The use of selection to improve sow longevity

Caitlyn Elizabeth Abell

Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

Part of the Agriculture Commons, and the Animal Sciences Commons

Recommended Citation

The use of selection to improve sow longevity

by

Caitlyn Elizabeth GeneAnn Hoots Abell

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Co-Majors: Animal Breeding and Genetics (Quantitative Genetics); Statistics

Program of Study Committee:
Kenneth Stalder, Co-Major Professor
Philip Dixon, Co-Major Professor
Max Rothschild
Jack Dekkers
Rohan Fernando
Jarad Niemi

Iowa State University
Ames, Iowa
2013

Copyright © Caitlyn Elizabeth GeneAnn Hoots Abell, 2013. All rights reserved.
# TABLE OF CONTENTS

LIST OF FIGURES iii

LIST OF TABLES iv

ACKNOWLEDGMENTS v

ABSTRACT vi

CHAPTER 1: GENERAL INTRODUCTION 1

CHAPTER 2: LITERATURE REVIEW 6

CHAPTER 3: USING CLASSIFICATION TREES TO DETECT INDUCED SOW LAMENESS WITH A TRANSIENT MODEL 30
  Abstract 30
  Implications 31
  Introduction 32
  Materials and Methods 34
  Results 38
  Discussion 41
  References 45

CHAPTER 4: GENETIC PARAMETER ESTIMATES AND RELATIVE ECONOMIC VALUES FOR PUREBRED AND CROSSBRED SOW LIFETIME PRODUCTIVITY 55
  Summary 55
  Introduction 56
  Materials and Methods 58
  Results 62
  Discussion 64
  References 68

CHAPTER 5: TOTAL COST ESTIMATION FOR IMPLEMENTING GENOME-ENABLED SELECTION IN A MULTI-LEVEL SWINE PRODUCTION SYSTEM 75
  Abstract 75
  Background 76
  Methods 78
  Results and Discussion 82
  Conclusions 91
  References 91

CHAPTER 6: GENERAL DISCUSSION AND CONCLUSIONS 97

REFERENCES 104
LIST OF FIGURES

**Figure 3.1a-b.** Classification tree to detect lameness at day 1 post-injection using measurements collected from the microcomputer-based embedded force plate system to detect induced sow lameness. 51

**Figure 3.2a-b.** Classification tree to detect lameness at day 1 post-injection using measurements collected from the GaitFour walkway system to detect induced sow lameness. 52

**Figure 5.1.** Value of expected improvement in genetic merit at commercial level. 94

**Figure 5.2.** Feasible region for profitability when incorporating genome-enabled selection in a terminal line selection program. 95

**Figure 5.3.** Feasible region for profitability when incorporating genome-enabled selection in a maternal line selection program. 96
LIST OF TABLES

Table 3.1. Subjective visual scores and objective weight distribution measures least squares means (standard error) from a study where lameness was induced in the sow’s rear feet. 48

Table 3.2. Subjective visual scores and objective weight distribution measures least squares means (standard error) from a study where lameness was induced in the sow’s front feet. 49

Table 3.3. Subjective visual scores and objective walking gait measures least squares means (standard error) from a study where lameness was induced in the sow’s feet. 50

Table 3.4. Classification tree error rates and mean decrease in accuracy for each weight distribution variable used to detect induced sow lameness. 53

Table 3.5. Classification tree error rates and mean decrease in accuracy for each walking gait measurement used to detect induced sow lameness. 54

Table 4.1. Removal parity distribution for crossbred and purebred sows in a study to evaluate the relationship between purebred and crossbred sow performance. 70

Table 4.2. Descriptive statistics for purebred and crossbred cumulative born alive traits in a study to evaluate the relationship between purebred and crossbred sow performance. 71

Table 4.3. Frequentist and Bayesian heritability estimates (±SE) for purebred and crossbred sow longevity and cumulative born alive traits using an animal model. 72

Table 4.4. Genetic correlation estimates (SE) between and crossbred sow longevity and cumulative born alive traits and a purebred longevity trait using an animal model. 73

Table 4.5. Relative emphasis of economically important swine traits based on genetic standard deviation and economic value per trait unit. 74
ACKNOWLEDGMENTS

First and foremost, I would like to thank my major professor, Dr. Kenneth Stalder. He has provided me with many opportunities and experiences to expand my knowledge and understanding of the swine breeding industry. Dr. Stalder has also been a great mentor and helped me to reach my full potential. I am truly grateful for the all the time and effort he has put into ensuring that I am fully prepared for my future goals.

My other committee members, Dr. Philip Dixon, Dr. Max Rothschild, Dr. Jack Dekkers, Dr. Rohan Fernando, and Dr. Jarad Niemi have been beneficial in ensuring my dissertation is satisfactory by offering their comments and suggestions. Dr. Rothschild has been especially helpful in contributing ideas to improve the projects in this dissertation. Dr. Fernando and Dr. Dekkers have given me a lot of guidance regarding the methods used for these projects. I would also like to thank the faculty, staff, and students of the Animal Breeding Genetics Group as well as the Department of Animal Science for their support and collaboration.

The financial support for this project was made possible through the National Pork Board and Iowa Pork Producers Association. For this, I am grateful. I would also like to thank the companies that have supplied the data for this project. Additionally, I want to acknowledge all the students who have been involved with these projects in any way.

I owe many thanks to my parents, Paul and Cathy Abell, who have provided me with the means to pursue all of my goals. I appreciate their encouragement as I have worked toward this degree. Finally, I would like to thank my fiancé, Alex Bruns, who has been patient and supportive throughout my career as a graduate student.
ABSTRACT

The objective of this dissertation was to evaluate multiple approaches to incorporating sow longevity or lifetime sow productivity into a selection program. Sow longevity can be selected for using indicator traits, such as structural soundness and lameness. In the first study, objective measurements to detect sow lameness were examined. Lameness was chemically induced for a short time period in multiparous sows and their weight distribution and walking gait were objectively measured in the days following lameness induction. Using a classification tree analysis, it was determined that the mean weight being placed on each leg was the most predictive measurement when determining whether the leg was sound or lame after injection. The weight distribution measures had a greater predictive ability compared to the walking gait indicators. These measures could be used to select sows that are less likely to become lame and be removed from the breeding herd. While reducing the lameness instances in a herd would improve sow longevity, direct selection for longevity would be desirable. In the second study, genetic correlations between purebred and crossbred sow longevity were estimated. Most genetic improvement programs are based on an assumed relationship between purebred performance in a nucleus herd and their relatives’ crossbred performance in a commercial herd; however, this study found that there was little to no genetic correlation between purebred and crossbred sow longevity for this population. While longevity is heritable at both the nucleus and commercial levels, results from this study indicate that little improvement would be made in crossbred longevity if selection relies solely on purebred information. One way to select for sow longevity would be to estimate purebred genomic breeding values using records from a related crossbred population. A spreadsheet for estimating the total costs associated with incorporating
genome-enabled selection into a swine breeding program was developed as the final part of this dissertation. This tool will aid producers in estimating the economic viability of incorporating genome-enabled selection into their specific breeding program. Based on the results from these projects, it is recommended that a commercial test herd be implemented as part of a selection program to improve longevity or sow productive lifetime. If a genetic company can succeed in improving sow longevity through an effective breeding program, production efficiency and profitability can be improved for commercial swine operations.
CHAPTER 1: GENERAL INTRODUCTION

In a commercial swine breeding herd, sow longevity is a contributing factor to a swine operation’s overall success and profitability. A sow must remain in the herd between 3 and 4 parities to produce enough piglets to pay for herself (Stalder, 2000). Increasing the number of lifetime piglets produced per sow reduces the proportion of the sow’s gilt replacement and development costs that must be recovered by each pig. Because of this, a sow should not be voluntarily culled from the breeding herd as long as she is producing at an acceptable level, or level equal to the herd average, in terms of litter size and weight and no animal wellbeing issues are present (too big for penning, injury, etc.).

