February 1999 for genetic evaluations at AIPL (VanRaden et al., 1999). Lactation records were retrospectively recomputed by AIPL to 1997. AIPL did not obtain test-day information from the DRPCs prior to 1997, therefore, the TIM method calculated by the DRPCs were utilized prior to 1997 (Powell and Sanders, 1999). Even though test day information is transferred between DRPCs and AIPL consistently after 1997, estimation of 305-d lactation milk yield is performed at each individual DRPC and by AIPL. Dairy Record Management Systems plans to change to the BPM method in late 2008 (J. S. Clay, DRMS, Raleigh, NC, personal communication) and other DRPCs in the United States have implemented or are in the process of implementing BPM. Once that is done, there will be consistent methods of calculating current records between AIPL and DRPCs. However, without a retrospective recomputation of historical production data at the point of BPM implementation at DRPCs, historical milk information based on TIM calculations will not be comparable to current standardized data.

### **Recomputation of Previous Days Open**

When a breeding date is reported, DHIA uniform operating procedures require DRPCs to calculate days open as effective breeding date minus previous fresh date (National Dairy Herd Information Program, 2002). The assumption made is that for records without breeding dates, DRMS calculates Previous Days Open (PDO) on cows past the voluntary waiting period as 10 days after the last test day (Fetrow et al., 1990). It was

not possible to confirm the DRMS calculation on cows without breeding dates because the last test day of each record was not included in the DRMS data source.

Estimated effective breeding dates were calculated for non-termination records on those lactations not beginning with an abortion. Estimated effective breeding date or pregnancy date was determined by using interval between successive calvings and subtracting 280 days, the average gestation length for Holsteins, from the difference (Fetrow et al., 1990). The estimated effective breeding date resulted in a close approximation to actual days opens for the lactation.

A 14-day difference was considered “non-error” or “within” an acceptable difference when comparing to days open information provided by DRMS, with days open calculated by subtracting 280 day from the interval between successive calvings. An overall error rate of 19.7% occurred in the DRMS database, with 24.6 % (4.8 % of total) of those occurring for cows with a reported breeding date and 75.4% (14.8 % of total) of those on cows without a reported breeding date.

## **Retrospectively Recomputed Lactation Values**

Dairy Records Management Systems uses a two-file system in handling lactation information for cows processed by their system. The two-file system used by DRMS may be consistent with other DRPCs but that has not been verified. Any cow that is currently in a producer’s herd is stored in an “active” file containing her current and previous lactation

information. Once a cow is removed from the herd, all her lactation records are moved to an historical file.

Adjustments are performed for standardizing milk information at the time of record completion while the record is maintained in the “active” file. When new factors are implemented, DRMS recomputes the standardized milk for cows in the “active” file, so that the most current information is available on active cows for purposes of herd management. However, no retrospective recalculation is performed in the historical file (J. S. Clay, DRMS, Raleigh, NC, personal communication).

Because of the nature of genetic research in predicting genetic trend and ancestor merit, it becomes necessary for AIPL to recompute consistently and retrospectively all data fields available as changes arise in record standardization. In contrast, a DRPC does not have the same need to recompute retrospectively their data files. When research is required on historical information, data from various DRPCs can be a result of a mixture of adjustments and changes made over time.

### **Adjustment Factors for Lactation Data**

Historically, adjustments were made to a cow’s lactation record for the following reasons: (1) to remove biases from comparisons of cows (or groups of cows) of different ages, (2) to reduce sampling variations due to unequal ages, and (3) to estimate what a specific record most probably would have been if all conditions were the same except for

the age of the cow (Miller, 1973). Adjusting lactation records to a standard basis, provides a producer with a simple way to compare all cows in their herd on an equal basis.

From 1974 to 1994, adjustment factors developed by McDaniel et al. (1967) were deployed for calving age and season. During that period, management improvements resulted in reduction in seasonal calving, earlier maturity of cows, and reduced impact of summer heat and humidity (Schutz, 2004; Norman et al., 1995). These factors were established from models that did not consider relationships among animals or differences due to genetic trend. Because genetic relationships had not been considered, research showed an over-adjustment on records of young cows resulting from genetic superiority being reflected in the age solutions (McDaniel, 1976).

In January 1995, age-parity-season factors for cows of the same age, but differing in lactation numbers and in time periods, were implemented (VanRaden et al., 1995). Genetic trend was reduced, and accuracies of comparisons of animals in different years were improved with the new age-parity-season lactation factors (VanRaden et al., 1995).

Yield traits are affected by a cow's current and previous reproductive status, age, parity, and previous days open and these factors are all interrelated (VanRaden et al., 1995). The January 1995 adjustment is multiplicative in nature, simultaneously adjusting for effects of age, parity, season, and previous days open to standardize records.

As pointed out earlier, Figure 2-1 demonstrates difference between BPM and TIM across time. AIPL data are recomputed retrospectively with the exception of the BPM change in 1997 due to AIPL lack of test day information prior to that period (Powell and Sanders, 1999), whereas DRMS data are not. After 1997, consistent adjustment factors for

the two groups were used. For a researcher using DRPC data, the January 1995 adjustment will be different for lactation completing prior and after 1995.

## **Consistent Standardization**

The data received from the DRPC consisted of mature equivalent value for milk standardized using two standardization methods. Age-season adjustments from 1974 through 1995 developed by McDaniel et al. (1967) were used. After 1995, DRMS used an updated version of age-parity-season adjustments (VanRaden et al., 1995). Standardized values for milk, protein, fat, somatic cell score, and pregnancy rate are available through the AIPL formats.

## **Conclusions**

At the current time, when considering the source of lactation data for genetic research, there are differences in data quality edits implemented between AIPL and DRPC data. Because there is no requirement for the DRPCs to transfer ancestral and historical information, when utilizing DRPC's data in research, length of time the herd or state has been processed at the particular DRPC will determine the amount of historical information received.

AIPL uses BPM for estimating milk and component yields, whereas DRMS uses TIM at this time. AIPL data are recomputed retrospectively, whereas data obtained from DRMS are calculated as a cow completes the lactation. Because DRMS does not recompute retrospectively its data, the change in age-parity-season factors that occurred in

1995 causes their historical data to contain a mixture of estimates based on 1974 age-season adjustments and 1995 age-parity-season adjustments.

All data collection systems have to balance accuracy and completeness with cost to maintain that accuracy. For an informational system to be successful, benefits to management from data quality, quantity, and distribution must clearly outweigh any financial and time costs associated with equipment and labor (Wiggans, 1994). DRPCs and AIPL have found successful balances in their areas of expertise.

A researcher must keep in mind the purpose of the organization in determining the best data source for a project. The primary objective of a DRPC is to provide dairy records services to producers, consultants, veterinarians, nutritionists, DHIA affiliates, and Extension Specialists. In contrast, the objective of AIPL is to conduct research to discover, test, and implement improved genetic evaluation techniques for economically important traits of dairy cattle. Use of consistently standardized historical data is more critical for AIPL objectives than for those of various DRPCs.

## References

- AIPL. USDA.** 2008. AIPL summary of changes (chronological) Online: [http://aipl.arsusda.gov/reference/changes/summary\\_chrono.html#history](http://aipl.arsusda.gov/reference/changes/summary_chrono.html#history). Accessed April 7, 2008.
- AIPL. USDA.** 2005. Dairy Records Processing Center Activity Summary. Natl. Coop. DHI Prog. Handbook, Fact Sheet K-6. Washington, D.C. Online: <http://aipl.arsusda.gov/publish/dhi/dhi06/drpcall.html>. Accessed December 13, 2007.
- Fetrow, J., D. McClary, R. Harman, K. Butcher, L. Weaver, E. Studer, J. Ehrlich, W. Etherington, W. Guterbock, D. Klingborg, J. Reneau, and N. Williamson.** 1990. Calculated selected reproductive indices: recommendations of the American association of bovine practitioners. *J. Dairy Sci.* 73:78-90.
- Majeskie, J. L.** 1992. Status of United States dairy cattle. Natl. Coop. DHI Progr. Handbook. Fact Sheet H-7. Ext. Sev. USDA, Washington, DC.
- McDaniel, B. T., R. H. Miller, E. L. Corley, and R. D. Plowman.** 1967. DHIA age adjustment factors for standardizing lactations to a mature basis. Dairy Herd Improvement Letter, ARS 44-188, USDA.
- McDaniel, B. T.** 1976. Selection goals for dairy cattle. Proc. Natl. Workshop Genet. Improv. Dairy Cattle. St Louis, MO.
- Miller, P.** 1973. A recent study of age adjustment. *J. Dairy Sci.* 56:952-958.
- Murrill, F. D.** 1985. What is NCDHIP? Natl. Coop. DHI Progr. Handbook. Fact Sheet H-1. Ext. Sev. USDA, Washington, DC.
- National Dairy Herd Information Program.** 2002. National Dairy Herd Information Program - Uniform Operating Procedures (3 parts) Online: <http://www.dhia.org/mission.asp>. Accessed April 7, 2008.
- National Dairy Herd Information Program.** 2000. Standard transfer format references. Online: <http://www.dhia.org/stfref.asp> Accessed: April 6, 2008.
- Norman, H. D., L. G. Waite, G. R. Wiggans, and L. M. Walton.** 1994. Improving accuracy of the United States genetics database with a new editing system for dairy records. *J. Dairy Sci.* 77:3198-3208.

- Norman, H. D., T. R. Meinert, M. M. Schutz, and J. R. Wright.** 1995. Age and seasonal effects on Holstein yield for four regions of the United States over time. *J. Dairy Sci.* 78:1855-1861.
- Norman, H. D., P. M. VanRaden, and J. R. Wright.** 1999. Comparison of test interval and best prediction methods for estimation of lactation yield from monthly, A.M.-P.M., and trimonthly testing. *J. Dairy Sci.* 82:438-444.
- Powell, R. L. and A. H. Sanders.** 1999. AIPL State and National Standardized Lactation Averages by Breed for Cows Calving in 1999. Online: <http://aipl.arsusda.gov/publish/dhi/dhi01/k2.shtml>. Accessed November 30, 2008.
- Sargent, F. D., V. H. Lytton, and O. G. Wall Jr.** 1968. Test Interval Method of calculating Dairy Herd Improvement Association records. *J. Dairy Sci.* 51:170-179.
- Schutz, M. M.** 2004. Age-season standardization. AIPL Res. Rep. Online: <http://aipl.arsusda.gov/reference/ageseason.htm>. Accessed December 12, 2007.
- Shook, G. E., E. L. Jensen, and F. N. Dickinson.** 1980. Factors for estimating sample-day yield in AM-PM sampling plans. *Dairy Herd Improvement Letter - U. S. Department of Agriculture, Science and Education Administration, Northeastern Region.* 56:25-30.
- VanRaden, P. M.** 1997. Lactation yields and accuracies computed from test day yields and (co)variances by best prediction. *J. Dairy Sci.* 80:3015-3022.
- VanRaden, P. M., G. R. Wiggans and C. P. Van Tassell.** 1999. Changes in genetic evaluations (February 1999). Online: <http://aipl.arsusda.gov/reference/changes/chng9902.html>. Accessed December 13, 2007.
- VanRaden, P. M., G. R. Wiggans, R. L. Powell and H. D. Norman.** 1995. Changes in USDA-DHIA genetic evaluations (January 1995). AIPL Res. Rpt. 1995; CH3. Online: <http://aipl.arsusda.gov/reference/changes/chng951.pdf>. Accessed December 12, 2007.
- Voelker, D. E.** 1981. Dairy Herd Improvement Associations. *J. Dairy Sci.* 64:1269-1277.
- Wiggans, G. R.** 1994. Meeting the needs at the national level for genetic evaluation and health monitoring. *J. Dairy Sci.* 77:1976-1983.



Table 2-1 Change in proportions of cows in Traditional versus A.M.-P.M. testing plans over time for herds processing through Dairy Records Management Systems (DRMS)

Year	Traditional		A.M.-P.M.	
	N	Percent	N	Percent
1980-1984	582,160	97.65	13,990	2.35
1985-1989	797,650	87.37	115,333	12.63
1990-1994	932,521	75.18	307,886	24.82
1995-1999	689,214	56.63	527,864	43.37
2000-2004	1,021,324	42.85	1,362,073	57.15

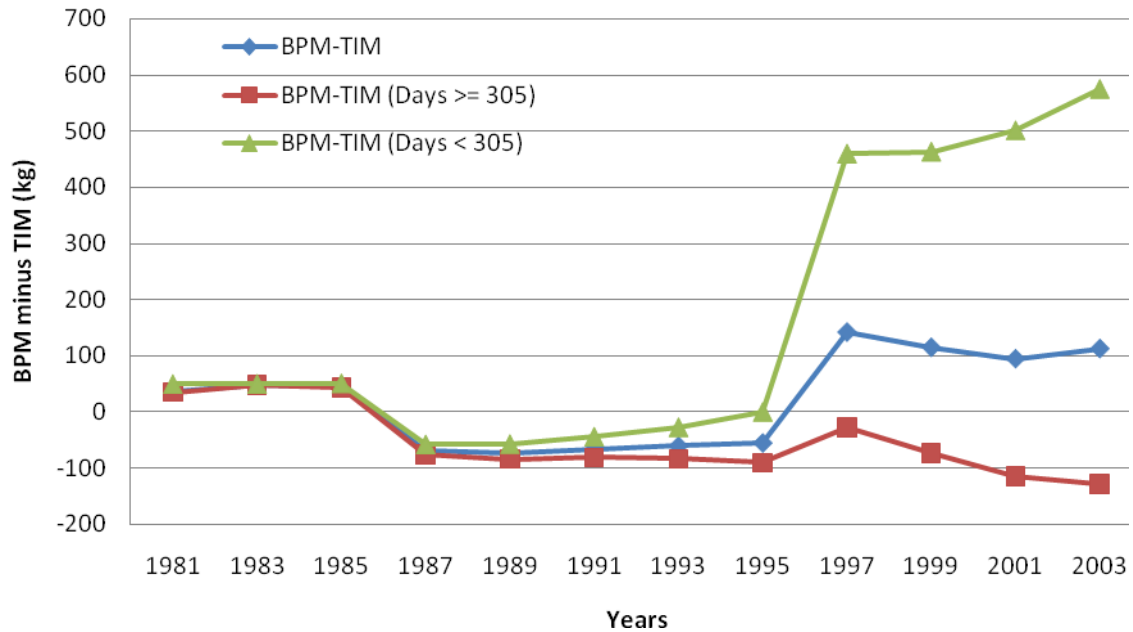


Figure 2-1. Differences between 305-d standardized milk values for best prediction method (BPM)<sup>1</sup> minus test-interval method (TIM) overall or for cows completing lactations longer or shorter than 305 d.

<sup>1</sup>The best prediction method (BPM) was calculated at USDA Animal Improvement Programs Laboratory (AIPL) and test-interval method (TIM) was calculated at Dairy Records Management Systems (DRMS)

## CHAPTER 3

# FIELD DATA HERD EVALUATION OF THE EFFECTS OF MATERNAL LINEAGE ON PRODUCTION AND FERTILITY TRAITS OF HOLSTEIN CATTLE

### Introduction

Reproductive inefficiency in dairy cattle has been the subject of great debate (Call and Stevenson, 1985). Annual milk yield per cow has increased 3.3 fold from 2,410 to 8,061 kg since the 1950s, whereas pregnancy rates of lactating cows have continued to decline (Butler and Smith, 1989). Royal et al. (2000) showed that pregnancy rate to first insemination declined from 56% in late 1970s to 40% in late 1990s. In a 100-cow herd, using Pennsylvania and United States information, net cost of rearing dairy replacements to get them in the milking string was estimated at \$32,344 or \$323 per cow (Tozer and Heinrichs, 2001). Reducing calving age from 25 to 21 months resulted in an 18% reduction in estimated rearing costs (Tozer and Heinrichs, 2001). Reproductive issues are very important to the producer because economic costs of reproductive inefficiency are second only to those of mastitis (Mee, 2007).