Additionally, removing females in early parities does not allow a swine operation to benefit from the increased sow and downstream grow-finish litter performance compared to the gilt’s productivity and that from her litter’s performance. Not only do sows tend to have greater litter sizes compared to gilts (Roehe and Kennedy, 1995), piglets from sow litters tend to have decreased mortality and higher performance at the finishers compared to piglets from gilt litters (Carney-Hinkle et al., 2013). When comparing the replacement gilt to the older sow in a commercial breeding herd, the gilt’s genetic advantage is not sufficient to cover the variable costs associated with gilt development until the sow has reached at least parity 7 under maximum genetic gain conditions (Abell et al., 2010). Using more realistic genetic improvement and generation interval values, the optimal culling time is more realistically the 10th or 11th parity.

Improving the average removal parity at the sow farm would increase the swine operation’s production efficiency. Besides improving management practices, sow longevity can be increased by improving genetics and management practices. While sow longevity has
a genetic component (Mészáros et al., 2010; Serenius and Stalder, 2004), there are multiple issues associated with incorporating the trait into a swine breeding program. Developing tools and methods to integrate sow longevity into a selection scheme can aid genetic suppliers in improving their breeding programs, as well as improve the profitability of many commercial swine operations.

One issue associated with selecting for sow longevity is that many sows are involuntarily removed from the herd due to non-reproductive reasons, such as lameness. Sows removed from the herd due to lameness issues tend to have produced fewer parities than sows removed from the herd due to undesirable production levels (Stein et al., 1990). Because of this, many sows do not get the opportunity to express their true potential for remaining in the herd based on superior reproductive performance. Selecting animals that are less likely to become lame would allow producers to have a greater proportion of females in the breeding herd with decreased risk for becoming lame due to their body conformation. This would result in fewer sows being removed from the herd at the early parities.

Currently, sow lameness is identified using a subjective visual scoring method (Zinpro, 2008). Subjective measurements can be influenced by intra- and inter-scorer variation, reducing the lameness score repeatability. The extensive scorer training required for visual scoring methods combined with high employee turn-over can become costly. Furthermore, increased employee turn-over causes more inter-scorer variation which lowers score accuracy. This reduced accuracy ultimately decreases the expected trait improvement over time. Developing an objective measurement system to detect lameness could lead to traits that can be easily incorporated into a breeding program. Weight distribution and gait while walking measures have shown the ability to detect differences between sound and lame
sows (Karriker et al., 2013). Further investigation needs to be done to determine the best measure(s) to detect sow lameness in terms of predictive ability and feasibility.

Selecting for sow lameness can reduce the involuntary culling percentage in a commercial swine breeding herd, allowing producers to better assess the true longevity or productive lifetime potential for sows in the herd. While increasing sow longevity is economical at a commercial level, it is not beneficial for nucleus herds that require an increased turnover rate (or reduced generation interval) in order to make the rapid genetic improvement needed to maintain genetic progress. Because of rapid turnover, nucleus herds do not allow sows to fully express their genetic potential for longevity, which is expressed at the end of the sow’s true production herd life. Furthermore, many genetic suppliers do not have easy access to commercial level records to monitor the longevity from a pedigreed crossbred offspring population descending from their purebred nucleus animals. A pedigreed crossbred commercial test herd can add value to a genetic supplier’s selection program by providing more accurate information, particularly for maternal traits or for lowly heritable traits, regarding commercial level performance which is the true breeding objective.

Since the genetic correlation between nucleus level performance and commercial level performance is not perfect or 1 (Wong et al., 1971) and selection decisions are often made based on nucleus level performance, realized genetic improvement observed at the commercial level is less than the genetic improvement rate occurring at the nucleus level. Variance component and heritability estimates differ between purebred and crossbred populations (Ehlers et al., 2005). It is important to determine if sows remaining in the nucleus herd for 2 parities are more likely to produce offspring that can remain in the commercial herd for at least 3 parities. In other words, is remaining in the nucleus herd a good indicator.
or predictor for sow longevity at the commercial production level? If this genetic correlation has an acceptable magnitude in the desired direction, then selecting sows based on nucleus level performance will increase longevity in commercial breeding herds. If the genetic correlation absolute value is small or near 0, then other methods to improve sow longevity would be necessary.

Since sow longevity is a sex-limited trait that is measured late in life, it is a trait where the breeding value accuracy can be greatly improved by using genome-enabled selection. With molecular information, breeding values for selection candidates that have not expressed the trait and whose parents may not have expressed the trait at the time of selection can be improved. However, there is a large cost associated with collecting genotypic information and integrating it into genetic evaluations. Thus, the accuracy increase must be enough to recover the additional costs associated with incorporating genome-enabled selection into a swine breeding program. The genetic gain rate must proportionally increase when compared to traditional selection methods so that the investment is viable.

**Dissertation Organization**

The following chapters in this dissertation include a comprehensive literature review, three scientific papers, and the overall project conclusions in this dissertation. The literature review in the second chapter outlines the current sow longevity knowledge and the issues associated with incorporating sow longevity into a swine breeding program. In the third chapter and first scientific paper, the predictive ability and accuracy of objective weight distribution and gait while walking measures relative to each other and to a visual lameness identification method to detect sow lameness are compared. This paper has been submitted to *Animal*. Chapter 4 contains a paper that estimates the genetic correlation between sow
Retention rates at the nucleus and commercial levels. This paper will be submitted to *Journal of Animal Breeding and Genetics*. In Chapter 5, a paper outlines the development of a tool to determine the total costs associated with incorporating genome-enabled selection into a swine breeding program. This paper has been submitted to *Genetics Selection Evolution*. The conclusion and implications of the three projects comprising this dissertation are summarized in the final chapter.

**Author contributions**

Chapter 3: AJ, LK and KS were involved in experimental design. GS, SH, KS and RF designed and developed the force plate measurement system. MR provided technical support for the project. CA and KS analyzed the data collected and drafted the manuscript.

Chapter 4: RF and TS developed the Bayesian analysis program. KG and MF provided technical support for the project. CA and KS developed the project and drafted the manuscript.

Chapter 5: JD, MR, and JM provided technical information and support for the project. CA and KS conceived the project, participated in spreadsheet development, and drafted the manuscript.
CHAPTER 2: LITERATURE REVIEW

Sow Longevity

There are multiple definitions for sow longevity. Longevity can be defined as the length of productive life in days, parity at removal, or some other measure of how long a sow remains in the breeding herd. Improving management practices, such as reducing voluntary culls and documenting correct removal reasons, can aide in improving sow longevity by understanding the real removal causes in a breeding herd. In the U.S. swine industry, the percent of sows culled annually is 45.2% with 8.2% sow mortality (PigCHAMP, 2012). Friendship et al. (1986) reported that increased culling rate decreased number born alive and number weaned.

When sows are culled after producing only 1 or 2, producers do not take advantage of the improved performance for sows compared to gilts. Schneider et al. (1982) reported that the total litter weight born alive from parity 2 sows was ~2 kg greater compared to litter weights from parity 1 females. Additionally, it has been shown that sows have greater litter weaning weight and greater litter weights 42 days post-weaning (Carney-Hinkle et al., 2013; Smith et al., 2007). Multiple studies have reported greater total born, number born alive, and number weaned from sow compared to gilts (Roethe and Kennedy, 1995; Tummaruk et al., 2001). However, not all studies found significant litter size differences when comparing sow and gilt litters (Carney-Hinkle et al., 2013; Southwood and Kennedy, 1991).

When evaluating removal reasons at commercial swine farms, it was noted that 23% of culling reasons were inaccurate (Knauer et al., 2007a). Farm-level assessments rely on accurate information, and these errors can negatively impact management decisions regarding sow removal. For example, if reproductive failure is the number one recorded
reason for sow culling, farm managers may concentrate on the sows’ reproductive performance. If many of the sows did not actually have reproductive issues resulting in removal, time and effort would be expended for a non-issue and another issue that should be addressed would be missed. Additionally, having erroneous information can slow genetic progress by decreasing trait heritability. There have been several studies that estimate genetic parameters for longevity. Mészáros et al. (2010) reported longevity heritabilities ranging from 0.05 to 0.18 for Landrace sows and from 0.14 to 0.25 for Large White sows. Guo et al. (2001) reported heritabilities for lifetime traits of ~0.25. Since longevity, like most reproductive traits, has low to moderate heritability, having inaccurate performance and/or pedigree information can almost stop genetic improvement.