Sire analysts and producers have long claimed that certain cow families, referred to as maternal lineages, have special attributes. It is common for a sire analyst to require a potential bull-dam not only to have a high genetic evaluation for yield traits, but also to originate from several generations of outstanding females. This thought process is not

without merit, as higher heritability for milk production has been reported from daughter-dam regression than from paternal half-sib analysis (Rendel et al., 1957; Van Vleck and Bradford, 1965; Thompson and Loganathan, 1968; Seykora and McDaniel, 1983). Cows born into outstanding lineages, which then receive preferential treatment, could possibly cause higher heritability (Gibson et al., 1997). Other explanations, from a genetic standpoint, could be nuclear maternal inheritance, cytoplasmic maternal inheritance, or maternally based parental imprinting (Gibson et al., 1997). In principle, those three genetic explanations of maternal inheritance can be separated from one another because there is a unique pattern of covariance among relatives for each (Rothschild and Ollivier, 1987; Gibson et al., 1988). In reality, distinguishing among those effects, and ruling out other effects such as sex linkage and maternal environmental effects, can be extremely difficult (Gibson et al., 1997).

Even though maternal effects may arise through several pathways, two maternal paths frequently examined in maternal lineage research are additive effects and cytoplasmic effects. Additive maternal effects are inherited in a Mendelian manner. These effects are strictly environmental in regards to offspring upon which they are measured. Phenotypic differences among dams for maternal effects are expressed only in the phenotypic values of their offspring. Maternal effects can be defined as any influence from a dam on its offspring, excluding effects of directly transmitted genes, that affects performance of the offspring (Legates, 1962).

National genetic evaluations currently do not use maternal lineages (ML). Disregarding ML, if its effects are significant, would decrease accuracy of selection. Researchers disagree on whether selection for cytoplasmic effects in genetic evaluations would be feasible.

O'Neill and Van Vleck (1988) concluded that total genetic gain would be increased only slightly in selection for cytoplasmic effects. Selection for cytoplasmic effects can be done directly through only the dam-to-dam pathway. A cow's cytoplasmic genes are transmitted to her offspring, but her sons will not transmit them to their daughters nor express maternal lineage effects. Selection pressure in the dam of sire dam is not very intense. Overestimation of heritability would have little effect on genetic selection differentials as compared to using correct estimates of heritability (O'Neill and Van Vleck, 1988).

Expected genetic improvement per year is calculated as estimated genetic superiority of selected individuals over their contemporaries, averaged over the age of selected animals (Rendel and Robertson, 1950). Even though the numerator of this formula has low selection intensity, the denominator has shown dramatic change over years. Estimates of genetic trend for dam to sire dam pathway showed a doubling of genetic gain per year from 1968 to 1979 (Van Tassell and Van Vleck, 1991).

Accuracy of bull-dam selection is very important for maximizing genetic gains. With simulated data, Boettcher et al. (1996a) compared correlations between true and predicted breeding values of two genetic evaluations. One evaluation accounted for

cytoplasmic effects at the level of 2.5% of the phenotypic variance and the other did not account for cytoplasmic effects. With a mitochondrial effect of 10% of the phenotypic variance, Boettcher et al. (1996a) concluded that progeny testing programs could decrease number of bulls sampled by 8% by correctly accounting for cytoplasmic effects. An 8% reduction in number of bulls sampled would result in a savings to the AI industry of \$2.5 million per year because of the costs associated with testing.

The objective of this study was to determine effect maternal lineage has on production and fertility traits in Holstein cattle.

## **Materials and Methods**

Four sources of data were obtained for analysis: 1) Lactation records from Dairy Records Management Systems (DRMS) in Raleigh, NC, 2) Genetic information from Animal Improvement Programs Laboratory (AIPL) of the USDA for sire data, 3) Genetic information from AIPL for cow data on the November 2007 genetic evaluation, and 4) Standardized lactation yields from AIPL.

Data were obtained through the dairy records processing center (DRPC) in Raleigh, NC in August 2005. Data included historical lactation records dating from January 1980 to August 2005 from 13 states. States included Alabama, Georgia, Florida, Indiana, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, Virginia, and Vermont. The 13 states used in this analysis accounted for 14% of cows on official Dairy Herd Improvement (DHI) testing plans and 15% of cows on management DHI testing plans (Wiggans, 2000). The subset of states had 760 kilograms (kg) less milk

for official DHI testing plans than the United States average of 9,491 kg. Wiggans (2000) reported similar results with the same subset of states resulting in 689 kg less milk production than the United States average.

All cows in the study were required to be Holsteins from a herd with at least 70% sire identification of their cows. A total of 1,132 herds with 1,070,568 cows were available for family lineage. Of herds having 70% sire identification, 728,715 (68.07%) lactation records were obtained that had dam calving dates matching birthdates of their daughters. Lactations after fifth parity were excluded because of their reduced value in estimating genetic merit (Wiggans and VanRaden, 1989).

### **Daughter-Dam Estimates of Heritability**

Heritabilities for milk, daughter pregnancy rate (DPR), and somatic cell score (SCS) were estimated by daughter-dam regression. Calculation of regression coefficients was obtained by PROC MIXED procedure of SAS (SAS Institute Inc., 2004). Daughter mature equivalent milk, DPR, and SCS were regressed on the corresponding dam record. Heritability estimates were performed on families having a fifth generation daughter.

### **Management Group**

Standardized values received from AIPL contained pre-adjustments for environmental effects. These pre-adjustments included parity, age at calving and month of calving within breed, region, time (5-year groupings across time), number of milkings per

day (adjusted to twice daily milking), previous days open, and heterogeneous variance for milk. Base age for mean and variance adjustments was 36-month-old second parity cows. Unequal variances across time, herds, and breed were adjusted to the Holstein base variance calculated from standardized records of first lactation cows that calved in 2002 (Schutz, 2004). Daughter pregnancy rate (DPR) contains pre-adjustments by season based on month of calving, with heterogeneous variance adjustments using the same procedures developed for yield traits (VanRaden et al., 2004).

Each individual lactation record was assigned a herd-year-season (HYS). Season was split into summer (April through September) and the remainder of months categorized as winter. Means and variances were obtained per herd-year-season, resulting in a management group per record. Each management group was required to have at least three records per herd-year-season.

Standardized values for milk, DPR, and SCS were obtained using the following formula.

$$\text{Standardized trait} = \frac{\text{Individual measure} - \text{Management group mean}}{\text{Management group standard deviation}}$$

## **Family Structure**

Cows without sister relationships from historical records with birth years of 1980 or 1981 were included as foundation cows. Sister relationships were removed to avoid across maternal family relationships. Records of cows born before 1980 were excluded from the



analysis. Historical records included cows calving and completing lactations through August 2005. Families were required to have a 5th generation daughter for analysis. Maternal lineages on 29,736 5th generation daughters were established in 767 herds accounting for 9,875 lineages. Table 3-1 contains number of cows per generation in the maternal lineage.

A maternal family mean of performance was calculated in two ways. First, the unweighted mean (U) of all the individual's daughters for generations one through four was determined. Second, a weighted mean (W) of the individual's performance was determined by weighting each daughter's record by degree of relationship (DOR) with the individual (free of progeny information). DOR was determined by coefficient of relationship between relatives (Falconer, 1989). Coefficient of relationship is the probability that alleles at a particular locus chosen at random from two individuals are identical by descent (Falconer, 1989). Mean weighted by DOR was calculated as shown below where PGdor is used to represent the previous generation DOR.

$$\frac{(.5 \times \text{Dam})(.5 \times \text{PGdor})[(.25 \times \text{Sister})(\text{Num of Sisters})][(.125 \times \text{Cousin})(\text{Num of Cousins})]}{(.5)(.5)(.25 \times \text{Num of Sisters})(.125 \times \text{Num of Cousins})}$$

## **Procedure 1**

Cows having historical records dating back until 1980 with milk production and conception equal to or higher than contemporary herdmates were identified. Daughters and granddaughters of those cows were followed, using the same screening criterion, to determine how many resultant cow families remained out of the initial population

investigated. Family characteristics by parities and generations were examined by Model [1] described below. For this and subsequent models, analyses were conducted separately on 1st parity, 1st and 2nd parity and 1st through 5th parity cows, respectively.

Analyses were conducted by combining different numbers of parities. The 1st and 2nd parity were combined for analysis because of results from previous work completed on cow fertility (VanRaden et al., 2004). The 1st through 5th parities were grouped because national yield and DPR evaluations include up to 5 parities per cow (VanRaden et al., 2004). Lactations after fifth parity were excluded because of their reduced value in estimating genetic merit (Wiggans and VanRaden, 1989).

Screening criteria used two types of selection methods for multiple-trait selection. The first selection method was independent culling level (C), resulting in selection of animals exceeding a minimum level for each trait (Falconer, 1989). Family average per generation for milk and DPR were required to be at or above herdmates for both traits. The second selection method involved index selection (I). It combined production and conception on an equal emphasis basis into an index score (Falconer, 1989). Families were grouped into three categories (above, mixed, and below). To be included in the “above” category, the average of family members by generations had to be above herdmates for C and I methods of selection. The “below” category included those below herdmate families by generation. A family that had one trait above and the other below was placed in the “mixed” category.

## **Model**

Characteristics of families by generation were evaluated using Model [1]. Characteristics included months of productive life, percent inbreeding, net merit dollars, and predicted transmitting ability (PTA) for milk, pregnancy rate (DPR), somatic cell score (SCS) and productive life (PL).

$$Y_{ijk} = \mu + \text{Gen}_i + \text{Herd}_j + e_{ijk} \quad [1]$$

where,

$Y_{ijk}$	=	Families for months of PL, percent inbreeding, net merit dollars, and PTA for milk, DPR, SCS, and PL
$\mu$	=	Overall mean
$\text{Gen}_i$	=	Fixed effect of $i_{\text{th}}$ generation
$\text{Herd}_j$	=	Random effect of $j_{\text{th}}$ herd
$e_{ijk}$	=	Error

## **Procedure 2**

Daughters from cow families identified in Procedure 1 that had four successive generations of above average fertility and above average production when compared to their contemporary herdmates were analyzed. Each family was placed into one of three categories based on results of Procedure 1. To be included in the “above” category the average of family members by generations were required to be above herdmates utilizing C and I methods of selection. The “below” category included those below herdmate families

by generation. A family that had one trait above and the other below was placed in the “mixed” category.

Analyses were performed on fifth generation family members to determine if milk production, DPR and SCS in the fifth generation were associated significantly with previous performance of the respective cow families using Model [2] below.

### **Model**

Fifth generation daughters using the first four generations of maternal family means unweighted or weighted by degree of relationship to daughter was evaluated using Model [2].

$$Y_{ijkl} = \mu + Fam_i + PTAs_j + PTAm_k + PTAs_l + PTAm_m + e_{ijklmn} \quad [2]$$

where,

$Y_{ijkl}$	=	5th Generation daughter yield deviation for milk, DPR, or SCS
$\mu$	=	Overall mean
$Fam_i$	=	Effect of $i_{th}$ family group for selection method
$PTAs_j$	=	Sire predicted transmitting ability for milk
$PTAm_k$	=	Maternal grand sire predicted transmitting ability for milk
$PTAs_l$	=	Sire predicted transmitting ability for daughter pregnancy rating
$PTAs_m$	=	Sire predicted transmitting ability for daughter pregnancy rating
$e_{ijklmn}$	=	Error

In an effort to achieve unbiased estimates, family groups were treated as a random effect. Boettcher et al. (1996b) reported that treating maternal lineage as a random effect

rather than a fixed effect when maternal lineage was statistically significant resulted in a superior model.

### **Procedure 3**

In an effort to consider potential impact of subsequent generations on selection, family values from generations one through four were calculated 1) based on unweighted average of family members or 2) averages weighted by degree of relationship to the individual. The hypothesis of procedure 2, that daughters from cow families identified in 4 successive generations of high fertility and above average production will be at least equal in production but superior in reproduction to contemporary herdmates, being false, single trait selection was one method used to determine if maternal family lineage effects exist. Each family was ranked and placed into one of three groups based on average deviations for the traits analyzed. The three groups were the top 25%, intermediate 50%, and bottom 25% for each trait. Analyses were performed on fifth generation members to determine if milk production and pregnancy rates in the fifth generation were significantly associated with historical performance of respective cow families.

### **Model**

Fifth generation daughters were evaluated using the first four generations of maternal family members unweighted or weighted by degree of relationship to daughter.

$$Y_{ijkl} = \mu + Fam_i + PTAs_j + PTAm_k + PTAs_l + PTAm_m + e_{ijklmn} \quad [2]$$

where,

$Y_{ijkl}$	=	5th Generation daughter yield deviation for milk, DPR, or SCS
$\mu$	=	Overall mean
$Fam_i$	=	Effect of $i_{th}$ family group for respective trait
$PTAs_j$	=	Sire predicted transmitting ability for respective trait
$PTAm_k$	=	Maternal grand sire predicted transmitting ability for respective trait
$PTAs_l$	=	Sire predicted transmitting ability for milk, DPR, or SCS
$PTAs_m$	=	Sire predicted transmitting ability for milk, DPR, or SCS
$e_{ijklmn}$	=	Error

## Results

Table 3-2 contains heritability estimates and standard errors calculated by regression of offspring on maternal parent (Falconer, 1989). Daughter mature equivalent milk, DPR, and SCS were regressed on the corresponding dam record. Heritability estimates decreased slightly from 0.37 for milk for a cow's first parity to 0.34 when including the first through the fifth parities. Heritabilities of DPR and SCS increased from 0.04 to 0.10 and 0.18 to 0.25, respectively, as number of parities increased.

Table 3-3 compares average standardized data for milk, DPR, and SCS of families by parities within generations using Model [1] for analysis. Generation effect was

significant ( $P < 0.05$ ) for all traits. Significant differences existed overall for generations for all traits measured ( $P < 0.05$ ). Significant differences existed between generations with the exception of generations 4 and 5 in SCS. Somatic cell score for the first generation was unavailable because initial calving was in 1982 and 1983 while AIPL began reporting SCS in 1984 (AIPL. USDA., 2006).

From generations 1 to 5, an increase of 1,811 kg in milk occurred for first and second parities while DPR decreased -8.57 percentage units, and SCS increased 0.14 points from generation 2 to 5. Coefficients of determination were 0.32, 0.10, 0.10 for milk, DPR, and SCS, respectively. The coefficient of determination was greatest for milk and smallest for DPR and SCS, which is consistent with heritability estimates for the three traits. An increase in coefficient of determination occurred for all traits as number of parities included in calculations increased. However, increase in coefficient of determination when going from two parities to five parities was small.

Deviation of family average from herd merit by parities within generation for Milk, DPR, and SCS using model [1] are reported in Table 3-4. In this case, the amount of variability explained by the model decreased as number of parities increased for all traits. Overall, generational effect was significant for all traits while generational effects between generations were significant, with the exception of generation 4 to 5 for SCS.

When first and second parities were combined, difference from contemporary herdmates for milk decreased from generation 1 to 5 by 117 kg and 1.47 percentage units for milk and DPR, respectively, and increased by -0.09 points for SCS. Family difference

from contemporaries for milk stayed above herdmates throughout all generations whereas DPR decreased below herdmates in the fourth generation. These results may reflect higher heritability for milk compared with DPR or the producers' selection goals. Lack of selection criteria for DPR may have contributed to the decline in observed in DPR.

Table 3-5 and Table 3-6 utilize information from the November 2007 genetic evaluation with Model [1] for analysis. Proportion of cows in the database that had genetic evaluations reported through AIPL was 92% (not shown in table). Productive life, the length of time an individual cow remains in the herd after first calving, decreased by 14.5 months over five generations (Table 3-6). Family foundation cows born in 1980 or 1981 (generation 1) had a difference from herdmates of 9.0 months (Table 3-6). Productive life for herdmates of foundation cows was 4.1 months longer than breed average in 1980 of 31.4 months (AIPL. USDA., 2006). The families' foundation animals stayed in the herd an average of 13.1 months longer than breed average. The 5<sup>th</sup> generation family members had 0.12 months more longevity over herdmates and 2.4 months greater than breed average for months of productive life in 1996 (AIPL. USDA., 2006).