There is a relatively high genetic correlation among the different sow longevity definitions, i.e. productive life, stayability from first to second litter, stayability from first to third litter, length of productive life, and lifetime production (Engblom et al., 2009) suggesting that the many of the same genes impact all the traits. Longevity has been reported to have unfavorable genetic correlations with average daily gain and backfat (Knauer et al., 2010; López-Serrano et al., 2000) and a favorable correlation with number weaned (Serenius et al., 2008). The favorable genetic correlation with number weaned is expected since low number weaned is often a culling criterion at many swine breeding herds. Culling on number weaned when selecting to improve longevity could be considered selection bias depending on the selection objective. For example, if the selection objective is to increase the number of parities produced by each sow, breeders could use the selection bias to their advantage by preventing low producing sows from remaining in herd for extended parities. If lifetime weaned is the selection objective to improve longevity and the trait where selection is
practiced is sow lifetime productivity, then the selection bias can be used directly in the selection objective to improve lifetime pigs weaned. Knauer et al. (2010) found that increased age at first farrowing decreased sow longevity while increased lactation feed intake improved longevity.

While sow longevity is heritable and can be improved through direct selection, improvement in other traits, such as lameness and structural soundness, can have a positive impact on sow longevity. Based on a study by Anil and co-workers (2009), lame sows were almost two times more likely to be removed within 350 days after lameness diagnosis compared to non-lame sows. Sows with poor leg conformation and abnormalities had a lower chance of survival compared to normal sows (de Sevilla et al., 2008). More sows are removed due to lameness or feet and leg problems at early parities compared to later parities (Anil et al., 2005; Dagorn and Aumaitre, 1979; Stein et al., 1990). Culling sows at early parities decreases productivity, efficiency, and ultimately, profitability of commercial breeding operations and can reduce genetic progress at nucleus and multiplication levels when culling occurs for something other than performance.

Since longevity is a low to moderately heritable, sex-limited trait that is measured late in life, it is a good candidate for genome-enabled selection. Onteru et al. (2011a) reported that ~10% of the genetic variation in lifetime traits can be explained by markers. Moreover, the authors reported multiple QTL regions associated with longevity. Likewise, Mote and co-workers (2009) identified specific markers associated with sow longevity. Onteru et al. (2011b) detected multiple QTL regions with significant associations for litter size traits in parities 1, 2, and 3, which indicates that markers can be used to predict a sow’s reproductive performance across her lifetime. Additionally, studies have reported significant associations
between QTL and structural soundness traits (Fan et al., 2009; Fan et al., 2011; Laenoi et al., 2012). Since structural soundness and lameness is associated with sow longevity, there is opportunity to indirectly select for increased sow lifetime through QTL associated with leg and conformation traits.

**Lameness**

Lameness is defined as “impaired movement or deviation from normal gait” (Wells, 1984). Lameness in swine, poultry, horses, and cattle has a large negative economic impact to livestock producers (Corr et al., 2003). The abnormal pig locomotion has been described as having a shorten stride length, stiff movements, and lowered ability to accelerate and change direction (Main et al., 2000). Locomotor disorders can be associated with neurological disorders, hoof or limb lesions, a mechanical-structural problem, trauma, or metabolic and infectious disease (Smith, 1988; Wells, 1984). A cull sow evaluation by Knauer (2006) found that 85% of sows evaluated at harvest had at least one lesion impacting at least one foot. The same authors further noted that lameness is a common reason why sows leave the breeding herd, typically second only to reproductive failure. However, many sows having been culled reportedly for reproductive failure may be a result of lameness that occurred earlier in the sows’ life cycle. Additionally, cull sows from a U.S. Midwestern harvest facility were evaluated, and it was reported that sows with a higher body condition score have more heel lesions on the front and rear feet compared to sows with lower body condition scores (Knauer et al., 2007b). From a study conducted at a Midwestern U.S. swine integrator, it was shown that sows with hoof lesions and abnormalities had lower performance compared to control sows (Fitzgerald et al., 2012). That study showed sows
with foot lesions and abnormalities tended to have greater pre-weaning mortality, fewer pigs weaned, and lower litter weaning weights (Fitzgerald et al., 2012).

There are numerous methodologies that can be employed to subjectively and objectively measure and evaluate the relative degree of lameness for an individual animal at a given point in time. Subjective lameness identification systems are designed to categorize lameness expressed primarily while the animal is walking and have been developed for cows (Manson and Leaver, 1988), dogs (Quinn et al., 2007), sheep (Welsh et al., 1993; Kaler et al. 2009), horses (Keegan et al., 2010), and finishing pigs (Main et al., 2000; Nielsen et al., 2001; Rothschild and Christian, 1988). The scoring systems used in the livestock industries have been implemented so that caretakers can quickly and affordably quantify lameness prevalence in the herd on any particular day. However, lameness scoring disagreement assigned to an individual animal can occur (Flower and Weary, 2006). These disagreements result from either inter- or intra-scorer variation meaning that different scorers may provide different scores or that the same individual may provide different scores when scoring the same animal twice.

While subjective lameness and leg structure measures are heritable and can be improved through selection (Rothschild and Christian, 1988; Serenius et al., 2001), an objective and standardized method for assigning lameness scores to an animal would likely be more accurate when compared to subjective scoring and provide producers with a useful tool to assess lameness resulting in more timely identification and treatment of lame sows. Additionally, increased accuracy would allow for faster genetic progress when improving the trait through traditional selection methods.
**Force Plate Measurement System**

One such method that shows promise is the force plate measurement system. This device quantifies the amount of force each limb applies to the assessment tool surface (Pastell et al., 2008; Sun et al., 2011). Pigs put more weight on front feet compared to rear feet, and force measurements vary depending on floor conditions (von Wachenfelt et al., 2009). Force plate measurement systems can measure variables that have been associated with objectively classifying structural abnormalities into lameness scores. An animal will distribute less weight on the limb(s) that is(are) painful or structurally unsound (Corr et al., 2003). The use of such equipment has been evaluated in other species such as dogs (Evans et al., 2005), chickens (Corr et al., 2003), and dairy (Pastell and Kujala, 2007).

Force plate measurements are typically obtained during locomotion for these species, which is acceptable because these animals spend a greater portion of their time moving. However, in current housing systems, sows spend the majority of their time in stalls. In this housing system type, lameness identification more commonly occurs while sows are standing. When sows are housed in pen gestation systems, a force plate application could be made to electronic sow feeding systems so that the weight distribution between limbs can be evaluated daily and total weight over time can be monitored. Available force plate measurement systems either do not allow for measurements over time (i.e. current force plates measure approximately 3-5 seconds) or the instruments are too large to measure sows in the gestation stall, the most commonly utilized gestation housing system throughout the U.S.

Sun and co-workers (2011) developed an embedded microcomputer-based force plate system as a lameness research platform to measure sow weight distribution. The results
indicate that this measuring system was able to identify sow lameness by separately measuring the weight placed on each leg. As mentioned above, few research projects in swine have evaluated the relationship between sow leg weight distribution with different lameness degrees and automatically detected lameness in sows prior to clinical veterinary diagnosis. Karriker and co-workers (2013) reported that lame sows will reduce the amount of weight placed on their lame foot and compensate by placing more weight on the foot lateral to the lame foot.

**Gait Measurements**

Gait measurements used in studies monitoring swine locomotion include stride length, stride time, and stance time (von Wachenfelt et al., 2008; von Wachenfelt et al., 2010; Grégoire et al., 2013). Stride length is defined as the length between each consecutive step made by the same foot. Stance time is the amount of time a foot activates a sensor on a given step. Stride time is defined as the time between two successive steps of the same foot. Gait measurements can be impacted by the walking surface (concrete or rubber mat) and walking path (curved or straight) (von Wachenfelt et al., 2008; von Wachenfelt et al., 2010).

The GaitFour walkway system (CIR Systems Inc., Sparta, NJ) is a tool that has been developed to record gait measures in animals while they are walking. The GaitFour walkway system has been utilized to define gait characteristics in lame and non-lame sows (Grégoire et al., 2013; Karriker et al., 2013). Grégoire and co-workers (2013) reported that lame sows tended to have a shorter stride length and longer stance time than non-lame sows. Kariker et al. (2013) determined that lame sows will spend less stance time on lame feet compared to the stance time on non-lame feet. Boars selected for poor front leg structure had significantly
shorter stride length compared to boars selected for desirable front leg structure (Morrow et al., 1991).