Percent of inbreeding increased from 0.93 to 4.25 over five generations for families in study. This 3.32 percentage point increase in inbreeding is consistent with a 3.07 percentage point breed average increase over the same time period (AIPL. USDA., 2006). Due to the complete nature of the pedigree information of the families, a higher inbreeding coefficient is expected. Incomplete pedigree information will reduce estimates of inbreeding (Lutaaya et al., 1999). Inbreeding depression for lifetime performance traits for



Holsteins was 2 to 2.5 times larger than that in grade populations that had higher levels of missing ancestor information (Smith et al., 1998).

## **Procedure 1**

Table 3-7 and Table 3-8 display the resulting family members by generation screened for milk and DPR by multi-trait selection methods: independent culling or selection index, respectively. Table 3-7 demonstrates family average change for milk and DPR compared to herdmates through the four generations. Independent culling level selection methodology was utilized; both milk and DPR were required to be above contemporary herdmates. When using family members first parities and family members averaged per generation, 561 (11%) families had their first and second generations above contemporary herdmates. As the number of generations increased, families with generational averages above contemporary herdmates for milk and DPR for each generation decreased to 66 (1%) by the fourth generation. Increasing the number of lactations included in the analysis did not change the overall percentage of families above contemporary herdmates for four generations for milk and DPR, which remained at less than 1 percent when each generation was averaged. Weighting families by degree of relationship resulted in number of families above herdmates increasing to 551 (10%) in the fourth generation.

Table 3-8 exhibits results for the second selection method, index selection that combined milk and DPR on an equal basis into an index score. When utilizing the

unweighted mean of all individual's daughter performance per generation with the first parity, 1,750 (33%) families had index scores for first and second generations above contemporary herdmates. As the number of generations increased, the number of families with generational average index scores above contemporary herdmates for each generation decreased to 634 (12%) by the fourth generation. Weighting maternal families by degree of relationship resulted in more families being above contemporary herdmates over four generations. For first parities in the fourth generation, weighting by DOR resulted in 1,836 (35%) of the families above contemporary herdmates for all four generations.

## **Procedure 2**

Two types of selection methods for multiple-trait selection were used in Procedure 1. The first was independent culling where the family average per generation was required for milk and conception rate at or above herdmates for both traits. Results from Table 3-7 indicate extremely unbalanced numbers of families by grouping that would potentially affect further analysis by multi-trait independent culling. For example, when using family members first parities averaged per generation the treatment groups would contain 66, 5,196, and 10 families for above, mixed, and below groups, respectively. I cannot find where the numbers come from in Table 3-7.

Generations 1 to 4 from Table 3-7 were utilized for the grouping in Table 3-9. Unbalanced treatment groups affects heterogeneity of variance across cells and problems arise in obtaining valid estimates of standard errors (Littell et al., 2002). Another critical

issue with unbalanced data is the inability to construct meaningful linear combinations of model parameters for hypothesis testing and estimation (Littell et al., 2002). In an effort to minimize the effect of unbalanced data, PROC MIXED was used for analyses to assign appropriate weights regardless of the number of observations per cell (SAS Institute Inc., 2004) and to calculate adjusted least squares means.

The extreme unbalance of the treatment groups may have contributed to the non-significance observed in Table 3-9. When using multi-trait independent culling levels, significance was observed with first parity milk when maternal family members were averaged and with first parity index when degree of relationship to daughters was utilized.

Index selection reported in Table 3-8 and Table 3-10 was the second of the multiple-trait selection methods analyzed. Generation 1 to 4 from Table 3-8 was utilized for the grouping in Table 3-10. Treatment data is still unbalanced when utilizing index selection but not to the extreme level observed with the independent culling method. Increases in the “above” group were 11%, 11%, and 8% percentage points by parity group, respectively, when maternal family members were averaged. Large percentage (23%, 23%, 21%) increases were observed when families were weighted by degree of relationship. Coefficients of determination were greater, on average, when index selection was utilized in comparisons involving 5<sup>th</sup> generation daughters. Significance resulted for milk and index for most of the groupings.

The lack of significance in pregnancy rate for any categories in multi-trait independent culling and index selection indicates that multi-trait selection for fertility mediated through maternal lineage in Holstein cattle may not be effective.

### **Procedure 3**

The hypothesis of procedure 2, that daughters from cow families identified in 4 successive generations of high fertility and above average production will be at least equal in production but superior in reproduction to contemporary herdmates, being false, single trait selection was utilized in Table 3-11 to Table 3-13 to determine if maternal family lineage effects exist.

Table 3-11 utilized milk for single trait selection. For 1<sup>st</sup> & 2<sup>nd</sup> parities combined, the coefficient of determination was substantially greater for milk (0.12) than for DPR (0.01) or SCS (0.02). Higher coefficients of determination found in analyses for milk were likely due to a higher heritability for milk (0.30) than for SCS (0.12) and DPR (0.04). For milk, the coefficients of determination decreased as later parities were included, whereas, for pregnancy rate use of later parities increased coefficients of determination. This result indicates that selection for family DPR becomes more relevant as the cow completes more parities. Increase in coefficient of determination as later parities are included could be a result of heterogeneous environmental variances over time (Gibson et al., 1997). When selecting for milk only, DPR decreased by -1.05 from bottom to the top 25% of the families selected for milk.

Although significant differences between groups were found for the three traits for 1<sup>st</sup> and 2<sup>nd</sup> lactation combined (Table 3-12), all coefficients of determination were small. A cow's family explained very little of the variation between cows for the variables studied. However, only families with five generations of daughters within a herd contributed to the results throughout this study. Using data for two parities, differences of -128 kg and 7.73 percentage units were shown between the 5<sup>th</sup> generation daughters from the top 25% families and the bottom 25% for milk and DPR, respectively.

Results of single trait selection for SCS are shown in Table 3-13. Milk difference from herdmates doubled in the 1<sup>st</sup> lactation from bottom 25% to top 25% when selecting for SCS. Even with negative correlation between milk and DPR, DPR increased by 0.45 percentage points. When the number of parities increased, DPR was not significant.

Table 3-14 compares 5<sup>th</sup> generation daughters using the first four generations of maternal family members weighted by degree of relationship grouped by PTA productive life. Cows leave herds for multiple reasons, including low yield, mastitis, reproductive failure, sales for dairy purposes, and death (Dentine et al., 1987). Addition of genetic evaluation for productive life in 1995 was the first genetic evaluation that was directly associated with longevity. VanRaden and Wiggans (1995) reported that a single-trait evaluation for productive life was more valuable in prediction of herd life than the collective predicted transmitting ability of 14 type traits.

In August 2006, an economic definition was introduced to replace the 1995 version of productive life calculations. The 2006 revision was important because more than 50% of

the cows had lactations longer than 305 days and weight was given beyond the 305-d period,. With the 2006 revision, weights were based on standard lactation curves, with the highest weight at the peak of lactation and diminishing credit given across the remainder of the lactation (VanRaden and Wiggans, 1995; VanRaden et al., 2006). In addition, the 2006 version gives slightly more weight to later lactations, thereby ensuring that cows with multiple lactations receive more total weight than cows with a single, long lactation. Given that the current PTA for productive life is a multi-trait index that combines traits of significant importance to longevity and fertility, no attempt was made to build another multi-trait index.

Conformation data were not obtained for this study. Therefore, PTA productive life was used for analyses. Since PTA productive life was obtained through AIPL, it includes progeny information, while all other analyses were free of progeny information. Correlations of productive life with milk, SCS, and DPR, were 0.08, -0.38, 0.08, respectively (VanRaden, 2006). PTA DPR received a weight of 15%; milk weight, 1%; and SCS weight, -10% of the relative weight for PTA productive life (VanRaden and Wiggans, 2003).

The group effect remained significant ( $P < 0.05$ ) for DPR of 5<sup>th</sup> generation daughters when families were weighted by degree of relationship. When single trait selection for DPR was executed (Table 3-14), group effects were significant. Coefficients of determination were larger for all parity groupings when family members were averaged with multi-trait selection of productive life compared to a single trait selection of DPR from

Table 3-12. When comparing DPR, the coefficient of determination is greater with single trait selection.

When maternal families were grouped by PTA for productive life, first parity daughters were not significant for milk. As the number of parities increased, the coefficient of determination also increased and group effects for milk became significant. Low relative weight for milk in the productive life model may be a reason that first parity for the 5<sup>th</sup> generation daughters was not significant for milk.

## **Discussion**

An increase of 1,811 kg in milk occurred for first and second parities while DPR decreased -8.57 percentage units and SCS increased 0.14 from generation 1 to 5. From the years 1974 to 2004, annual milk yield per cow has increased from 7,565 to 12,252 kg (AIPL. USDA., 2006). However, the correlation between production and fertility is unfavorable in dairy cows (Pryce and Veerkamp, 2001). The increased milk production has been accompanied by a 6 percentage unit decline in 21-d pregnancy rate since 1980 (Shook, 2006), which is equivalent to an increase of 24 days open. Genetic factors have accounted for over 55% of the phenotypic gains in yield traits and about one-third of the decline in pregnancy rate (Shook, 2006).

Foundation animals of families stayed in the herd an average of 13.1 months longer than breed average and 9.0 months longer than herdmates. Hare et al. (2006) reported the number of parities, productive herd life, and survival rates primarily decreased due to more intense culling based on management decisions rather than genetic factors. Culling rate is

influenced by management style. Farms with moderately intensive grazing had lower culling rates than those with extensive grazing (Hanson et al., 1998). White et al. (2002) reported that seasonal calving Holsteins and Jerseys on pasture in North Carolina had lower culling rates and culling costs than did confined cows.

Inbreeding increased from 0.93 to 4.25 % over five generations for families in this study. Estimated effects of inbreeding on nonproduction traits include decreased reproductive efficiency (Hermas et al., 1987; Smith et al., 1998)), decreased longevity (Huizinga et al., 1986; Smith et al., 1998), and a slight increase in somatic cell score (Miglior et al., 1995). Other research has indicated that somatic cell score was not affected by inbreeding (Thompson et al., 2000). Smith et al. (1998) reported a decrease of 13.1 d in length of productive life for each 1% increase in inbreeding. Inbreeding levels greater than 10% decreased lactation lengths by 2 to 8 d, with survival rates reduced for all levels of inbreeding (Thompson et al., 2000). Hermas et al., 1987 reported a decrease of 3.7 d in age at first calving per 1% increase in inbreeding. Effects of inbreeding have been shown to be more severe as the level of inbreeding increases (Wall et al., 2005).

Independent culling level selection methodology was utilized. Both milk and DPR were required to be above contemporary herdmates. As the number of generations increased, families with generational averages above contemporary herdmates for milk and DPR at each generation decreased to 66 (1%) by the fourth generation. As the number of generations increased, the number of families with generational average selection index scores above contemporary herdmates for each generation decreased to 634 (12 %) by the



fourth generation. An intrinsic problem of fertility analyses is that these traits may have been subjected to censoring due to selection based on milk yield. Censoring is also likely involved in the data used in the current study. Compared to cows with lower milk production, producers may give more opportunities to higher yielding cows to conceive and, possibly, even delay their breeding (Kadarmideen et al., 2003). One of the main advantages of a selection index is that it can provide an unbiased estimate for a trait that is observed only on animals highly selected for milk yield based on early lactations, which can lead to biased fertility observations in later lactations (Kadarmideen et al., 2003).

Use of weighting by degree of relationship resulted in greater differences between the top 25% families and bottom 25% of families for all variables compared to weighting family members equally. Weighting families by degree of relationship placed more emphasis on the dam and later generations of the families; therefore, preferential treatment could play a larger role in the elite families. Elite lineages will generally be subjected to preferential treatment only for a few generations (Gibson et al., 1997).

Cytoplasmic genes are transferred directly from dam to offspring. With no mutation or heterogeneity of the cytoplasmic components within a dam, all her offspring will receive an identical copy of cytoplasmic genes (Southwood et al., 1989). Maternal effects include cytoplasmic inheritance, intrauterine and postpartum nutrition provided by the dam, antibodies and pathogens transmitted from dam to offspring, and maternal behavior (Hohenboken, 1985 ). When weighted by degree of relationship, intrauterine and postpartum nutrition provided by the dam may play a larger role in the daughter

performance. Even though dairy dams do not nurse their young for an extended period, maternal effects from postpartum nutrition could play a role, since dams' colostrum usually is fed to their young for the first few days of life.

## **Conclusions**

The amount of variability explained by family differences by parities within generation for Milk, DPR, and SCS decreased as the number of parities increased. Overall, generational effects were significant for all traits while generational effect between generations were significant with the exception of generations 4 and 5 for SCS. The amount of variation explained by the model decreased with the increase in number of lactations for all three traits. Foundation animals of families stayed in the herd an average of 13.1 months longer than breed average. Inbreeding increased from 0.93 to 4.25% over five generations for families in the study.

With multi-trait independent culling levels, as the number of generations increased, families with generational average above contemporary herdmates for milk and DPR for each generation decreased to 66 (1%) by the fourth generation. When using independent culling selection levels, significance was observed with first parity milk when maternal family members were averaged and first parity index when degree of relationship to daughters was utilized. Extreme unbalance among treatment groups may have been a factor on the non-significance observed in DPR or milk.

With weighting DPR and milk equally to make up an index selection, family members' first parities and family members averaged per generation, 1,750 (33 %) families

had index scores for first and second generations above contemporary herdmates. As the number of generations increased, the number of families with generational average index scores above contemporary herdmates for each generation decreased to 634 (12%) by the fourth generation. Weighting maternal families by degree of relationship resulted in more families ranking above contemporary herdmates over four generations. For first parities of the fourth generation, 1,836 (35%) families had maintained above contemporary herdmates for all four generations.

Using multi-trait selection on productive life, group effect was significant ( $P < 0.05$ ) for DPR of 5<sup>th</sup> generation daughters. Coefficient of determination was larger for all parity groups when family members were averaged with multi-trait selection on productive life compared to single trait selection on DPR. When comparing DPR, coefficient of determination was greater than when single trait selection was practiced. These findings provide evidence that pregnancy rate and milk production, when taking into account degree of relationship, are partially inherited through maternal lineage.

Multi-trait selection for pregnancy rate when weighting milk and DPR equally in Holstein cattle, using independent culling and index selection, was evaluated. No effect from maternal lineage on fertility traits were significant for any categories. It appears there may be potential economic value in considering maternal family history for pregnancy rate when selecting future bull dams using multi-trait selection on productive life, which increases the weight of DPR over milk.

Utilizing single trait selection, differences between groups were significant for three traits tested, but all coefficients of determination were small. A cow's family explained very little of the variation between cows for those variables.

## References

- AIPL. USDA.** 2006. AIPL inbreeding coefficients for Holstein cows (Calculated February, 2006). Online. <http://aipl.arsusda.gov/dynamic/inbrd/current/HOt.html>. Accessed August 14, 2006.
- AIPL. USDA.** 2008. Cow Genetic Trends. Online: [http://aipl.arsusda.gov/eval/summary/trend.cfm?R\\_Menu=HO.s#StartBody](http://aipl.arsusda.gov/eval/summary/trend.cfm?R_Menu=HO.s#StartBody). Accessed March 16, 2008.
- Boettcher, P. J., A. E. Freeman, S. D. Johnston, R. K. Smith, D. C. Beitz, and B. T. McDaniel.** 1996a. Relationships between polymorphism for mitochondrial deoxyribonucleic acid and yield traits of Holstein cows. *J. Dairy Sci.* 79:647-654.
- Boettcher, P. J., M. T. Kuhn, and A. E. Freeman.** 1996b. Impacts of cytoplasmic inheritance on genetic evaluations. *J. Dairy Sci.* 79:663-675.
- Butler, W. R. and R. D. Smith.** 1989. Interrelationships between energy balance and postpartum reproductive function in dairy cattle. *J. Dairy Sci.* 72:767-783.
- Call, E. P. and J. S. Stevenson.** 1985. Current challenges in reproductive management. *J. Dairy Sci.* 68:2799-2805.
- Falconer, D. S.** 1989. *Introductions to quantitative genetics.* 3rd rev. ed. John Wiley & Sons, Inc, New York, NY.
- Gibson, J. P., A. E. Freeman, and P. J. Boettcher.** 1997. Cytoplasmic and mitochondrial inheritance of economic traits in cattle. *Livest. Prod. Sci.* 47:115-124.
- Gibson, J. P., B. W. Kennedy, and L. R. Schaeffer.** 1988. Genetic models for estimation of autosomally inherited genetic effects that are expressed only when received from either a male or female parent. *J. Dairy Sci.* 71: 143(Abstr.).
- Hanson, G. D., L. C. Cunningham, M. J. Morehart, and R. L. Parsons.** 1998. Profitability of moderate intensive grazing of dairy cows in the northeast. *J. Dairy Sci.* 81:821-829.
- Hare, E., J. R. Wright, and H. D. Norman.** 2006. Survival Rates and Productive Herd Life of Dairy Cattle in the United States. *J. Dairy Sci.* 89,:3713-3720.
- Hermas, S. A., C. W. Young, and J. W. Rust.** 1987. Effects of mild inbreeding on productive and reproductive performance of Guernsey cattle. *J. Dairy Sci.* 70:712-715.