Additionally, the GaitFour walkway has been used to evaluate gait characteristics in healthy Labrador Retrievers (Light et al., 2010). Differences between lame and sound dog gaits were detected using the GaitFour walkway system (Lequang et al., 2009). Results from this study indicated that dogs will place less weight on their injured leg compared to their sound leg which agrees with the reported work using sows.

Classification Trees

Classification trees can be used to analyze categorical data. Classification trees are structured as dichotomous trees using nodes or decision tests to define the variable threshold values for each classification. The threshold values determine the point or value separates different classifications. Classification trees have been used to determine genetic similarities among livestock breeds (Li et al., 2004; Martínez et al., 2000; Megens et al., 2008; Moazami-Goudarzi et al., 1997). These studies utilized microsatellite data as decision variables for the classification trees. Additionally, classification trees can be used in association studies (Zhang and Bonney, 2000).

In a classification tree analysis, the distribution for each variable is examined for each classification (for example lame or sound). Variables where the distributions for each classification do not overlap explain the largest proportion of variation between classifications. Recursive partitioning is used to determine which variables provide the most informative divisions for predicting the classification category for each observation. Classification trees are robust to outliers since they are nonparametric (Verbyla, 1987). However, bias can result when exhaustive search algorithms are used to form the
classification trees. Methods where not all binary splits are examined can reduce the bias and still lead to trees with similar accuracies as those created with exhaustive search algorithms (Loh and Shih, 1997).

The decision to add a node to the tree is often based on the improvement in the Gini coefficient or how often a randomly chosen observation would be incorrectly classified. However, other splitting criteria can be used. For example, Shih (1999) showed the accuracy of trees developed using chi-squared and entropy criterion or mean posterior improvement criterion can result in more accurate trees compared to trees split using the Gini criterion.

**Improving Crossbred Performance**

Crossbred sows outperform their purebred parents for reproductive traits due to the heterosis advantage for the crossbreds (Johnson, 1980). Because of this, sows in commercial breeding herds are crossbred, typically a two-way cross between two maternal lines. In addition to improved litter size, crossbred animals perform better at the finisher in terms of average daily gain (Smith and McLaren, 1967). Smith and McLaren (1972) compared 2, 3, 4, 5, and 6-breed crosses and reported that livability and growth rate were optimized for a 4-way cross finishing animal, suggesting that more breeds in the cross does not necessarily lead to improved performance. Since different breeds have different combining abilities, meaning that crosses between different breeds will have varying performance levels, it is important to carefully choose which breeds or pure lines are used to make crossbred animals (Magee and Hazel, 1959).

Additionally, production efficiency could be improved using a rotational crossbreeding system compared to a purebred herd (Smith and McLaren, 1972). Breeding programs for crossbred animals can be classified into three categories: rotational, terminal, or
rota-terminal. In a rotational crossbreeding program, replacement gilts are saved from the market crosses and mated to a pure line boar (Ahlschwede et al., 1987). The breed of the boar changes every generation. In a terminal cross system, replacement females are not saved from the market crosses. Replacement gilts are derived from maternal line matings and are mated to a boar from another breed to make market animals (Ahlschwede et al., 1987). A rota-terminal system combines the rotational and terminal breeding programs. Replacement gilts are derived from a rotational crossbreeding system utilizing different maternal line breeds and then mated to a terminal sire to produce market animals (Ahlschwede et al., 1987).

A rotational swine crossbreeding system relies on selecting replacement females from a market hog group. After gilts are selected, producers must carefully track breed makeup to ensure that the correct boar breed is used to generate the next generation of offspring. Within a sow group multiple breed composition may be represented. Because of this, boars from multiple breeds (up to the number of breeds utilized in the rotational system) may need to be used to ensure market hogs have the correct genetic makeup. It is clear that a rotational crossbreeding system requires a great deal of record keeping just to maintain genetic makeup information for the breeding herd females. If this fails to occur, performance suffers and the crossbreeding advantage may be lost. To avoid the record keeping associated with proper implementation of a rotational crossbreeding system, producers may replace females after they have farrowed a single litter in order to reduce the number boar breeds needed at mating. Because of this, the swine operation does not take advantage of the increased performance of sows and piglets from sow litters compared to gilts and their offspring.
Terminal cross systems do not require the meticulous attention paid to matings as in rotational systems since gilts are not saved from market crosses. This system is the easiest to implement by barn workers, because all commercial females are mated to the same terminal cross boar breed or line. The terminal crossbreeding system fully takes advantage of hetorosis if implemented correctly. That is, both maternal and market hog heterosis is maximized. Additionally, if a crossbred terminal boar is used, then the entire commercial breeding program maximizes heterosis at all levels (maternal, paternal, and offspring).

Since the rota-terminal system is a combination of a rotational crossbreeding system and a terminal cross system, it has the advantage of using only one terminal line to produce market hogs; however, producers must monitor the sow genetic makeup in the maternal herd and the breed of boars mated to them to ensure replacement gilts have the correct breed composition.

Today, most producers employ a terminal cross system. A terminal cross system is often incorporated by utilizing internal multiplication programs to generate crossbred replacement gilts for commercial level production (i.e. Landrace × Large White) and producing market crosses by mating the crossbred females to a terminal line sire (i.e. Duroc). Crossbred replacement gilts could be purchased from a seedstock supplier, but this creates a potential risk of introducing new diseases into the breeding herd.

Improving crossbred performance is the breeding objective for swine breeders and swine breeding stock suppliers. However, most swine genetic companies make selection decisions based on purebred performance records and the genetic correlation between purebred and crossbred performance is not perfect or 1 (Cecchinato et al., 2010; McLaren et al., 1985; Wong et al., 1971). Genetic gain for crossbred performance can be improved by
estimating purebred breeding values using both crossbred and purebred information. Ehlers et al. (2005) showed that heritabilities estimated by pooling crossbred and purebred data are similar to estimates from purebred data alone and crossbred data alone. Provided that there is a genetic correlation between the purebred and crossbred data, pooling data could be used to increase the accuracy of breeding value estimates through the added information. Ehlers et al. (2006) showed that the accuracy increased when pooled data were used to estimate breeding values instead of purebred data alone. The increase in accuracy was 0.087 for number born alive (0.557 to 0.644), 0.078 for adjusted litter weaning weight (0.575 to 0.653), and 0.073 for weaning to first service interval (0.586 to 0.659).

Heritabilities and genetic correlations between traits in a selection index vary between crossbred and purebred populations (Louca and Robison 1967; Lutaaya et al., 2001; Stanislaw et al., 1967). Since selection index coefficients are derived based on genetic correlations and trait heritabilities, genetic parameter differences among purebred and crossbred populations can impact selection index coefficients. If index weights are incorrect, selection would not be optimized, decreasing the rate of genetic gain.

Genomic selection on purebred animals can be conducted using crossbred populations with minimal loss in accuracy if the pure line contributed to the genetic makeup of the crossbred animals (Ibáněz-Escriche et al., 2009; Toosi et al., 2010). Zeng and co-workers (2013) reported that a dominance model is superior to the additive model and the breed specific model in improving crossbred performance when selection is performed on purebred animals and dominant gene action is present. The results from these studies combined suggest that crossbred information can be used when incorporating genomic information into
a purebred selection program when the breeding objective is improved crossbred performance.

**Categorical Traits**

Since categorical traits are not normally distributed, standard genetic evaluation methods cannot be used to include these traits in a breeding program. When compared to a typical linear animal model, a threshold hold model will have greater accuracies and better genetic parameter estimates for categorical traits (Ramirez-Valverde et al., 2001). A threshold model will analyze the categorical data as if there is an underlying normally distributed random variable that is determining in which category an observation will exist (Albert and Chib, 1993). Based on the normally distributed trait, thresholds are used to determine value ranges for each category (Korsgaard et al., 2003). When performing multivariate analysis with categorical traits, closed form solutions are not easily computed.