- Hohenboken, W. D.** 1985. Maternal effects, page 135-150 in general and quantitative genetics. A.B chapman.4:408. Elsevier Science Pub. Co.Amsterdam ; New York.
- Huizinga, H. A., S. Korver, B. T. McDaniel, and R. D. Politiek.** 1986. Maternal effects due to cytoplasmic inheritance in dairy cattle. Influence on milk production and reproduction traits. *Livest. Prod. Sci.* 15:11-26.
- Kadarmideen, H. N., R. Thompson, M. P. Coffey, and M. A. Kossaibati.** 2003. Genetic parameters and evaluations from single- and multiple-trait analysis of dairy cow fertility and milk production. *Livest. Prod. Sci.* 81:183-195.
- Legates, J. E.** 1962. Heritability of fat yields in herds with different production levels. *J. Dairy Sci.* 45:990-996.
- Littell, R. C., W. W. Stroup, R. J. Freund and R. C. Littell.** 2002. SAS for linear models.466. 4thSAS InstituteCary, NC, USA.
- Lutaaya, E., I. Misztal, J. K. Bertrand, and J. W. Mabry.** 1999. Inbreeding in populations with incomplete pedigrees. *Journal of Animal Breeding and Genetics = Zeitschrift Für Tierzüchtung Und Züchtungsbiologie.* 116:475-480.
- Mee, J. F.** 2007. The role of the veterinarian in bovine fertility management on modern dairy farms [electronic resource]. *Theriogenology.* 68,:257-265.
- Miglior, F., E. B. Burnside, and J. C. M. Dekkers.** 1995. Nonadditive genetic effects and inbreeding depression for somatic cell counts of Holstein cattle. *J. Dairy Sci.* 78:1168-1173.
- O'Neill, K. and L. D. Van Vleck.** 1988. Potential of cytoplasmic effects for selection in dairy cattle. *J. Dairy Sci.* 71:3390-3398.
- Pryce, J. E. and R. F. Veerkamp.** 2001. The incorporation of fertility indices in genetic improvement programmes. Pages 223–236 in *Fertility in the High-Producing Dairy Cow.* M. Diskin, ed. British Society of Animal Science Occasional Publication No. 26., Edinburgh, Scotland.
- Rendel, J. M. and A. Robertson.** 1950. Estimation of genetic gain in milk yield by selection in a closed herd of dairy cattle. *J. Genet.* 50:1-8.
- Rendel, J. M., A. Robertson, A. A. Asker, S. S. Khishin, and M. T. Rabab.** 1957. The inheritance of milk production characteristics. *J. Agric. Sci.* 48:426-432.
- Rothschild, M. F. and L. Ollivier.** 1987. Expectation of variance due to mitochondrial genes from several mating designs. *Genetique, Selection, Evolution.* 19:171-180.

- Royal, M. D., A. O. Darwash, A. P. F. Flint, R. Webb, J. A. Woolliams, and G. E. Lamming.** 2000. Declining fertility in dairy cattle: changes in traditional and endocrine parameters of fertility. *Anim. Sci.* 70:487-501.
- SAS Institute Inc.** 2004. SAS/STAT® 9.1 User's guide. 7 volumes:5180. SAS Institute Inc., Cary, NC.
- Schutz, M. M.** 2004. Age-season standardization. AIPL Res. Rep. Online: <http://aipl.arsusda.gov/reference/agesseason.htm>. Accessed December 12, 2007.
- Seykora, A. J. and B. T. McDaniel.** 1983. Heritabilities and correlations of lactation yields and fertility for Holsteins. *J. Dairy Sci.* 66:1486-1493.
- Shook, G. E.** 2006. Major Advances in Determining Appropriate Selection Goals. *J. Dairy Sci.* 89:1349-1361.
- Smith, L. A., B. G. Cassell, and R. E. Pearson.** 1998. The effects of inbreeding on the lifetime performance of dairy cattle. *J. Dairy Sci.* 81:2729-2737.
- Southwood, O. I., B. W. Kennedy, K. Meyer, and J. P. Gibson.** 1989. Estimation of additive maternal and cytoplasmic genetic variances in animal models. *J. Dairy Sci.* 72:3006-3012.
- Thompson, J. R., R. W. Everett, and N. L. Hammerschmidt.** 2000. Effects of inbreeding on production and survival in Holsteins. *J. Dairy Sci.* 83:1856-1864.
- Thompson, N. R. and S. Loganathan.** 1968. Composition of cow's milk. II. Genetic influences. *J. Dairy Sci.* 51:1933-1935.
- Tozer, P. R. and A. J. Heinrichs.** 2001. What affects the costs of raising replacement dairy heifers: a multiple-component analysis. *J. Dairy Sci.* 84:1836-1844.
- Van Tassell, C. P. and L. D. Van Vleck.** 1991. Estimates of genetic selection differentials and generation intervals for four paths of selection. *J. Dairy Sci.* 74:1078-1086.
- Van Vleck, L. D. and G. E. Bradford.** 1965. Comparison of heritability estimates from daughter-dam regression and paternal half-sib correlations. *J. Dairy Sci.* 48:1372-1375.
- VanRaden, P. M.** 2006. Net merit as a measure of lifetime profit: 2006 version. AIPL Res. Rep. NM\$1 (11-000). Online: <http://aipl.arsusda.gov/reference/nmcalc.htm>. Accessed September 4, 2006.

- VanRaden, P. M., A. H. Sanders, M. E. Tooker, R. H. Miller, H. D. Norman, M. T. Kuhn, and G. R. Wiggans.** 2004. Development of a national genetic evaluation for cow fertility. *J. Dairy Sci.* 87:2285-2292.
- VanRaden, P. M., M. E. Tooker, R. E. Pearson, and C. M. B. Dematawewa.** 2006. Productive life including all lactations and longer lactations with diminishing credits. *J. Dairy Sci.* 89:3213-3220.
- VanRaden, P. M., and G. R. Wiggans.** 2003. Methods used to predict multi-trait productive life: 2006 version. *AIPL Res. Rep. PL2 (11-03)*. Online: <http://aipl.arsusda.gov/reference/multi-pl.htm>. Accessed March 4, 2008.
- VanRaden, P. M. and G. R. Wiggans.** 1995. Productive life evaluations: calculation, accuracy, and economic value. *J. Dairy Sci.* 78:631-638.
- Wall, E., S. Brotherstone, J. F. Kearney, J. A. Woolliams, and M. P. Coffey.** 2005. Impact of nonadditive genetic effects in the estimation of breeding values for fertility and correlated traits. *J. Dairy Sci.* 88:376-385.
- White, S. L., G. A. Benson, S. P. Washburn, and J. T. J. Green.** 2002. Milk production and economic measures in confinement or pasture systems using seasonally calved Holstein and Jersey cows. *J. Dairy Sci.* 85:95-104.
- Wiggans, G. R.** 2000. USDA Summary of 2000 Herd Averages DHI Report K-3. Online: <http://aipl.arsusda.gov/publish/dhi/dhi01/k3.shtml>. Accessed: January 9, 2008.
- Wiggans, G. R. and P. M. VanRaden.** 1989. USDA-DHIA Animal Model Genetic Evaluation Natl. Coop. DHI Progr. Handbook. Fact Sheet H-2. Natl. Coop DHI, Washington, D.C.



Table 3-1 Number of cows and daughters per generation across herds that had family for 5 generations

<b>Generation</b>	<b>Number of Cows</b>	<b>Cows Per Family</b>
1	9,872	1.00
2	19,919	2.02
3	27,504	2.79
4	31,340	3.17
5	29,736	3.01

Table 3-2 Heritability estimates as calculated using daughter-dam regression

Parities Included	<b>Heritability</b>					
	Milk		DPR <sup>1</sup>		SCS <sup>2</sup>	
	Pairs <sup>3</sup>	h <sup>2</sup> (SE)	Pairs <sup>3</sup>	h <sup>2</sup> (SE)	Pairs <sup>3</sup>	h <sup>2</sup> (SE)
1	86,514	0.37 (0.01)	70,460	0.04 (0.01)	54,405	0.18 (0.01)
1 & 2	86,433	0.35 (0.01)	71,001	0.07 (0.01)	55,261	0.22 (0.01)
1 thru 5	86,634	0.34 (0.01)	71,240	0.10 (0.01)	56,754	0.25 (0.01)

<sup>1</sup> Daughter pregnancy rate

<sup>2</sup> Somatic cell score

<sup>3</sup> Number of daughter-dam pairs

Table 3-3 Means across families by parities and generation for average Milk, DPR, and SCS<sup>2,3</sup>

<b>1st Parity Cows</b>						
Gen	Milk (kg) <sup>4</sup>		DPR <sup>4</sup>		SCS <sup>1,5</sup>	
	Mean	SE	Mean	SE	Mean	SE
1	8,422	36.30	28.73	0.23	--	--
2	8,789	33.61	26.55	0.19	3.13	0.02
3	9,327	32.93	25.25	0.17	3.19	0.02
4	9,904	32.73	22.75	0.17	3.26	0.02
5	10,383	32.84	20.95	0.17	3.27	0.02
R <sup>2</sup>	0.3147		0.0763		0.0869	
<b>1st and 2nd Parity Cows</b>						
Gen	Milk (kg) <sup>4</sup>		DPR <sup>4</sup>		SCS <sup>1,5</sup>	
	Mean	SE	Mean	SE	Mean	SE
1	8,522	35.56	28.13	0.21	--	--
2	8,844	33.01	25.70	0.17	3.12	0.02
3	9,349	32.37	24.04	0.16	3.18	0.02
4	9,887	32.19	21.38	0.16	3.24	0.02
5	10,333	32.29	19.56	0.16	3.26	0.02
R <sup>2</sup>	0.3235		0.0975		0.1027	
<b>1st through 5th Parity Cows</b>						
Gen	Milk (kg) <sup>4</sup>		DPR <sup>4</sup>		SCS <sup>1,5</sup>	
	Mean	SE	Mean	SE	Mean	SE
1	8,542	35.02	26.91	0.20	---	---
2	8,837	32.68	24.45	0.16	3.17	0.02
3	9,316	32.09	22.55	0.16	3.24	0.02
4	9,836	31.91	19.81	0.15	3.28	0.02
5	10,275	32.01	18.28	0.15	3.29	0.02
R <sup>2</sup>	0.3348		0.1077		0.1088	

<sup>1</sup> Genetic evaluation for SCS began in 1984 with lower values optimal

<sup>2</sup> Foundation families started in 1980 and 1981

<sup>3</sup> Generation effect is significant ( $P < 0.05$ ) for all traits overall

<sup>4</sup> Generation effect is significant ( $P < 0.05$ ) for traits between generations

<sup>5</sup> Generation effect is significant ( $P < 0.05$ ) for traits between generations except from 4 to 5

Table 3-4 Deviation of family average minus herd merit for Milk, DPR, and SCS<sup>2,3</sup>

<b>1st Parity Cows</b>						
Gen	Difference Milk (kg) <sup>3,4</sup>		Difference DPR <sup>3,4</sup>		Difference SCS <sup>1,5</sup>	
	Mean	SE	Mean	SE	Mean	SE
1	135	18.48	1.45	0.1839	---	---
2	10	13.66	1.12	0.1288	-0.03	0.0161
3	50	12.27	1.34	0.1113	0.00	0.0114
4	109	11.85	0.85	0.1054	0.04	0.0106
5	148	12.09	0.82	0.1084	0.06	0.0108
R <sup>2</sup>	0.0323		0.0137		0.0198	

<b>1st and 2nd Parity Cows</b>						
Gen	Difference Milk (kg) <sup>3,4</sup>		Difference DPR <sup>3,4</sup>		Difference SCS <sup>1,5</sup>	
	Mean	SE	Mean	SE	Mean	SE
1	171	16.29	1.10	0.1545	---	---
2	10	11.37	0.47	0.1085	-0.03	0.0132
3	10	9.83	0.36	0.0936	0.00	0.0088
4	28	9.35	-0.23	0.0886	0.03	0.0081
5	54	9.59	-0.37	0.0912	0.05	0.0082
R <sup>2</sup>	0.0166		0.0136		0.0132	

<b>1st through 5th Parity Cows</b>						
Gen	Difference Milk (kg) <sup>3,4</sup>		Difference DPR <sup>3,4</sup>		Difference SCS <sup>1,5</sup>	
	Mean	SE	Mean	SE	Mean	SE
1	72	15.08	0.37	0.1420	---	---
2	-94	10.32	-0.50	0.0996	0.02	0.0121
3	-114	8.77	-0.78	0.0856	0.05	0.0079
4	-97	8.28	-1.45	0.0808	0.08	0.0072
5	-41	8.52	-1.41	0.0832	0.09	0.0073
R <sup>2</sup>	0.0125		0.0131		0.0107	

<sup>1</sup> Genetic evaluation for SCS began in 1984 with lower values optimal

<sup>2</sup> Foundation families started in 1980 and 1981

<sup>3</sup> Generation effect is significant ( $P < 0.05$ ) for all traits overall

<sup>4</sup> Generation effect is significant ( $P < 0.05$ ) for traits between generations

<sup>5</sup> Generation effect is significant ( $P < 0.05$ ) for traits between generations except from 4 to 5 value

Table 3-5 Means across families by generation for PTA milk, DPR, and SCS for November 2007 genetic information<sup>1,2</sup>

Gen	Family Average					
	PTA Milk <sup>3</sup>		PTA DPR <sup>3</sup>		PTA SCS <sup>3</sup>	
	Mean	SE	Mean	SE	Mean	SE
1	-1,676	6.32	1.92	0.01	2.87	0.00
2	-1,312	3.28	1.35	0.01	2.90	0.00
3	-950	2.45	0.84	0.00	2.93	0.00
4	-588	2.14	0.39	0.00	2.96	0.00
5	-256	2.07	0.07	0.00	2.98	0.00
R <sup>2</sup>	0.3581		0.2529		0.0845	

Gen	Family Average					
	PTA Milk <sup>3</sup>		PTA DPR <sup>4</sup>		PTA SCS <sup>6</sup>	
	Mean	SE	Mean	SE	Mean	SE
1	104	6.35	0.09	0.01	0.00	0.00
2	54	4.76	0.07	0.01	-0.01	0.00
3	35	4.26	0.04	0.01	-0.01	0.00
4	21	4.11	0.00	0.01	0.00	0.00
5	22	4.21	-0.03	0.01	0.00	0.00
R <sup>2</sup>	0.0220		0.0214		0.0209	

<sup>1</sup> Foundation families started in 1980 and 1981

<sup>2</sup> Generation effect is significant ( $P < 0.05$ ) for all traits overall

<sup>3</sup> Generation effect is significant ( $P < 0.05$ ) for traits between generations

<sup>4</sup> Generation effect is significant ( $P < 0.05$ ) for traits between generations except from 1 to 2

<sup>5</sup> Generation effect is significant ( $P < 0.05$ ) for traits between generations except from 4 to 5

<sup>6</sup> Generation effect is significant ( $P < 0.05$ ) for traits between generations except from 1 to 2, 1 to 2, 1 to 3, 1 to 4, 1 to 5, 2 to 3, 2 to 4

Table 3-6 Means across families by generation for productive life, net merit dollars, and percent inbreeding coefficient<sup>1</sup>

<b>Family Average</b>								
Gen	Months of Productive Life <sup>3</sup>		PTA Productive Life <sup>3</sup>		Net Merit Dollars <sup>3</sup>		% Inbreeding Coefficient <sup>3</sup>	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
1	44.56	0.22	-0.98	0.02	-354	1.48	0.93	0.02
2	35.57	0.11	-0.49	0.01	-278	0.77	1.47	0.01
3	32.63	0.08	-0.18	0.01	-197	0.58	2.38	0.01
4	30.92	0.07	0.01	0.01	-120	0.50	3.40	0.01
5	30.03	0.07	0.15	0.01	-51	0.49	4.25	0.01
R <sup>2</sup>	0.0635		0.0872		0.3137		0.2146	

<b>Family Difference from Herdmates</b>								
Gen	Months of Productive Life <sup>3</sup>		PTA Productive Life <sup>3</sup>		Net Merit Dollars <sup>5</sup>		% Inbreeding Coefficient <sup>4</sup>	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
1	9.04	0.22	0.47	0.02	36	1.48	0.00	0.02
2	4.02	0.15	0.29	0.01	19	1.10	-0.12	0.02
3	2.04	0.13	0.21	0.01	12	0.99	-0.10	0.01
4	0.85	0.12	0.10	0.01	6	0.95	-0.01	0.01
5	0.12	0.12	0.01	0.01	4	0.98	0.11	0.01
R <sup>2</sup>	0.0179		0.0267		0.0229		0.0151	

<sup>1</sup> Foundation families started in 1980 and 1981

<sup>2</sup> Generation effect is significant ( $P < 0.05$ ) for all traits overall

<sup>3</sup> Generation effect is significant ( $P < 0.05$ ) for traits between generations

<sup>4</sup> Generation effect is significant ( $P < 0.05$ ) for traits between generations except from 1 to 4, 2 to 3

<sup>5</sup> Generation effect is significant ( $P < 0.05$ ) for traits between generations except from 4 to 5

Table 3-7 Number of families per generation in comparison to contemporary herd mates using independent culling selection levels for milk and pregnancy rate<sup>4</sup>

<b>Family members 1<sup>st</sup> Parity</b>													
Gen	Maternal families all family members averaged						Maternal families weighted by degree of relationship						
	Above <sup>1</sup>		Mixed <sup>2</sup>		Below <sup>3</sup>		Above <sup>1</sup>		Mixed <sup>2</sup>		Below <sup>3</sup>		
	N	%	N	%	N	%	N	%	N	%	N	%	
1-2	561	11	4,546	86	165	3	1,434	27	3,221	61	619	12	
1-3	196	4	5,041	96	35	1	926	18	4,029	76	319	6	
1-4	66	1	5,196	99	10	0	551	10	4,591	87	132	3	

<b>Family members 1st and 2nd Parities</b>													
Gen	Maternal families all family members averaged						Maternal families weighted by degree of relationship						
	Above <sup>1</sup>		Mixed <sup>2</sup>		Below <sup>3</sup>		Above <sup>1</sup>		Mixed <sup>2</sup>		Below <sup>3</sup>		
	N	%	N	%	N	%	N	%	N	%	N	%	
1-2	543	10	4,614	86	187	3	1,412	26	3,332	61	670	12	
1-3	188	4	5,120	96	36	1	873	16	4,186	76	355	6	
1-4	55	1	5,279	99	10	0	504	9	4,759	87	151	3	

<b>Family members 1st through 5th Parities</b>													
Gen	Maternal families all family members averaged						Maternal families weighted by degree of relationship						
	Above <sup>1</sup>		Mixed <sup>2</sup>		Below <sup>3</sup>		Above <sup>1</sup>		Mixed <sup>2</sup>		Below <sup>3</sup>		
	N	%	N	%	N	%	N	%	N	%	N	%	
1-2	405	8	4,733	88	218	4	1,168	21	3,511	64	808	15	
1-3	103	2	5,196	97	57	1	659	12	4,373	80	455	8	
1-4	19	0	5,317	99	20	0	304	6	4,936	90	247	5	

<sup>1</sup> Above: Family generational average were above herd merit for both milk and DPR

<sup>2</sup> Mixed: Family generational average were above and below herd merit for milk and DPR

<sup>3</sup> Below: Family generational average were below herd merit for both milk and DPR

<sup>4</sup> Independent culling levels involves selecting animals which exceed a particular level for each objective

Table 3-8 Number of families per generation in comparison to contemporary herdmates using index selection for milk and pregnancy rate<sup>4</sup>

<b>Family members 1<sup>st</sup> Parity</b>												
Gen	Maternal families all family members averaged						Maternal families weighted by degree of relationship					
	Above <sup>1</sup>		Mixed <sup>2</sup>		Below <sup>3</sup>		Above <sup>1</sup>		Mixed <sup>2</sup>		Below <sup>3</sup>	
	N	%	N	%	N	%	N	%	N	%	N	%
1-2	1,750	33	2,460	<b>47</b>	1,064	20	2,761	52	465	9	2,048	39
1-3	1,057	20	3,705	70	512	10	2,250	43	1,543	29	1,481	28
1-4	634	12	4,363	83	277	5	1,836	35	2,427	46	1,011	17

<b>Family members 1st and 2nd Parities</b>												
Gen	Maternal families all family members averaged						Maternal families weighted by degree of relationship					
	Above <sup>1</sup>		Mixed <sup>2</sup>		Below <sup>3</sup>		Above <sup>1</sup>		Mixed <sup>2</sup>		Below <sup>3</sup>	
	N	%	N	%	N	%	N	%	N	%	N	%
1-2	1,788	33	2,532	47	1,091	20	2,823	52	528	10	2,060	38
1-3	1,082	20	3,796	70	533	10	2,333	43	1,604	30	1,474	27
1-4	611	11	4,521	84	279	5	1,858	34	2,531	47	1,022	19

<b>Family members 1st through 5th Parities</b>												
Gen	Maternal families all family members averaged						Maternal families weighted by degree of relationship					
	Above <sup>1</sup>		Mixed <sup>2</sup>		Below <sup>3</sup>		Above <sup>1</sup>		Mixed <sup>2</sup>		Below <sup>3</sup>	
	N	%	N	%	N	%	N	%	N	%	N	%
1-2	1,663	30	2,525	46	1,299	24	2,706	49	541	10	2,240	41
1-3	883	16	3,889	71	715	13	2,142	39	1,595	29	1,750	32
1-4	446	8	4,618	84	423	8	1,565	29	2,615	48	1,307	24

<sup>1</sup> Above: Family generational average were above herd merit when equally combining milk and DPR

<sup>2</sup> Mixed: Family generational average were above and below herd merit when combining milk and DPR

<sup>3</sup> Below: Family generational average were below herd merit when equally combining milk and DPR

<sup>4</sup> Index selection involves selecting animals according to their overall merit by weighting factors. Milk and DPR were weighted equally



Table 3-9 Comparison of daughters in 5th generation using first four generations of maternal family members weighted by degree of relationship with families grouped by independent culling levels for milk and pregnancy rate<sup>2,3,6</sup>

<b>Family members 1st Parity</b>				
<b>Trait Deviations</b>	<b>Above<sup>3</sup></b>	<b>Mixed<sup>4</sup></b>	<b>Below<sup>5</sup></b>	<b>R<sup>2</sup></b>
Milk (kg)	239	184	48	0.0103
Pregnancy Rate	0.56	0.69	-1.66	0.0064
Somatic Cell Score <sup>1</sup>	-0.02	0.05	0.05	0.0012
<b>Family members 1st through 2nd Parities</b>				
<b>Trait Deviations</b>	<b>Above<sup>3</sup></b>	<b>Mixed<sup>4</sup></b>	<b>Below<sup>5</sup></b>	<b>R<sup>2</sup></b>
Milk (kg)	143	80	6	0.0177
Pregnancy Rate	-0.72	-0.52	-1.02	0.0105
Somatic Cell Score <sup>1</sup>	0.10	0.06	0.41	0.0020
<b>Family members 1st through 5th Parities</b>				
<b>Trait Deviations</b>	<b>Above<sup>3</sup></b>	<b>Mixed<sup>4</sup></b>	<b>Below<sup>5</sup></b>	<b>R<sup>2</sup></b>
Milk (kg)	23	-28	-9	0.0185
Pregnancy Rate	-1.67	-1.55	-2.04	0.0116
Somatic Cell Score <sup>1</sup>	0.01	0.09	-0.07	0.0020

<sup>1</sup> Lower value is optimal

<sup>2</sup> Model 2 above was utilized in this analysis

<sup>3</sup> Above: The four prior family generational averages were above herd merit for both milk and DPR

<sup>4</sup> Mixed: The four prior family generational averages were above and below herd merit for milk and DPR

<sup>5</sup> Below: The four prior family generational averages were below herd merit for both milk and DPR

<sup>6</sup> Independent culling levels involves selecting animals which exceed a particular level for each objective

Table 3-10 Comparison of daughters in 5th generation using first four generations of maternal family members weighted by degree of relationship with families grouped by index selection for milk and pregnancy rate<sup>2,3,6</sup>

<b>Family members 1st Parity</b>				
<b>Trait Deviations</b>	<b>Above<sup>3</sup></b>	<b>Mixed<sup>4</sup></b>	<b>Below<sup>5</sup></b>	<b>R<sup>2</sup></b>
Milk (kg)	190	136	25	0.0151 <sup>7</sup>
Pregnancy Rate	0.53	0.74	0.83	0.0043
Somatic Cell Score <sup>1</sup>	0.11	0.07	0.05	0.0015
<b>Family members 1st through 2nd Parities</b>				
<b>Trait Deviations</b>	<b>Above<sup>3</sup></b>	<b>Mixed<sup>4</sup></b>	<b>Below<sup>5</sup></b>	<b>R<sup>2</sup></b>
Milk (kg)	76	47	-57	0.0208 <sup>7</sup>
Pregnancy Rate	-0.69	-0.41	-0.47	0.0106
Somatic Cell Score <sup>1</sup>	0.11	0.07	0.05	0.0013
<b>Family members 1st through 5th Parities</b>				
<b>Trait Deviations</b>	<b>Above<sup>3</sup></b>	<b>Mixed<sup>4</sup></b>	<b>Below<sup>5</sup></b>	<b>R<sup>2</sup></b>
Milk (kg)	-29	-58	-138	0.0225 <sup>7</sup>
Pregnancy Rate	-1.78	-1.49	-1.42	0.0111
Somatic Cell Score <sup>1</sup>	0.13	0.11	0.07	0.0017

<sup>1</sup> Lower value is optimal

<sup>2</sup> Model 2 above was utilized in this analysis

<sup>3</sup> Above: The prior family generations averages were above herd merit when equally combining milk and DPR

<sup>4</sup> Mixed: The prior family generations averages were above and below herd merit combining milk and DPR

<sup>5</sup> Below: The prior family generations averages were below herd merit when equally combining milk and DPR

<sup>6</sup> Index selection involves selecting animals according to their overall merit by weighting factors. Milk and DPR were weighted equally

<sup>7</sup> Group effect and between groups are significant ( $P < 0.05$ ) except from “middle to top”

Table 3-11 Comparison of daughters in 5th generation using first four generations of maternal family members weighted by degree of relationship with families grouped by Milk (kg) deviated from herdmates

<b>Family members 1st Parity</b>				
<b>Trait Deviations</b>	<b>Bottom 25%</b>	<b>Middle 50%</b>	<b>Top 25%</b>	<b>R<sup>2</sup></b>
Milk (kg)	-391	118	657	0.1243 <sup>2</sup>
Pregnancy Rate	1.22	0.63	-0.12	0.0064 <sup>3</sup>
Somatic Cell Score <sup>1</sup>	0.12	0.08	0.06	0.0137
<b>Family members 1st and 2nd Parities</b>				
<b>Trait Deviations</b>	<b>Bottom 25%</b>	<b>Middle 50%</b>	<b>Top 25%</b>	<b>R<sup>2</sup></b>
Milk (kg)	-402	15	511	0.1192 <sup>2</sup>
Pregnancy Rate	0.01	-0.60	-1.04	0.0109 <sup>5</sup>
Somatic Cell Score <sup>1</sup>	0.11	0.09	0.02	0.0217 <sup>3</sup>
<b>Family members 1st through 5th Parities</b>				
<b>Trait Deviations</b>	<b>Bottom 25%</b>	<b>Middle 50%</b>	<b>Top 25%</b>	<b>R<sup>2</sup></b>
Milk (kg)	-475	-72	378	0.1177 <sup>2</sup>
Pregnancy Rate	-0.99	-1.63	-2.08	0.0116 <sup>4</sup>
Somatic Cell Score <sup>1</sup>	0.14	0.11	0.06	0.0233 <sup>5</sup>

<sup>1</sup> Lower value is optimal

<sup>2</sup> Group effect is significant ( $P < 0.05$ ) for trait between groups and overall

<sup>3</sup> Group effect and between groups are significant ( $P < 0.05$ ) except from “bottom to middle”

<sup>4</sup> Group effect and between groups are significant ( $P < 0.05$ ) except from “middle to top”

<sup>5</sup> Group effect and between groups are significant ( $P < 0.05$ ) except from “bottom to middle” and “middle to top”

Table 3-12 Comparison of daughters in 5th generation using first four generations of maternal family members weighted by degree of relationship with families grouped by daughter pregnancy rate deviated from herdmates

<b>Family members 1st Parity</b>				
<b>Trait Deviations</b>	<b>Bottom 25%</b>	<b>Middle 50%</b>	<b>Top 25%</b>	<b>R<sup>2</sup></b>
Milk (kg)	216	104	100	0.0160 <sup>4</sup>
Pregnancy Rate	-3.80	0.58	4.82	0.0640 <sup>2</sup>
Somatic Cell Score <sup>1</sup>	0.11	0.08	0.07	0.0136

<b>Family members 1st and 2nd Parities</b>				
<b>Trait Deviations</b>	<b>Bottom 25%</b>	<b>Middle 50%</b>	<b>Top 25%</b>	<b>R<sup>2</sup></b>
Milk (kg)	92	49	-36	0.0212 <sup>3</sup>
Pregnancy Rate	-4.26	-0.79	3.47	0.0759 <sup>2</sup>
Somatic Cell Score <sup>1</sup>	0.13	0.07	0.04	0.0215 <sup>5</sup>

<b>Family members 1st through 5th Parities</b>				
<b>Trait Deviations</b>	<b>Bottom 25%</b>	<b>Middle 50%</b>	<b>Top 25%</b>	<b>R<sup>2</sup></b>
Milk (kg)	21	-61	-128	0.0258 <sup>2</sup>
Pregnancy Rate	-4.94	-1.81	2.14	0.0740 <sup>2</sup>
Somatic Cell Score <sup>1</sup>	0.14	0.09	0.08	0.0230 <sup>5</sup>

<sup>1</sup> Lower value is optimal

<sup>2</sup> Group effect is significant ( $P < 0.05$ ) for trait between groups and overall

<sup>3</sup> Group effect and between groups are significant ( $P < 0.05$ ) except from “bottom to middle”

<sup>4</sup> Group effect and between groups are significant ( $P < 0.05$ ) except from “middle to top”

<sup>5</sup> Group effect and between groups are significant ( $P < 0.05$ ) except from “bottom to middle” and “middle to top”

Table 3-13 Comparison of daughters in 5th generation using first four generations of maternal family members weighted by degree of relationship with families grouped by Somatic Cell Score