Gibbs sampling is one method of Markov chain Monte Carlo and is often used to obtain variance component estimates for categorical traits (Scott et al., 1991; Sorensen et al., 1994). Gibbs sampling procedures draw from joint distributions which are produced from the fully conditional posterior distribution for each parameter (Sorensen and Gianola, 2010). In other words, each parameter in the distribution is sampled individually, conditional on the current values for all other parameters. Blocked Gibbs procedures use similar sampling procedures, but rather than estimating a single parameter conditional on all other parameters remaining the same, several parameters are sampled at time. Parameter groups are chosen based on the joint distributions formed when the parameter groups are conditioned on the remaining parameters. In each iteration, all parameter groups are sampled individually, conditional on the remaining parameters. This sampling creates a Markov chain which can be
described using a stationary distribution. This stationary distribution will be equivalent to the complete likelihood and can be used to estimate variance components. When using these methods to estimate variance components, initial starting points must be chosen based on prior knowledge. The prior distribution used in the analysis can impact the resulting estimates (Stock et al., 2007). When using a threshold model and Bayesian analysis for categorical traits, the predictive error is reduced if contemporary groups are fitted as random rather than fixed effects (Luo et al., 2001). Additionally, Luo et al. (2001) reported that the sire model outperformed, the animal model, in terms of predictive error, with estimating variance components for categorical traits due to some subclasses having only values in one category in the animal.

Genetic parameters for categorical traits have been estimated for single traits (Bennewitz et al., 2007; Eler et al., 2002; van der Westhuizen et al., 2001), but few studies have attempted to determine genetic correlations between categorical traits (van Tassell et al., 1998) even though the methodology for the analysis has been established (Korsgaard et al., 2003). Understanding the genetic correlations between categorical will allow for a wider range of traits to be incorporated into a selection program.

**Single Nucleotide Polymorphism (SNP) Chip**

The development of the PorcineSNP60 Bead Chip has opened the door for the use of genomic selection in swine breeding improvement programs (Ramos et al., 2009). Duroc, Pietrain, Large White, Landrace, and Wild Boar animals were used to develop the chip. The chip will allow studies that aim to estimate breeding values for economically important traits using molecular markers to be conducted. The pig genome has recently been sequenced
(Archibald et al., 2010). This should allow for more SNPs to be mapped, thus increasing the knowledge concerning the associations between genes and phenotypes for traits.

The 60K SNP chip shows how far the technology has come in the last 15 years when the only tests available were for single markers affecting single traits. The first genetic tests developed were for markers affecting economically important traits such as number born alive (Estrogen Receptor gene or ESR), growth rate and feed intake (Melanocortin-4 Receptor or MC4R), and meat quality (protein kinase adenosine monophosphate-activated gamma(3)-subunit or PRKAG3) (Stalder et al., 2005; Kim et al., 2000; Rothschild et al., 2007). Including genotyping in genetic evaluations was considered in Stalder et al. (1997). The QTL evaluated in this study was associated with Porcine Stress Syndrome. The study showed that there was no significant difference between the maternal performance of normal (NN) and carrier (Nn) Landrace females.

**Genomic Selection**

The idea of genomic selection was first introduced by Meuwissen and co-workers, (2001). Genomic selection is based on determining the individual’s breeding value by summing the allele effects at each SNP across the individual animal’s entire genome. Ultimately, the genotypic information combined with the individual’s phenotypic performance records and both genotypic and phenotypic information from relatives of the selection candidates are used in combination to more accurately estimate the breeding value for individual animals. Genomic selection should allow selection decisions to be made earlier in the animal’s life or among non-producing animals in the sex limited trait case, thus decreasing the generation interval for the herd and improving the rate at which genetic progress should occur (Goddard and Hayes, 2007).
The breeding value obtained from traditional best linear unbiased prediction (BLUP) selection is modified to obtained genomically enhanced breeding values. This is accomplished by calculating a breeding value for each animal based solely on each animal’s genomic information. The effect of each position (or the SNP) is estimated after accounting for other fixed effects such as contemporary group as is typically done in a traditional genetic evaluation (Erbe et al., 2010). The effects from the genomic information at each position evaluated are summed to estimate the genetic merit or the value of the animal as a parent. These effects are estimated from a base population with relevant phenotypic records and genomic information. A weighted sum is computed by combining the genomic estimated breeding value (EBV) and the EBV from traditional BLUP selection. The weights are determined based on the amount of information included in each set of EBVs, the variation from each set of EBVs, and the associations or correlations between the EBV sets. This sum is used as the final genetic index value for each animal (combines both phenotypic and genomic information) for that particular genetic evaluation. The information from phenotypes and genomics are combined in order to ensure that all information available is used in the selection process to provide producers with the most accurate information as possible when they are making selection decisions.

**Imputation**

Inferring high density genotypes from a low density panel is known as imputation (Habier et al., 2009). When imputation is used, selection candidates are genotyped using low density panels and the actual selected animals are often re-genotyped using a high density marker panel. A low density marker panel has fewer markers spread farther out along the DNA compared to a high density marker panel. Utilizing the low density marker panel can be
an effective tool to reduce genotyping costs once the initial training data set has been established. A training data set is used to determine population haplotypes so that imputation can be used to infer or guess high density genotypes from a low density marker panel. To further reduce costs, companies may choose to genotype only males. While this saves costs, EBV accuracy may be reduced, especially for sex-limited traits or novel traits where phenotypic data are not routinely collected. Henryon et al. (2012) showed that when 20% or fewer of the selection candidates were genotyped, genetic gain rates were greatest when only males were genotyped. When 40-50% of the candidates were genotyped, the genetic gain rate was maximized when the male to female ratio for the genotyped candidates was 25:75. Genotyping only females resulted in the poorest genetic gain and the greatest inbreeding no matter which proportion of candidates was genotyped.

Huang and co-workers (2012) analyzed several genotyping strategies to determine an optimal level between reducing costs and maintaining high imputation accuracy. From this study, it was concluded that a sizable (~2,500 animals) base training population genotyped with a high density marker panel was needed to maintain an acceptable imputation accuracy regardless of the relationship between the training and test populations. Additionally, the authors illustrate that parents of the testing population could be genotyped using a low density marker panel and still maintain accuracy if a base training population has already been established. This will reduce the costs associated with genotyping selection candidates.

The EBV accuracy from genome-enabled selection can be predicted using the genomic relationship between the training animals and selection candidates (Goddard et al., 2011; Habier et al., 2010). Increasing the average genomic relationship between the training
data and selection candidates may improve EBV accuracy (Goddard et al., 2011; Habier et al., 2010).

**Genetic Improvement Expected from Genomic Selection**

Muir (2007) showed that there is potential for increased genetic improvement when using genomic selection in addition to traditional BLUP selection. The added information from genomic information increases the breeding value accuracy and therefore, increases response to selection. Pszczola and co-workers (2010) showed that when the heritability of a trait was 0.01, the accuracy can be increased from 0.50 to 0.72 when going from using traditional BLUP selection methods to using genomic breeding value estimations with 2000 genotyped animals. Genomic breeding value accuracy for young dairy bulls can be increased to over 0.70 if at least 4 generations of animals are used in the training data set (Pszczola et al., 2010). Furthermore, this study showed that genomic selection could have an advantage over traditional methods when applied to traits that have low heritability and are difficult to measure (i.e. expensive to measure, sex-limited, or observed later in life).

Weigel et al. (2009) used 4,703 genotyped Holstein bulls to estimate the correlation between the genomic predicted transmitting ability (PTA) and the PTA estimated based on progeny performance 5 years after the genomic PTA was estimated. The correlation was 0.38 when the entire genome was used to estimate the PTA. Su et al. (2010) used 3,330 bulls to estimate PTA reliabilities with 38,134 SNPs compared to PTAs estimated using conventional parent averages. The GEBV PTA reliability was 0.26 higher when compared to the conventional parent average PTA.

Meuwissen et al. (2001) showed that the genomic selection accuracy can be as high as 0.85 when genomic breeding values are estimated using Bayesian methods, 1-cM spacing
between markers, and 2200 phenotypic records. Using BLUP to estimate genomic breeding values resulted in an accuracy of 0.73 when the same number of phenotypic records and marker spacings were used. Solberg et al. (2008) similarly demonstrated that increasing the marker density increased the genomic breeding value accuracy. However, when Erbe et al. (2012) compared the genomic prediction accuracy from 50K and 800K marker panels, the estimates from the 800K marker panel did not consistently have a greater accuracy than the 50K marker panel estimations depending on which trait and what breed was being analyzed.