<b>Family members 1st Parity</b>				
<b>Trait Deviations</b>	<b>Bottom 25%</b>	<b>Middle 50%</b>	<b>Top 25%</b>	<b>R<sup>2</sup></b>
Milk (kg)	81	137	162	0.0155 <sup>5</sup>
Pregnancy Rate	-0.16	0.92	0.60	0.0061 <sup>6</sup>
Somatic Cell Score <sup>1</sup>	0.50	0.07	-0.31	0.0837 <sup>2</sup>
<b>Family members 1st and 2nd Parities</b>				
<b>Trait Deviations</b>	<b>Bottom 25%</b>	<b>Middle 50%</b>	<b>Top 25%</b>	<b>R<sup>2</sup></b>
Milk (kg)	-17	47	72	0.0208 <sup>4</sup>
Pregnancy Rate	-0.96	-0.31	-0.73	0.0103
Somatic Cell Score <sup>1</sup>	0.47	0.06	-0.28	0.1005 <sup>2</sup>
<b>Family members 1st through 5th Parities</b>				
<b>Trait Deviations</b>	<b>Bottom 25%</b>	<b>Middle 50%</b>	<b>Top 25%</b>	<b>R<sup>2</sup></b>
Milk (kg)	-87	-53	-41	0.0239
Pregnancy Rate	-2.22	-1.27	-1.60	0.0116 <sup>6</sup>
Somatic Cell Score <sup>1</sup>	0.47	0.10	-0.24	0.1014 <sup>2</sup>

<sup>1</sup> Lower value is optimal

<sup>2</sup> Group effect is significant ( $P < 0.05$ ) for trait between groups and overall

<sup>3</sup> Group effect and between groups are significant ( $P < 0.05$ ) except from “bottom to middle”

<sup>4</sup> Group effect and between groups are significant ( $P < 0.05$ ) except from “middle to top”

<sup>5</sup> Group effect and between groups are significant ( $P < 0.05$ ) except from “bottom to middle” and “middle to top”

<sup>6</sup> Group effect and between groups are significant ( $P < 0.05$ ) except from “bottom to top” and “middle to top”

Table 3-14 Comparison of daughters in 5th generation using first four generations of maternal family members weighted by degree of relationship with families grouped by PTA productive life

<b>Family members 1st Parity</b>				
<b>Trait Deviations</b>	<b>Bottom 25%</b>	<b>Middle 50%</b>	<b>Top 25%</b>	<b>R<sup>2</sup></b>
Milk (kg)	102	159	178	0.0150
Pregnancy Rate	-1.07	0.66	1.94	0.0121 <sup>2</sup>
Somatic Cell Score <sup>1</sup>	0.23	0.07	-0.04	0.0215 <sup>2</sup>
<b>Family members 1st and 2nd Parities</b>				
<b>Trait Deviations</b>	<b>Bottom 25%</b>	<b>Middle 50%</b>	<b>Top 25%</b>	<b>R<sup>2</sup></b>
Milk (kg)	-24	54	93	0.0198 <sup>3</sup>
Pregnancy Rate	-2.33	-0.52	0.96	0.0213 <sup>2</sup>
Somatic Cell Score <sup>1</sup>	0.24	0.06	-0.05	0.0322 <sup>2</sup>
<b>Family members 1st through 5th Parities</b>				
<b>Trait Deviations</b>	<b>Bottom 25%</b>	<b>Middle 50%</b>	<b>Top 25%</b>	<b>R<sup>2</sup></b>
Milk (kg)	-118	-46	-5	0.0230 <sup>2</sup>
Pregnancy Rate	-3.14	-1.62	-0.12	0.0212 <sup>2</sup>
Somatic Cell Score <sup>1</sup>	0.26	0.09	-0.01	0.0338 <sup>2</sup>

<sup>1</sup> Lower value is optimal

<sup>2</sup> Group effect is significant ( $P < 0.05$ ) for trait between groups and overall

<sup>3</sup> Group effect and between groups are significant ( $P < 0.05$ ) except from “bottom to middle”

<sup>4</sup> Group effect and between groups are significant ( $P < 0.05$ ) except from “middle to top”

<sup>5</sup> Group effect and between groups are significant ( $P < 0.05$ ) except from “bottom to middle” and “middle to top”

## GENERAL CONCLUSIONS

All data collection systems have to balance accuracy and completeness with cost to maintain that accuracy. For an informational system to be successful, benefits to management from data quality, quantity, and distribution must clearly outweigh any financial and time costs associated with equipment and labor (Wiggans, 1994). DRPCs and AIPL have achieved success in their areas of expertise.

A researcher must keep in mind the purpose and objectives of the organization when determining the best source of data for a project. The following questions should be considered before requesting dairy research data from any data provider:

- Data availability
- Data quality
- Retrospectively recomputed lactation values
- Lactation records calculated with test interval or best prediction method
- Adjustment factors of lactation data
- Recomputation of previous days open
- Consistent standardization

Effects of maternal lineage on production and fertility traits of Holstein cattle was evaluated using field data with different selection methods. Independent culling level selection methodology, requiring both milk and DPR to be above contemporary herdmates for four consecutive generations, was studied. As the number of generations increased, families with generational averages above contemporary herdmates for milk and DPR per generation decreased to 66 (1%) by the fourth generation when weighting the families equally by generation. Number of families increased to 323 (6%) when families were weighted by degree of relationship. DPR was not significant for parity group with family

weighting due to unbalanced numbers in the groupings. Milk and DPR, with equal weightings on each, were analyzed using a multi-trait selection index. Average for each family by generation was required to be above zero for four generations. Even with 1,836 (35 %) of the families maintained above contemporary herdmates for all four generations, DPR was not significant ( $P > 0.05$ ). The original hypothesis that 5th generation daughters would have above average fertility and above average production, when family averages were above contemporaries in preceding generations, was rejected.

Comparison of 5th generation daughters using the first four generations of maternal family members, either averaged or weighted by degree of relationship to daughter, by single trait selection resulted in significance of group effect ( $P < 0.05$ ). These findings provide evidence that DPR and milk production could be improved through single trait selection on maternal lineage.

Maternal families were evaluated on PTA for productive life, which includes heavier weighting for DPR, using multi-trait selection methods. Pregnancy rate was observed to be significant ( $P < 0.05$ ). These findings provide evidence that pregnancy rate and milk production may be inherited through maternal lineage. Thus, there may be potential economic value in considering maternal family history for pregnancy rate when selecting future bull dams.



## **APPENDICES**

## **APPENDIX A**

### **DATA SOURCES EDITS**

Four sources of data were used for analyses: 1) Lactation records were obtained from the Dairy Records Management Systems (DRMS) in Raleigh, NC; one of four Dairy Records Processing Centers (DRPC) in the United States; 2) Genetic information was obtained from the Animal Improvement Programs Laboratory (AIPL) of the USDA for sire data; 3) Genetic information was obtained from AIPL for cow data on the November 2006 genetic evaluation; and 4) Standardized lactation yields were obtained from AIPL.

#### **Sources of data**

##### **Data Source 1: Lactation data from DRPC**

Data obtained from Dairy Records Processing Center (DRPC) in Raleigh, NC in August 2005 included historical lactation records from 13 states dating from January 1980 to August 2005. States included Alabama, Georgia, Florida, Indiana, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, Virginia, and Vermont.

Quality of pedigree accuracy is critical in matching maternal lineage from herds obtained from DRPC. Due to the nature of the DRPC business model, additional edit corrections were required for merging with AIPL data sources and pedigree accuracy determination.

Data cleansing, a process of detecting and correcting (or removing) corrupt or inaccurate data, was completed on the DRPC by accessing the data error codes assigned by

AIPL (Norman et al., 1994). Detailed information on the error codes was available on the AIPL website. Of the 172 edits AIPL applies, 48 edits were applied to the DRPC data (Table A-1). The remaining 124 were not relevant in these analyses or they lacked the cross-reference data required to apply them. Appendix B contains the program used for data cleansing of DRPC data.

The 0-general records (subgroup C), 1-animal identification (ID) (subgroup A, B, D, F), 2-sire ID (subgroup A, B, D, E, F), 3-dam ID (Subgroup A, D, E, F), 5-birth date (A,C), 6-Calving date errors (F, H, J) were used for editing DRPC data. From these error codes, 20 were checked for occurrence in DRPC data but did not occur. All cow, dam, and sire identification (ID) fields were expanded from 9 to 12 characters for edits 1Ba, 2Ba, and 3Ba. DRPC identification fields had a length of 9 while AIPL data sources had a length of 12 characters. Identification fields were required to have a length of 12 characters in order to combine databases. Expanding the length of identification fields occurred on 3,902,239 records. For complete detail on data cleansing and record removal, see Table A-1 and Table A-2 below.

Table A-1 Data Source 1: Data cleansing of lactation data from DPRC using AIPL error codes

<b>Code</b>	<b>Error Code Description</b>	<b>Action</b>	<b>Records Affected</b>
0Cb	Species code is not 0 or 1 (code of 4 appeared in data)	Change	4,508
<b>1-Animal Identification Edits</b>			
1Aa	Animal ID is all zeroes	Change	N/A
1Ab	Animal ID is all blanks	Reject	85,101
1Ac	Animal ID has embedded blanks	Reject	4,183
1Ad	Animal ID has embedded "UNK".	Reject	8,825
1Ae	Animal ID has embedded "NEEDID", "NOID", "99PJS0104", "99PJD1002", or "000099999999"	Reject	1
1Af	Animal ID has embedded "IMPORT" and is not verified	Reject	N/A
1Ba	Animal ID has leading blanks that are changed to zeroes	Change	All
1Bc	Animal ID '00004' that is changed to '00000'. Animal country code is changed to CAN	Change	2,574
1Bd	VIP animal ID is '000 B' or '000 C'. Animal number is changed to '00000B' or '00000C'	Change	N/A
1Da	Animal breed code is invalid.(Kept breed "H" and "W")	Reject	791,525
1Dc	USA animal ID has non-numeric characters where not allowed. (Character in Field 1-5 and 9-12)	Reject	6,373
1Fb	Animal ID equals sire ID on the input record.	Reject	70
1Fc	Animal ID equals dam ID on the input record	Reject	125
<b>2 – Sire Identification</b>			
2Aa	Sire ID is all zeroes	Change	2,820,296
2Ad	Sire ID has embedded "UNK"	Change	7,944
2Ae	Sire ID has embedded "NEEDID", "NOID", "99PJS0104", or "000099999999"	Reject	N/A
2Af	Sire ID has embedded "IMPORT" and is not verified		N/A
2Ba	Sire ID has leading blanks that are changed to zeroes	Change	All
2Bb	Sire ID is '00004' that is changed to '00000' Sire country code is changed to CAN	Change	37,912
2Bc	Canadian sire ID has non-numeric @30; possible grade sire ID. Sire country code is changed to USA	Change	N/A
2Bl	Breed code of sire ID is invalid. Breed code is changed.	Reject	333

	(Removed breed A,B,G,J,M,X)		
2Dc	USA sire ID has non-numeric characters where not allowed	Change	23,479
2Dm	Invalid old-type registered sire ID and last 7 digits equal 0	Change	1,703
2Do	Old registered sire ID is '00036' for registered animal	Change	12
2Dp	Grade sire ID and last 4 digits equal 0000	Change	1,047
2Ea	CAN sire ID has non-numeric characters in (25-35)	Change	N/A
2Ed	CAN sire ID is less than '000000010000' or greater than '000018000000' and is not verified	Change	N/A
2Fb	Sire ID equals dam ID number	Reject	N/A

### 3 – Dam Identification

3Aa	Dam ID is all zeroes	Change	333
3Ad	Dam ID has embedded "UNK"	Change	N/A
3Ae	Dam ID has embedded "NEEDID", "NOID", "99PJD1002", "000099999999"	Change	N/A
3Af	Dam ID has embedded "IMPORT" and is not verified	Change	N/A
3Ba	Dam ID has leading blanks that are changed to zeroes	Change	All

### 5 – Birth date

5Ab	Birth date is invalid	Reject	3
5Ae	Birth year is less than 1950	Reject	N/A
5Ca	Birth date is not at least 548 days before first calving date	Reject	71,986

### 6 – Nontest-Day Production

6Fa	Calving date is not numeric	Reject	N/A
6Fb	Calving date is after the current date	Reject	N/A
6Fc	Calving date is before 1950	Reject	N/A
6Fe	Calving date is before the birth date	Reject	N/A
6Fg	Calving date is before 15 months (456 days) of age	Reject	N/A
6Hb	Previous days dry is > 600	Change	681
6Hc	Previous days dry is not zero and lactation number is 1. Age at calving is less than 884 days and is not verified	Change	1,461
6Jb	Conception date is before calving date plus 14 days	Change	13,493
6Jd	Conception date is after current date	Change	N/A
6Je	Conception date is after calving date plus days in milk - 1	Change	8,943
6Jf	Conception date is more than 50 days after the termination of the lactation record so it was set to missing	Change	982,357

Table A-2 Data Source 1: Removal of lactation data from DPRC using AIPL error codes

Source 1 Remaining	Removed	% of Total	
7,962,672			Lactation records from 2,472,669 Cows
7,878,251	84,421	1.06	Exact duplicate records
7,822,344	37,657	0.47	Duplicate record per lactation per cow. Keeping greatest days in milk
7,821,026	19,568	0.25	Duplicate calving date for the same cow
7,748,461	72,565	0.91	DRMS test herds (Herd codes 55999001 to 55999999)
7,663,360	85,101	1.07	1Aa- Animal ID is all zeroes
7,659,177	4,183	0.05	1Ac- Animal ID has embedded blanks
7,650,352	8,825	0.11	1Ad- Dam ID has embedded "UNK"
7,650,351	1	0.00	1Ae- Dam ID has embedded "NEEDID", "NOID", "99PJD1002", "00009999999"
6,858,830	791,525	9.94	1Da- Animal breed code is invalid (Kept breed "H" and "W")
6,852,457	6,373	0.08	1Dc- USA animal ID has non-numeric characters where not allowed.
6,852,387	70	0.00	1Fb- Animal ID equals sire ID
6,852,262	125	0.00	1Fc- Animal ID equals dam ID
6,851,929	333	0.00	2B1- Breed code of sire ID is invalid (Removed breed A,B,G,J,M,X)
6,851,766	163	0.00	2Fb- Sire ID equals dam ID number
6,851,763	3	0.00	5Ab- Birth date is invalid
6,779,786	71,986	0.90	5Ca- Birth date is not at least 548 days before first calving date
		<b>86.07%</b>	<b>Total Percent remaining after edits</b>

## Data Source 2: Genetic data for bulls from AIPL

A file containing genetic information on all domestic and foreign bulls in November 2006 evaluation was acquired from AIPL. The bull evaluation file included pedigree, cross-reference, and genetic evaluation information. The initial data file included 266,438 bulls. Non-Holstein breed animals, making up 20.5% of records, were removed. Bulls with country codes from Canada, Germany, Netherlands, and USA were retained in the data file. Duplicate bull identification numbers between countries did exist in the initial bull evaluation file. Bulls from countries not listed above accounted for an additional 16.2% of bulls and they were removed. Source 1 (DRPC data) did not contain a country code for sire. In order to utilize Source 1, only one sire with the same identification could be used. If multiple bulls occurred in the bull file, all occurrences of a duplicate sire would match with individual cow records resulting in duplication in the cow file. Therefore, bulls were eliminated due to duplication of bull identification across countries. Edited data source included 168,745 bulls making up 63.3% of the original bull data file.

See Table A-3 below for summary of edits that were performed.

Table A-3 Data Source 2: Data cleansing of bull genetic evaluation for yield traits

<b>Remaining</b>	<b>Removed</b>	<b>% of Total</b>	
266,438			Number of Sires
211,834	54,604	20.5	Non-Holstein Bulls
168,745	43,089	16.2	Sire Country: Canada, Germany, Netherlands, USA
		<b>63.33</b>	<b>Total Percent remaining after edits</b>

### Data Source 3: Genetic data for cows from AIPL

All cows resulting from data Source 1 (DRPC data) were sent to AIPL merge with genetic information stored at AIPL for cows. Genetic information for cow data was from the November 2006 genetic evaluation. A total of 2,773,097 cows (not listed in table below) were transmitted to AIPL in the initial upload and 1,303,918 cows were transmitted in the received download. Edits included removal of duplicates and non-Holstein animals with 99.2 percent of records remaining after edits (Table A-4). AIPL data included a variable for first herd the cow calved for her first lactation and a field for the last herd the cow housed in. , First and last herd variables were identical for a cow that was owned by the same producer for her entire life. No herd identification was on file in Source 3 for 134,934 (10.4%) cows of the 1,303,918. Cows lacking herd identification were not initially removed during edits, since the returned file should include animals in the initial upload file.