Genomic breeding value application has been investigated for pig populations. Nielsen et al. (2010) showed the correlation between genomic breeding values and traditional BLUP breeding values was 0.62 for the 170 boars used in their data set. Cleveland et al. (2010) reported the genomic breeding value accuracy for total pigs born per litter to be between 0.64 and 0.82, depending on the training data set used. The authors reported that the accuracy for the number of stillborns per litter ranged from 0.33 to 0.68. Hayes and Goddard (2003) reported that there was almost no difference in response to selection for growth traits when comparing traditional BLUP selection and marker-assisted selection. A larger genetic gain rate differential between BLUP and marker-assisted selection method was shown for net feed intake and pigs born alive.

Dekkers (2007b) developed a method using selection index theory to calculate the genetic response expected from incorporating genomic selection into a selection index. The method deterministically calculated the genetic response anticipated from using genomic selection with defined genetic parameters. The study showed that, for a trait recorded on both sexes prior to selection, selection based on markers alone can improve response by 8.5% compared to selection based on solely on phenotypic information. If genomic information is
used alone, the genetic improvement resulting from selection may not exceed the genetic improvement based on BLUP breeding values (Dekkers, 2007b; Muir, 2007). All available phenotypic and genomic information should be incorporated into the EBVs to ensure the most accurate EBV for selection and thus, the most rapid genetic gain.

Jannick (2010) showed that genomic selection results in greater gain compared to phenotypic selection for the first seasons of a barley breeding program. However, after about 15 seasons of selection, the cumulative gain from phenotypic selection was approximately equal to the cumulative gain from genomic selection. This equilibrium occurred sooner with a large training population. The long-term performance associated with genomic selection compared to traditional BLUP selection in a livestock population has yet to be determined.

Based on stochastic simulations, annual genetic gain could be increased by 23 to 91% for a maternal line (Lillehammer et al., 2011) and 27 to 33% for a terminal line (Tribout et al., 2012). Based on these increases in annual genetic gain, there is potential for genome-enabled selection to be profitable for both maternal and terminal line selection programs. Ibáñez-Escriche et al. (2009) demonstrated that crossbred data can be used to determine genomic estimated breeding values (GEBVs) for purebred individuals. When the training population consisted of 1000 crossbred animals (2 breed cross using closely related breeds), the GEBVs accuracy was over 0.78 for 500 markers and over 0.81 for 2000 markers. When the crossbreds were developed from distantly related breeds, the accuracies were over 0.72 and 0.81 for 500 and 2000 markers, respectively. The individuals selected using these breeding values will be those that can be expected to have the best performing crossbred offspring. Dekkers (2007a) showed that the use of crossbred performance records and marker-assisted selection resulted in 74% and 43% greater response than purebred selection
or combined crossbred and purebred selection when the marker-based EBVs accuracy was 0.9.

Genomic information can also be utilized to correct pedigree errors. Using genomic information in this way, not only leads to increased accuracy through genotypic data, but traditional BLUP EBVs are improved due to a more correct or accurate pedigree. Correcting pedigree errors will improve the herd connectedness and properly associate relatives’ records.

**Economic Value of Genomic Selection**

König et al. (2009) calculated the economic benefit for using genomic selection in the dairy industry. They determined the value on a population-wide basis rather than on an individual herd basis. The genomic breeding value accuracy had to be at least 0.70 in order for the discounted profit to be greater for a genomic breeding program compared to a conventional progeny testing program. Because generation interval was reduced and the progeny test portion required to prove sires could be eliminated, genomic selection was more profitable when compared to using only traditional phenotypic selection. The discounted returns were calculated as the economic value from the increased genetic potential for the population over a 15-year investment period. The discounted costs included the expense for genotyping and progeny testing. The discounted profit was increased by a factor between 1.36 and 2.59 for each genomic breeding program considered in the study. Discounted profits were improved by the greatest amount when 100% of the cows were inseminated by young bulls that did not have daughter records. However, they did not account for the need to re-estimate genomic values. If they did, the re-estimation cost may decrease the discounted profits to a point where it is more profitable to use traditional selection methods. The swine
industry will greatly benefit from knowing the potential economic value gained from genomic selection.

Van Eenennaam and co-workers (2010) estimated the breakeven value for a DNA test from a hypothetical marker panel assumed to be associated with 7-41% of the selection criteria additive genetic variance to be between $142 and $256 for a seedstock supplier. This assumed that 3% of the bull calves were used as studs and the remaining top 50% were sold as commercial bulls. In the dairy industry, incorporating genomic selection into a breeding program would have the added benefit of reducing the generation interval by at least half of the length associated with traditional BLUP selection (Schaeffer, 2006; Schefers and Weigel, 2012). The generation interval decrease could contribute to a larger proportion of the increase in genetic gain rate compared to the increase associated with improved accuracy. Since pigs are selected and mated relatively soon after sexual maturity, the generation interval would not be impacted by incorporating genomic information into the selection program. Therefore, swine genetic suppliers must rely solely on increased accuracy to make genomic selection profitable.

**Single-Step Selection**

Another method of genomic selection is single-step selection. This method requires less computing power and time when running genetic evaluations. Single-step selection utilizes genomic information by modifying the relationship among animals involved in the analyses (relationship matrix) compared to what is used in traditional BLUP selection (Legarra et al., 2009). Typically, the relationship between animals used in genetic evaluations is based on pedigree relationships among animals in the evaluation. With single-step selection, the relationship among individuals (coded within the relationship matrix) is
modified to include genomic relationship information (Chen et al., 2011a). This method takes advantage of the ability to better estimate the actual relationship among relatives rather than just using the average expected relationship among relatives. The relationships are estimated by collected the genotype information on animals in the genetic evaluation and using this information to determine the actual proportion of DNA that is common among two animals compared to average amount of DNA common among all animals in the population. For example, all full siblings have a numerator relationship of 1/2; however; with genotype information, full siblings’ relationships can be differentiated between those that have more alleles in common compared to other full siblings and could range from 0 to 1. After the relationship matrix has been modified, the genetic evaluation is conducted using similar methods as traditional BLUP selection. When single-step selection is practiced, the improved accuracy is achieved due to better genetic relationship estimates among animals rather using expected genetic relationships based on pedigree information (Legarra et al., 2009). Hickey and co-workers (2012) showed when imputation methods are implemented the correlation between the single-step genomic EBV and traditional progeny test BLUP EBV was approximately 0.50.

Chen et al. (2011b) showed that the predictive ability for genomic breeding values estimated using the single-step selection method was ~0.1 greater for chicken body weight and ~0.15 greater for breast meat area when compared to traditional BLUP breeding values. In a study by Koivula and co-workers (2012), the genomic predictions from the single-step method were highly correlated with genomic predictions from genomic selection (r>0.95); however, the validation accuracies were greater for the single-step method compared to genomic selection for all traits analyzed (milk, protein, fat, and mastitis). While validation
accuracies for the single-step method were approximately 0.1-0.2 higher than parental average estimations, they were only ~0.02 higher for the youngest bulls in the data set. This suggests that genomic information may not have as much value at an early age when selection decisions would be made.
CHAPTER 3: USING CLASSIFICATION TREES TO DETECT INDUCED SOW LAMENESS WITH A TRANSIENT MODEL

Modified from a paper submitted to Animal

C. E. Abell¹, A. K. Johnson¹, L. A. Karriker², M. F. Rothschild¹, S. J. Hoff³, G. Sun¹,a, R. F. Fitzgerald¹,b, and K. J. Stalder¹

¹Department of Animal Science, Iowa State University, Ames, Iowa 50011, USA
²Swine Medicine Education Center, Department of Veterinary Diagnostic and Production Animal Medicine, Iowa State University, Ames, Iowa 50011, USA
³Department of Agriculture and Biosystems Engineering, Iowa State University, Ames, Iowa 50011, USA

aDepartment of Biological and Agricultural Engineering, Texas A&M University, College Station, Texas 77843, USA
bPIC North America, 100 Bluegrass Commons Blvd., Suite 2200, Hendersonville, TN 37075, USA

Abstract

Feet and legs issues are some of the main causes for sow removal in the U.S. swine industry. More timely lameness identification among breeding herd females will allow better treatment decisions and outcomes. Producers will be able to treat lame females before the problem becomes too severe and cull females while they still have salvage value. The objective of this study was to compare the predictive abilities and accuracies of objective weight distribution and gait while walking measures relative to each other and to a visual lameness identification method to detect induced lameness among multiparous sows.