Table A-4 Data Source 3: Data cleansing of cow genetic evaluation for yield traits

<b>Remaining</b>	<b>Removed</b>	<b>% of Total</b>	
1,303,918			Number of Cows
1,295,992	7,926	0.6	Non-Holstein Cows
1,292,810	3,182	0.2	Exact duplicate records
		<b>99.2</b>	<b>Total Percent</b>



## Data Source 4: Lactation data from AIPL

From data Source 1 (DRPC), 7,840,594 lactation records on 2,773,097 cows were sent to AIPL for a match of the cows AIPL calculated, standardized lactation yields. With duplicates removed (1.2%), the returned download included 3,461,552 lactation records on 1,440,577 cows. No additional quality edits were required for Source 4 received from AIPL since data cleansing occurred prior to receiving the data (Norman et al., 1994).

Table A-5 Data Source 4: Data cleansing of lactation data from AIPL

Remaining	Removed	% of Total	
3,461,552			Number of lactations on 1,440,577 Cows
3,420,771	40,781	1.2	Exact Duplicates
		<b>98.8</b>	<b>Total Percent remaining after edits</b>

Daughter pregnancy rate (DPR) edited by AIPL staff prior to receiving Source 4. Pregnancy rate measures how quickly a cow becomes pregnant. Pregnancy rate was defined as percentage of cows that became pregnant during each 21-d period, since each estrus cycle represents one chance for a cow to become pregnant (VanRaden et al., 2004). A Days open was transformed to pregnancy rate using the following simple linear function:

$$\text{Pregnancy Rate} = 0.25 \times (233 - \text{Days Open})$$

AIPL calculates pregnancy rate slightly differently than DRPCs and reproductive specialists. When calculating pregnancy rate, AIPL considers parities complete at 250 days

in milk (DIM). Cows not pregnant by 250 DIM are assigned 250 days open (VanRaden et al., 2004). AIPL researchers investigated sensitivity to the upper limit of DIM in estimating heritability from a range of days from 150 to 305 days open and concluded that 250 DIM was optimal.

## **Combination of Data Sources**

### **Combination 1: Source 3 x Source 4**

The first combination was performed on Source 3 (genetic cow) and Source 4 (AIPL lactation) acquired from AIPL. Since 10.4% of data had missing herd identification in data Source 4, a match with another AIPL data source was utilized to determine herd-of-origin for cows with missing herd identification. A total of 3,089,675 (90.3%) of Source 3 matched with Source 4. The 331,096 (9.7%) non-matching data from Combination 1, referred to as Remaining 1, were utilized for future matching.

Table A-6 Combination 1: Record match of combining data Source 3 with Source 4

Source 4 Remaining	Match Number	% of Total	
3,420,771			Starting record count
730,866	2,689,905	78.6	ME: Herd-Cow ID F105: Last Herd- Cow ID
601,929	128,937	3.8	ME: Herd-Cow ID F105: First Herd- Cow ID
539,724	62,205	1.8	ME: Herd-Cow ID F105: Last Herd- Cow Dual ID
536,260	3,464	0.1	ME: Herd-Cow ID F105: First Herd- Cow Dual ID
331,309	204,951	8.9	ME: Cow ID (herd value null) F105: Cow ID
331,096 <sup>a</sup>	213	0.0	ME: Cow ID (herd value null) F105: Cow ID- Cow Dual ID
	<b>3,089,675</b>	<b>90.3</b>	<b>Total Match Percentage</b>

<sup>a</sup> Remain 1 made up of 331,096 records from Source 4 not matching Source 3

### Combination 2: (Combination 1 + Remain 1) x Source 1

Second, Combination 1 built in the section above, was merged with Source 1 (DRPC lactation). A total of 3,403,445 records, 98.2% of data from Combination 1, matched with data from Source 1. A total of 121,360 records from Source 4 (AIPL lactation) that did not match Source 1 (DRPC data) or Source 3 (genetic cow) were deleted. Database Remain 2, composed of 3,421,236 (49.5%) records not matching Combination 1 from Source 1, was made up of DRPC records not matching any data received from AIPL.

Table A-7 Combination 2: Record match of combining Combination 1 and Remaining 1 with Source 1

<b>Combination 1 Remaining</b>	<b>Source 1 Remaining</b>	<b>Records Matched</b>	
3,420,771	6,779,786		Starting record count
138,686	3,497,701	3,282,085	C1: Herd ME-Cow ID ME-Calve Date ME S1: Herd-Cow ID-Calve Date
121,360			Remove Source 4 (not matching Source 3)
50,386	3,426,773	70,974	C1: Herd Last-Cow ID Last-Calve Date ME S1: Herd-Cow ID-Calve Date
48,786	3,425,175	1,600	C1: Herd Last-Cow ID Duel-Calve Date ME S1: Herd-Cow ID-Calve Date
44,939	3,421,328	3,847	C1: Herd First-Cow ID Last-Calve Date ME S1: Herd-Cow ID-Calve Date
44,847	3,421,236 <sup>a</sup>	92	C1: Herd First-Cow ID Duel-Calve Date ME S1: Herd-Cow ID-Calve Date
<b>3,403,445</b>	<b>3,358,550</b>		<b>Total Match</b>
<b>98.18</b>	<b>49.5</b>		<b>Percent of total</b>

<sup>a</sup> Remain 2 made up of 3,421,236 records from Source 1 not matching Combination 1

### Combination 3: Remain 2 x Source 3

Remain 2, non-matching observations with Combination 2 database assembled above, were matched with Source 3 (genetic cows) in assembling Combination 3a. Combination 3 was made up of: Combination 3a (7.7%), Combination 2 (49.9%), and Remain 3 (42.5%). Table A-9 reports the makeup of the comprehensive database, 6,825,172 lactation records, used to determine maternal lineage

Table A-8 Combination 3a: Record match of combining Remain 2 with Source 3

<b>Remain 2 Remaining</b>	<b>Records Matched</b>	
3,421,236 a		Starting record count
3,059,992	361,735	S1: Herd-Cow ID-Birth date F105: Last Herd- Cow ID-Birth date
3,048,160	11,832	S1: Herd-Cow ID-Birth date F105: First Herd- Cow ID-Birth date
3,039,404	8,756	S1: Herd-Cow ID-Birth date F105: Last Herd- Cow Dual ID-Birth date
3,039,144	260	S1: Herd-Cow ID-Birth date F105: First Herd- Cow Dual ID-Birth date
2,898,251	140,893	S1: Cow ID-Birth date-Dam ID F105: Cow ID-Birth date-Dam ID
2,897,438 b	813	S1: Cow ID-Birth date-Dam ID F105: Cow Dual ID-Birth date-Dam ID
	524,289	Total Match Percentage
	15.32	

<sup>a</sup> Remain 2 made up of 3,421,236 records from Source 1 not matching Combination 1

<sup>b</sup> Remain 3 made up of 2,897,438 records from Source 3 not matching Remain 2

Table A-9 Combination 3: Record counts from combining source of data

<b>Source Counts</b>	<b>% of Total</b>	
524,289	7.7	Table A-8: Database Combination 3a
3,403,445	49.9	Table A-7: Database Combination 2
2,897,438	42.5	Table A-8: Database Remain 3
6,825,172		Total count in Combination 3

### **Combination 4: Combination 3 x Source 2**

Combination 3, documented above and containing all data sources for cows, was combined with Source 2 (genetic bull) to make Combination 4. Database Combination 3 included a variable of sire identification from AIPL and another variable of sire identification from DRPC. For records having AIPL and DRPC information, sire identification from AIPL was utilized as sire on record and used in combining with Source 2. For 42.5% of records from DRPC not matching any AIPL data source, sire identification from DRPC was utilized as sire on record. Sire genetic information was matched on 51.3% of cows in database Combination 3. A temporary table of 2,659,789 cows, made up of 1,363,243 (51.3%) cows with sire genetic information and 1,296,546 (48.8%) cows without, was used to determine percent of sire-identified cows by herd. A sire identification requirement of 70% or greater was required for herds. This edit removed 1,391 herds with 1,589,221 cows from the analysis. Combination 4 database included 1,132 herds and 1,070,568 (2,659,789 minus 1,589,221) cows.

Table A-10 Combination 4: Record match of combining database Combination 3 with Source 2

<b>Combination 3 Remaining</b>	<b>Match Number</b>	<b>% of Total</b>	
6,825,172			Starting record count
2,659,789			Record count on a per cow basis
1,297,484	1,362,305	51.2%	C4: Sire ID S2: Sire ID
1,296,546	938	0.04%	C4: Sire ID S2: Sire Dual ID
48.75%	1,363,243	51.3%	Total Match Percentage

### **Combination 5: Combination 4 (Dam) x Combination 4 (Daughter)**

Family lineage was assembled by matching dam information with daughter information. A temporary data set containing all parities of cows in database Combination 4 was used to match calving dates with birthdates of daughter. A single record per cow from Combination 4 made up daughter table. Of cows from herds having 70% sire identification, 728,715 (68.1%) cows had dam records identified. Cows without dam identification remained in database Combination 5 because foundation animals in the first generation were not required to have dam information identified.

Table A-11 Combination 5: Record match of dam with daughter with Combination 4 (dam)

<b>Daughter Remaining</b>	<b>Match Number</b>	<b>% of Total</b>	
1,070,568			Starting record count (Number of cows)
366,883	704,086	65.8%	Cow: Herd-Dam ID-Birth Dam: Herd-Cow Dual ID-Calve ( $\pm 30$ days)
363,812	3,071	0.3%	Cow: Herd-Dam ID-Birth Dam: Herd-Cow ID-Calve ( $\pm 30$ days)
353,644	10,172	0.9%	Cow: Herd-Dam AIPL ID-Birth Dam: Herd-Cow AIPL Dual ID-Calve ( $\pm 30$ days)
348,638	5,128	0.5%	Cow: First Herd-Dam AIPL ID-Birth Dam: First Herd-Cow AIPL Dual ID-Calve ( $\pm 30$ days)
348,128	512	0.1%	Cow: First Herd-Dam AIPL ID-Birth Dam: Last Herd-Cow AIPL Dual ID-Calve ( $\pm 30$ days)
342,422	5,746	0.5%	Cow: Herd-Cow ID-Birth Dam: Herd-Calf ID-Calve ( $\pm 30$ days)
31.93%	728,715	68.1%	Total Match Percentage

## Data Quality Check

Figures 1 through 3 reflect changes in standardized values for milk, DPR, and Somatic Cell Score (SCS) over time. These figures include three categories: the United States Holstein population on DHIA test, all cows on DHIA in the research region of study, and cows qualifying to be part of the research study. The 13 states used in this analysis



made up 14% of cows on official (non-owner sampler) DHIA testing plans and 15% of cows on management plan (owner-sampler) type testing plans (Wiggans, 2000). The subset of states had 767 kg less milk for official DHIA testing plans than US average of 9,567 kg. A similar result occurred for cows enrolled on management testing plans. Cows on these plans had 694 kg less milk production than the United States average (Wiggans, 2000).

## References

- Norman, H. D., L. G. Waite, G. R. Wiggans, and L. M. Walton.** 1994. Improving accuracy of the United States genetics database with a new editing system for dairy records. *J. Dairy Sci.* 77:3198-3208.
- VanRaden, P. M., A. H. Sanders, M. E. Tooker, R. H. Miller, H. D. Norman, M. T. Kuhn, and G. R. Wiggans.** 2004. Development of a national genetic evaluation for cow fertility. *J. Dairy Sci.* 87:2285-2292.
- Wiggans, G. R.,** 2000. USDA Summary of 2000 Herd Averages DHI Report K-3. Online: <http://aipl.arsusda.gov/publish/dhi/dhi01/k3.shtml>. Accessed: January 9, 2008.

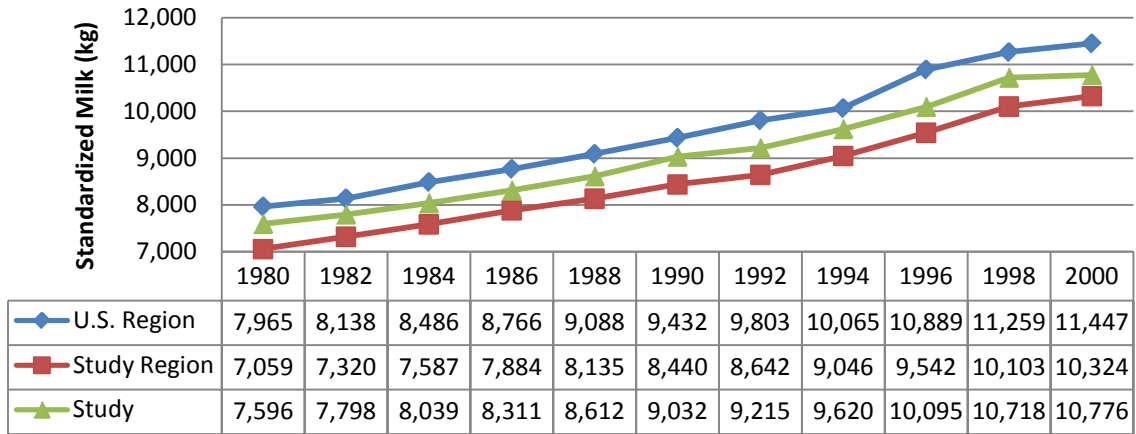


Figure A-1 United States Holstein population on DHIA test, cows in the research region of study, and cows in research study for standardized milk (kg)

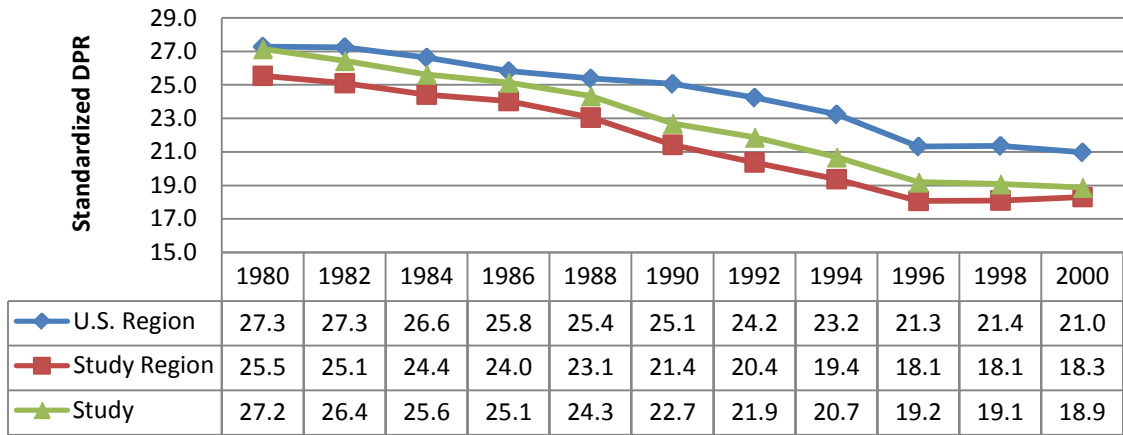


Figure A-2 United States Holstein population on DHIA test, cows in the research region of study, and cows in research study for daughter pregnancy rating

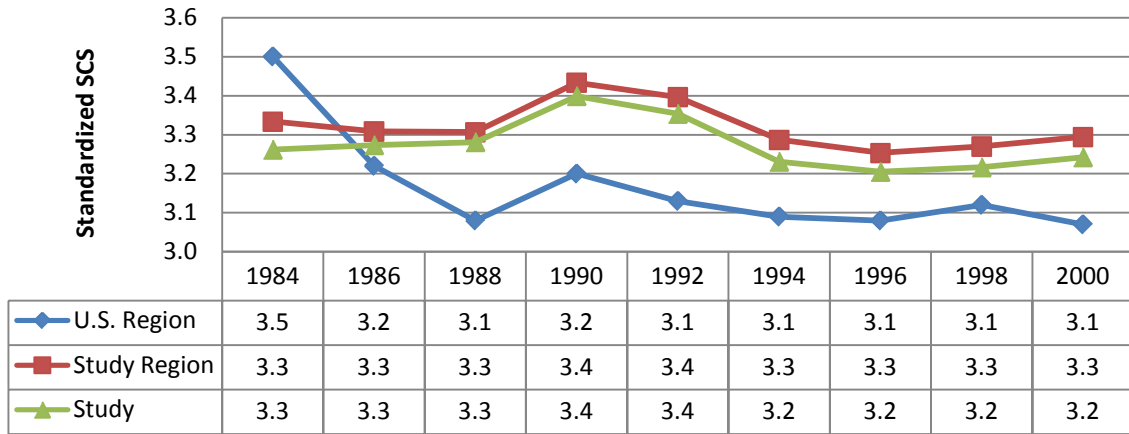


Figure A-3 United States Holstein population on DHIA test, cows in the research region of study, and cows in research study for somatic cell score.