Developing an objective lameness diagnosis algorithm will benefit animals, producers, and scientists in timely and effective identification of lame individuals as well as aid producers in their efforts to decrease herd lameness by selecting animals that are less prone to become lame. In the early stages of lameness, the weight distribution among legs and walking gait are impacted. The weight placed on each foot can be measured using a micro-computer based
force plate, and gait can be analyzed using a gait evaluation system. Lameness was chemically induced for a short time period in multiparous sows and their weight distribution and walking gait were measured in the days following lameness induction. Using a classification tree analysis, it was determined that the mean weight being placed on each leg was the most predictive measurement when determining whether the leg was sound or lame. The classification tree’s predictive ability decreased as the number of days post-lameness induction increased. The weight distribution measurements had a greater predictive ability compared to the walking gait measurements. The error rates associated with the weight distribution trees were 12.5% and 23.7% at 6 days post-lameness induction for front and rear injected feet, respectively. For the walking gait classification trees, the error rates were 25.0% and 32.6% at 6 days post-lameness induction for rear and front injected feet, respectively. More timely lameness identification can improve sow lifetime productivity as well as animal welfare.

**Key Words:** Gait, Lameness, Sow, Weight Distribution

**Implications**

Developing an automatic lameness diagnosis algorithm will benefit animals, producers, and veterinarians in timely and effective lameness identification among individual sows before clinical signs are apparent, as well as aid producers in their efforts to decrease herd lameness by selecting animals that are less prone to become lame. Being able to predict sow lameness can aid in delivering maximum animal health and welfare benefits, improving sow lifetime productivity, and optimizing sow farm labor.
Introduction

Lameness is defined as “impaired movement or deviation from normal gait” (Wells, 1984). Locomotor disorders can be associated with neurological disorders, lesions on the hoof or limb, a mechanical-structural problem, trauma, or metabolic and infectious diseases (Smith, 1988; Wells, 1984). A cull sow evaluation by Knauer (2006) found that 85% of sows evaluated at harvest had at least one lesion impacting at least one foot. The same authors further noted that lameness is a common reason why sows leave the breeding herd, typically second only to reproductive failure. However, many sows reportedly culled for reproductive failure may have been culled as a result of lameness that occurred earlier in the sows’ life cycle.

The high lameness incidence can result in a large profit loss for swine operations due to the sows’ early removal from the breeding herd before they have paid for themselves and mortalities that result in no salvage value for cull sows (Anil et al., 2005; Stalder et al., 2000). Additionally, costs associated with treating sow lameness and reduced production (Fitzgerald et al., 2012) further reduce the sow farm’s production efficiency. The reduced production results from reduced feed intake which impacts piglet weaning weight and pre-weaning mortality resulting in fewer full value pigs at weaning (Fitzgerald et al., 2012). Increased labor required to treat and remove lame sows is another expense to the swine operation. Impaired worker morale occurs when a large amount of time is spent removing sow mortalities. Combined, the reduced productivity and increased expense associated with lame breeding herd animals reduces the sow farm’s production efficiency, increases production costs, and decreases overall profitability.
Identifying a sow during early lameness can allow producers to retain the sow’s salvage value or provide more timely and effective treatment options for the sow. Currently, sow lameness is scored using subjective visual identification methods (Zinpro, 2008). Visual scoring methods require substantial training time to allow individuals to become both accurate and proficient when evaluating breeding herd females for lameness. With high employee turnover rates, this can become costly and inefficient because barn workers are constantly in a training state for lameness detection. Visual identification methods require more intensive sow evaluation and observation time in order for the scorer to accurately identify lameness.

Few swine research projects have attempted to detect lameness before it is clinically apparent using visual observation methods. An objective identification method is needed to identify sows during early lameness stages potentially when lameness is undetectable by visual identification methods. Early in the lameness process, sows will change the magnitude of the difference in weight distribution between legs (side-to-side, front-to-back, and contra-laterally) and will change their gait while walking. Using technologies to detect leg weight distribution differences can allow an objective, lameness detection method to be developed.

Detecting lameness before clinical signs are visually apparent or evident will allow producers to cull females while they still have salvage value rather than allowing lameness to progress where treatment delays marketing or where lameness results in mortality or necessitates euthanasia. The objective of this study was to compare the predictive abilities and accuracies of objective weight distribution and gait while walking measures relative to each other and to a visual lameness identification method to detect induced lameness among multiparous sows.
Materials and Methods

Sows were housed and fed individually according to the Swine Care and Use Guidelines (Federation of Animal Science Society, 1999), and protocols were reviewed and approved by the Iowa State University Animal Care and Use Committee before conducting experimental work. Twenty-four multiparous sows derived from the Large White breed with mean weight 198 kg were used in this study (range 162 kg to 252 kg). The sows’ parity ranged from 1 to 4 with a 2.5 average. Prior to lameness induction, weight distribution measurements, gait measurements while sows were walking, and a visual lameness score were collected for each sow to determine a baseline “sound” value for all measurements. The 24 sows were injected with 10mg amphotericin B in the distal inter-phalangeal joint at one of four injection sites (left front toes, right front toes, left rear toes, and right rear toes) according to the methods outlined in Karriker et al. (2013). The injection site was randomly assigned to each sow. The injection resulted in synovitis, or synovial membrane inflammation, causing the sows to become lame. Sows injected in rear feet did not completely resolved lameness until after 6 days post-injection (DPI), while sows injected in front feet appeared to resolve lameness by DPI +6 (Karriker et al., 2013). For this study, a sows’ foot was considered to be lame after it was injected with amphotericin B and clinical lameness signs were observed.

The sows’ weight distribution on each foot was measured using a microcomputer-based force plate for 9 days following lameness induction (Sun et al., 2011) and sows were scored for lameness using a visual analog scale. Additionally, each sows’ gait was evaluated while walking using the GaitFour walkway system (CIR Systems Inc., Sparta, NJ) on DPI +1 and +6. Each sow was injected a second time in the lateral joint compared to their first
lameness induction during the second treatment replication and subsequent measurements were recorded. This resulted in 48 lameness events (24 sows × 2 replications) with weight distribution measurements.

For the measurements collected using the microcomputer-based force plate (Sun et al., 2011), the weight distribution was measured 2 times per second for 15 minutes each day. Since sows place a disproportionate amount of their weight on their front feet, the data were analyzed separately based on which body half (front or rear) was induced lame. Weight distributions and gait measurements on the two feet from the half of the sow’s body where lameness was induced were analyzed with one foot labeled lame and the other labeled sound. The variables analyzed for each collection period were the mean weight placed on each foot, the interquartile range (QR), the 5th percentile of weight measurements (P5), the 95th percentile of weight measurements (P95), the standard deviation (SD), and the mode. Additionally, the skewness (SKEW) and kurtosis (KURT) of the weight distribution during the collection period was calculated.

Sow lameness was scored using a visual analog scale (Quinn et al., 2007). Scorers were asked to indicate degree of sow lameness on a 10 cm line with 0 cm being completely sound with no lameness indications or signs and 10 cm being completely lame and non-weight bearing. The scores were recorded using the millimeter distance from 0. All 24 sows were scored each day by at least two scorers; scorers were aware of the sow’s lameness stage relative to the injection day. This was unavoidable due to personnel availability.

Additionally, each sows’ gait was evaluated while walking using the GaitFour walkway system on DPI +1 and +6. Gait measurements captured while the sow was walking were recorded for three walking events each day. The active walkway was 0.76 × 4.27-m.
The sow walked across the GaitFour walkway system repeatedly until 3 acceptable walking events occurred. A walking event was considered acceptable if the sow maintained a fluid motion while walking across the walkway without running and/or stopping at any point during the recording process. The walking gait measurements used in the analysis were stride length (cm, STRL), stance time (s, STAT), stride time (s, STRT), maximum pressure (kg, MP), and number of sensors (NS). Stride length was defined as the distance between each consecutive step made by the same foot. Stance time was the amount of time a foot activates a sensor on a given step. Stride time was defined as the time between two successive steps of the same foot. Maximum pressure was defined as the maximum weight placed on a given foot each step. Number of sensors was defined as the number of sensors activated by a step from a given foot. Sensors collected data at 120 hertz.