## APPENDIX B

### DRMS DATA CLEANSING WITH AIPL ERROR CODES

Due to the nature of the DRPC business model, additional edit corrections may be required for DRMS data before using it for research purposes. Data cleansing, a process of detecting and correcting (or removing) corrupt or inaccurate data, was completed on the DRMS data by accessing the data error codes available on the AIPL website. Of the 172 AIPL edit error checks that AIPL applies, 48 applied to the DRMS data as shown in the below program. The remaining 124 were not applied due either to irrelevance of the edit in this analysis or to lack of cross-referencing data required for applying appropriate edit.

```
/******  
/*Editing of DRMS data file format 35 for purpose of  
*/  
/*including the edit standards of AIPL  
*/  
/*  
*/  
/*http://aipl.arsusda.gov/formats/geterr.cfm?ecode_in=0A#  
*/  
/*  
*/  
/* Author: Crystal Vierhout  
*/  
/* Date: 12/15/2007  
*/  
/******  
  
/*Remove DRMS test herds */  
data pass_test l_remove.drpc_test;  
  set l_data.lact2;  
  if (herdcode gt 55999001 and herdcode lt 55999999)  
    then output l_remove.drpc_test;  
    else output pass_test;  
  
/******  
/*1A-Obvious Unknown Animal Errors*/
```

```

data pass_la
    l_remove.edit_lAa
    l_remove.edit_lAb
    l_remove.edit_lAc
    l_remove.edit_lAd
    l_remove.edit_lAe
    l_remove.edit_lAf;
set pass_test;

/*1Aa-Animal identification number is all zero */
    if index(cowreg,'000000000') then output l_remove.edit_lAa;

*1Ab-animal identification number is all blanks*/
    else if index(cowreg,'          ') then output l_remove.edit_lAb;

/*1Ac-Animal identification number has embedded blanks */
    else if index(cowreg,' ') then output l_remove.edit_lAc;

/*1Ad-Animal identification number has embedded "UNK" */
    else if index(cowreg,'UNK') then output l_remove.edit_lAd;

/*1Ae-Animal identification number has embedded "NEEDID", "NOID",
"99PJS0104", "99PJD1002", or "000099999999" */
    else if index(cowreg,'NEEDID') then output l_remove.edit_lAe;
    else if index(cowreg,'NOID') then output l_remove.edit_lAe;
    else if index(cowreg,'99999999') then output l_remove.edit_lAe;

/*1Af-Animal identification number has embedded "IMPORT" */
    else if index(cowreg,'IMPORT') then output l_remove.edit_lAf;
    else output pass_la;

/*****/
/*1B-Obvious Misidentification Errors */;
data pass_lb_temp;
    set pass_la;
    length sirereg1 damreg1 cowreg1 $12.;
/*1Ba-Animal ID number has leading blanks that are changed to zeroes */
    sirereg1=trim('000'!!sirereg);
    damreg1=trim('000'!!damreg);
    cowreg1=trim('000'!!cowreg);
drop cowreg damreg sirereg;

data pass_lb;
    set pass_lb_temp;
    length sirereg damreg cowreg $12.;
    cowreg=tranwrd(cowreg1,' ','0');
    damreg=tranwrd(damreg1,' ','0');
    sirereg=tranwrd(sirereg1,' ','0');
    drop cowreg1 damreg1 sirereg1;

/*Animal ID number is '00004' that is changed to '00000' */
    if substr(cowreg,1,5)='00004'

```

```

        then do ;
            cowreg=tranwrd(cowreg, "00004", "00000");
            cow_cty="CAN";
        end;

/*****/
/*1D-Invalid Format Errors */;
data pass_1d
    l_remove.edit_1Da
    l_remove.edit_1Dc;
    set pass_lb;
    /*1Da-Animal breed code is invalid */
    if cowbrd not in ('H','W') then output l_remove.edit_1Da;

    /*1Dc-USA animal ID number has non-numeric where not allowed */
    else if indexc(substr(cowreg,1,5),'ABCDEFGHIJKLMNOPQRSTUVWXYZ') > 0
then output l_remove.edit_1Dc;
    else if indexc(substr(cowreg,9,4),'ABCDEFGHIJKLMNOPQRSTUVWXYZ') > 0
then output l_remove.edit_1Dc;
    else output pass_1d;
run;
/*****/
/*1F-Miscellaneous Errors */;
data pass_1f
    l_remove.edit_1Fb
    l_remove.edit_1Fc;
    set pass_ld;
/*1Fb-Animal ID number equals sire ID number on the input record */
    if cowreg=sirereg then output l_remove.edit_1Fb;

    /*1Fc-Animal ID number equals dam ID number on the input record */
    else if cowreg=damreg then output l_remove.edit_1Fc;
    else output pass_1f;
run;
/*****/
/*2A-Obvious Unknown Sire Errors (changed to blanks)*/;
data pass_2a;
    set pass_1f;
    /*2Aa-Sire identification number is all zeroes*/
    if sirereg='000000000000' then sirereg=' ';

    /*2Ad-Sire identification number has embedded "UNK"*/
    if index(sirereg,'UNK') then sirereg=' ';

/*2Ae-Sire ID has "NEEDID", "NOID", "99PJS0104", "000099999999"*/
    if index(sirereg,'NEEDID') then sirereg=' ';
    if index(sirereg,'NOID') then sirereg=' ';
    if index(sirereg,'000099999999') then sirereg=' ';

    /*2Af-Sire ID number has embedded "IMPORT" and is not verified*/
    if index(sirereg,'IMPORT') then sirereg=' ';
/*****/

```



```

/*2B-Obvious Unknown Sire Errors (changed to blanks)*/
data pass_2b
  l_remove.edit_2B1;
  set pass_2a;

  /*2Bb-Sire ID is '00004' (@25-29) that is changed to '00000'. */
  if substr(sirereg,1,5)='00004'
    then do ;
      sirereg=tranwrd(sirereg, "00004", "00000");
      cow_cty="CAN";
    end;

  /*2B1-Breed code of sire ID number is invalid. Breed code is
changed */
  if sirebrd in ('A','B','G','J','M','X') then output
l_remove.edit_2B1;
  else output pass_2b;
run;
/*****/
/*2D-Invalid Format Errors (changed to blanks)*/
data pass_2d
  l_remove.edit_2Fb;
  set pass_2b;
  /*2Dc-Reject USA sire ID number has non-numeric where not allowed*/
  if index(substr(sirereg,1,5),'ABCDEFGHIJKLMNOPQRSTUVWXYZ') > 0
then sirereg=' ';
  if index(substr(sirereg,9,4),'ABCDEFGHIJKLMNOPQRSTUVWXYZ') > 0
then sirereg=' ';

/*2Dm-Reject Invalid old-type registered sire ID and last 7 digits equal
0*/
  if substr(sirereg,6,7)='0000000' then sirereg=' ';

  /*2Do-Reject Old type reg sire ID is '00036' for registered
animal*/
  if substr(sirereg,1,5)='00036' then sirereg=' ';

  /*2Dp-Reject Grade sire identification and last 4 digits equal
0000*/
  if substr(sirereg,9,4)='0000' then sirereg=' ';

/*2Ed-Reject CAN sire ID is lt '000000010000' or gt '000018000000'*/
  if cow_cty="CAN"
    then do;
      if sirereg lt '000000010000' then sirereg=' ';
      if sirereg gt '000018000000' then sirereg=' ';
    end;
  /*2Fb-Sire ID equals dam ID number*/
  if sirereg ne " "
    then do;
      IF sirereg=damreg then output l_remove.edit_2Fb;
      else output pass_2d;
    end;

```

```

                end;
                else output pass_2d;
run;
/*****
/*2F-Miscellaneous Errors*/
data pass_2f
    l_remove.edit_2Fb;
    set pass_2d;

    /*2Fb-Sire ID number equals dam ID number*/
    if sirereg=damreg then output l_remove.edit_2Fb;
    else output pass_2f;
/*****
/*3A-Obvious Unknown Dam Errors (Changed to blanks)*/
data pass_3a;
    set pass_2f;

    /*3Aa-Dam identification number is all zeroes*/
    if index(damreg,'000000000') then damreg=' ';

    /*3Ad-Dam identification number has embedded "UNK"*/
    if index(damreg,'UNK') then damreg=' ';

    /*3Ae-Dam Id has "NEEDID", "NOID", "99PJD1002", "00099999999"*/
    if index(damreg,'NEEDID') then damreg=' ';
    if index(damreg,'NOID') then damreg=' ';
    if index(cowreg,'000099999999') then damreg=' ';

    /*3Af-Dam IDn number has embedded "IMPORT" and is not verified*/
    if index(cowreg,'IMPORT') then damreg=' ';
/*****
/*5A-Animal's First Calving Discrepancy*/
data pass_5a
    l_remove.edit_5ab
    l_remove.edit_5ca;
    set pass_3a;
    /*5Ab-Birth date is invalid*/
    if birthdte = . then output l_remove.edit_5ab;

    /*5Ca-Birth date is not at least 548 d before first calving date */
    else if calvdte-birthdte lt 457 then output l_remove.edit_5ca;
    else output pass_5a;
run;

data pass_6f
    l_remove.edit_6fb
    l_remove.edit_6fc
    l_remove.edit_6fe;
    set pass_5a;
    /*6Hb-Previous days dry is > 600 */
    if pdysdry > 600 then pdysdry=0;

```

```

/*6Hc-Previous days dry is not zero and lactatio number is 1 */
if lactno=1
  then do;
    if pdysdry gt 0
      then do;
        if calvdte-birthdte lt 884 then pdysdry=0;
      end;
    end;
  end;

/*6Jb-Conception date is before calving date plus 14 d */
if breeddte ne . then
  do;
    if breeddte-calvdte lt 14 then breeddte=.;
  end;
/*6Je-Conception date is after calving date + DIM - 1 for RIP input
*/
if calvdte+dim-1 lt breeddte then breeddte=.;

/*6Jf-Conception date is more than 50 d after the termination of
the lactation record so it was set to missing */
if status=5
  then do;
    if breeddte ne .
      then do;
        if statdte ne .
          then do;
            if statdte+50 gt breeddte then output
pass_6f;
          end;
        end;
      end;
    end;
  end;

/*6Fb-Calving date is after the current date*/
if calvdte - today() gt 0 then output l_remove.edit_6fb;

/*6Fc-Calving date is before 1950*/
else if year(calvdte) lt 1950 then output l_remove.edit_6fc;

/*6Fe-Calving date is before the birth date*/
else if calvdte lt birthdte then output l_remove.edit_6fe;
else output pass_6f;
run;

proc sort data=pass_5a out=l_data.Lact3 dupout=l_remove.lact3_dup
nodupkey; by herdcode cowreg birthdte calvdte;
run;

```

## APPENDIX C

### MODEL SELECTION

Proposed model:  $Y_{ijklm} = \mu + \text{HERD}_i + \text{FAM}_j + \text{PTAS}_k + \text{PTAM}_l + \varepsilon_{ijkl}$

where,

$Y_{ijklm}$	=	5 <sup>th</sup> Generation yield deviation for milk or pregnancy rate
$\mu$	=	Overall mean
$\text{Herd}_i$	=	Effect of $i_{\text{th}}$ herd
$\text{Fam}_j$	=	Effect of $j_{\text{th}}$ family quartile for milk or pregnancy rate
$\text{PTAS}_k$	=	Sire PTA for milk or daughter pregnancy rating
$\text{PTAM}_L$	=	MGS PTA for milk or daughter pregnancy rating
$\varepsilon_{ijkl}$	=	Error

Dummy variables for classification variables herd and family quartile ranking were added to the model. The Pearson correlation coefficient concluded no strong correlations between independent variables for milk or DPR as indicated by correlation statistics close to 1 or -1. The correlation coefficient of dependent variable, (5<sup>th</sup> Generation yield deviation) was analyzed with the dependent variables. The p-value for the correlation of 5<sup>th</sup> Generation yield deviation and herd were not significant for milk ( $p = 0.1367$ ) or DPR ( $p = 0.0534$ ) indicating that the population correlation coefficient was not significantly different from zero. Remaining independent variables were significantly ( $P < .05$ ) correlated with dependent variables.

Multicollinearity refers to the fact that some independent variables provide redundant information (Bowerman et al., 1986). With the ability to inflate standard errors of parameter estimates and standard errors of predictions, multicollinearity can hinder the model's ability to predict significance of independent variables. Multicollinearity of the model was evaluated by using the correlation statistics (proc corr) available through (SAS Institute Inc., 2004). The regression procedure with options VIF, COLLIN, COLLINOINT was used to examine variance inflation factors, conditional index values, and proportion of variation explained by principal components (SAS Institute Inc., 2004).

Variance inflation factors of parameter estimates did not yield indicators greater than 1.5 for milk or DPR. The largest conditional index for milk was 5.2435 and DPR of 8.2867, which did not indicate strong collinearity. Therefore, further diagnostics were performed.

Removal of the intercept was considered to avoid biasing of variables in the process, since the data could support a zero intercept. The decision was made not to remove the intercept for the following reasons: 1) the analysis was using a small subset of the herd with average yield deviation of the animal in the subset not equaling zero, and 2) class variables were utilized in the general linear model analysis with an intercept being implicitly fit with the no intercept option in the model.

The optimal model, which excluded variable herd for milk and included variable herd for DPR, resulted from Stepwise regression analysis. The proposed model was utilized with the DPR analysis in the process of using general linear model. The variable

herd was not significant ( $P = 0.0634$ ) in the DPR model and was dropped from any further use.

In an effort to achieve unbiased estimates, family groups were treated as a random effect. Boettcher et al. (1996) reported treating maternal lineage as a random effect rather than a fixed effect when doing so resulted in a statistically, significantly superior model.

$$\text{Model: } Y_{ijkl} = \mu + \text{FAM}_i + \text{PTAS}_j + \text{PTAM}_k + \varepsilon_{ijk} \quad [1]$$

where,

$Y_{ijkl}$	=	5 <sup>th</sup> Generation yield deviation for milk or pregnancy rate
$\mu$	=	Overall mean
$\text{Herd}_i$	=	Effect of $i_{\text{th}}$ herd
$\text{Fam}_i$	=	Effect of $j_{\text{th}}$ family quartile for milk or pregnancy rate
$\text{PTAS}_j$	=	Sire PTA for milk or daughter pregnancy rating
$\text{PTAM}_k$	=	MGS PTA for milk or daughter pregnancy rating
$e_{ijkl}$	=	Error

## References

- Boettcher, P. J., M. T. Kuhn, and A. E. Freeman.** 1996. Impacts of cytoplasmic inheritance on genetic evaluations. *J. Dairy Sci.* 79:663-675.
- Bowerman, B. L., R. T. O'Connell, and D. A. Dickey.** 1986. Linear statistical models : An applied approach.
- SAS Institute Inc.** 2004. SAS/STAT® 9.1 User's guide.7 volumes:5180.