When evaluating weight distributions and visual lameness scores, a mixed model analysis (SAS, Cary, NC) was conducted to determine at what day the measurements were no longer able to detect lameness in the sows relative to injection day (P>0.05). The model used can be written as:

\[ y = Xb + Zu + e \]

where \( y \) is the vector of observations, \( b \) is the vector of fixed effects, \( f \) is the vector of random contemporary group effects, \( u \) is the additive genetic animal effects, \( e \) is the vector of residual effects, and \( X \) and \( Z \) are known incidence matrices. When analyzing visual lameness data, the interaction between the injected feet pair (front or rear) and day relative to injection was fitted as the fixed effect (\( b \)), and sow and scorer within date were fitted as random effects (\( u \)). When the weight distribution data were evaluated, the 3-way interaction between injection half, day relative to injection, and lameness status was fitted as the fixed effect in
the model \( b \), and sow was fitted as a random effect \( u \). To analyze the gait measurements captured while the sows were walking, the analysis model included the 3-way interaction between half injected, day relative to injection, and lameness status as a fixed effect \( b \) and sow within day relative to injection and lameness status as a random effect \( u \). Since sows completed multiple walks each day, walk within a day was considered to be a repeated measurement for each sow.

A classification tree analysis was performed using the rpart package in R (Terneau et al., 2013). In the classification tree analysis, the distribution for each variable is examined for each classification, lame and sound for this study. Variables where the distributions for each classification (i.e. lame or sound) do not overlap explain the largest proportion of variation between classifications. Recursive partitioning is used to determine which variables provide the most informative division for predicting the classification category for each observation. Nodes or decision tests are created to define the variable threshold values within each test for each lame or sound classification. The threshold values determine at what point or value the animal is declared lame.

The randomForest package (Liaw et al., 2013) was used for a random forest analysis to classify feet as lame or sound. In the random forest analysis, multiple classification trees are created from the data. In this study, 1,000 trees were generated. The variables used in the greatest proportion of the trees created are considered to be the most informative for prediction. The response variable in both analyses was foot status (lame or sound). The relative importance of each variable in the random forest analysis was evaluated and compared with the variables used in the classification tree. The mean decrease in accuracy
quantifies the reduction in error rate observed when a variable is included in the decision tree and was used to measure variable importance in classifying observations as sound or lame.

Multiple analyses were performed where the weight distribution measurements were only included as potential variables in the tree, the walking gait measurements were only included and both measurements types were included. For the walking gait measurements, average of the sows’ three daily walking events was used in the analysis. When developing the decision tests, the measurements involving the weight and pressure a sow placed on each foot were expressed as a percentage of the sow’s total body weight. Cross-validation was used to determine the predictive ability of the classification trees. This involved removing a single observation from the analysis and predicting the observation based on the remaining observations. This process continued until all observations were removed and predicted by the remaining data. Error rates were calculated as the number of incorrect classifications for the observation predicted using the remaining observations.

**Results**

Least squares means for the subjective and objective measurements for sows injected in a rear or front toe are seen in Tables 1 and 2, respectively, for weight distribution measurements. The least squares means for the walking gait measurements are shown in Table 3. The visual analog scale scores for DPI +1 through +6 were significantly different (P<0.05) from the sound day visual score when a rear or front toe was lame. At DPI +7, the visual score was not significantly different (P>0.05) from the sound day score regardless of which body half (front or rear) was injected. In general, the weight distribution measurements were significantly different when compared to base line day up to DPI +7 and +4 when a rear and front foot were injected, respectively. The gait measures taken while the
sows were walking were significantly different (P<0.05) the day immediately following injection; in general, the measurements were resolved (P>0.05) by DPI +6 regardless of which body half the foot was from (front or rear).

The classification trees generated using the force plate measurements on DPI +1 are shown in Figures 3.1a and 3.1b, for rear and front feet injected, respectively. Regardless of which body half (front or rear) was injected, the only variable in the tree was mean total body weight percentage placed on the foot. This indicates that, on average and as expected, the sows placed less weight on the lame foot compared to the sound foot. The larger threshold for the front foot injected tree is a result of the weight a sow places on her front legs compared to her rear legs.

The classification trees using the GaitFour measurements on DPI +1 are shown in Figures 3.2a and 3.2b for rear and front feet injected, respectively. Both trees have a single node with maximum pressure as a percentage of total body weight being the decision variable. This is in agreement with the decision variable in the weight distribution trees. When a classification tree was developed using both the weight distribution information and the walking gait information, the resulting tree used the same decision variables and threshold values as the weight distribution classification tree for DPI +1 and +6 regardless of which body half (front or rear) was injected. This indicates that the weight distribution measurements have higher predictive ability to detect lameness compared to the walking gait measures.

The error rates for the weight distribution and gait classification trees by day relative to lameness induction are shown in Tables 4 and 5. The weight distribution trees were able to classify the lame and sound feet with less than 5% error the first three days following
lameness induction when a rear foot was injected. However, it is more important to detect lameness several days after lameness induction when clinical signs may not be as readily apparent and may be more like the onset for a lameness challenge that is just beginning to occur on a breeding female in a commercial herd setting. The cross-validation error rate was less than 50% up to DPI +6 and DPI +8 when a rear foot was injected and up to DPI +3 and DPI +6 and +7 when a front foot was injected. By DPI +9 for rear injected sows and DPI +8 for front injected sows, the classification trees were not able to accurately predict any of the observations when the observation was not included in the development of the tree. Because the treatments used in the present study involved injecting one foot (i.e. left or right) within the sow’s two foot pairs (i.e. front or rear feet), the maximum error rate based on classifying individual feet as lame or sound would be 50%. In this case, the classification tree would have only a single node and all observations would be given the same classification. The errors rates presented in this study are associated with the cross-validation method described above and thus, are based on the number of correctly classified observations, when the observation was not included in the tree development. Therefore, the maximum error rate is 100%, meaning that no observations were correctly classified.

In general, the weight distribution trees had a greater predictive ability to detect lameness compared to the gait measures captured on the sows while walking. This could be due to the longer time associated with collecting the weight distribution measurements when on the force plate compared to collecting the gait measures on the walkway system.

The mean decrease in accuracy for each variable included in the analysis for the force plate and GaitFour measurements are in Tables 4 and 5, respectively. Based on the results seen in Table 4, it is clear that the mean total body weight percentage and the 5th and 95th
percentiles for total body weight percentage placed on each foot were consistently important variables when classifying lameness regardless of which body half (front or rear) was injected. The results in Table 5 indicate that maximum pressure as a percentage of total body weight is most informative for detecting lameness regardless of which body half (front or rear) was injected. For most of the variables, the mean decrease in accuracy became negative at day +6, indicating a decrease in the predictive ability for the gait measures captured while the sows were walking.

**Discussion**

There are numerous methodologies that can be employed to subjectively and objectively measure and evaluate the relative degree of lameness for an individual animal at a given point in time. Subjective lameness identification systems are designed to categorize lameness expressed while the animal is walking and have been developed for cows (Manson and Leaver, 1988), dogs (Quinn et al., 2007), sheep (Welsh et al., 1993; Kaler et al. 2009), horses (Keegan et al., 2010), and finishing pigs (Main et al., 2000; Rothschild and Christian, 1988). The scoring systems used in the livestock industries have been implemented so that caretakers can quickly and affordably quantify lameness prevalence in the herd on any particular day. However, there can be disagreement between the lameness score assigned to an individual animal (Flower and Weary, 2006). This disagreement is the result of either inter- or intra-scorer variation, meaning that different scorers may provide different scores or that the same individual may provide different scores when scoring the same animal twice. An objective and standardized method for assigning lameness scores to an animal would likely be more accurate when compared to subjective scoring measures and provide
producers with a useful tool to assess lameness, resulting in more timely identification and treatment of lame sows.

One such method that shows promise is the force plate measurement system. This device quantifies the amount of force each limb applies to the assessment tool surface (Pastell et al., 2008). Force plate measurement systems can measure variables that have been associated with objectively classifying structural abnormalities into lameness scores. An animal will distribute less weight on the limb(s) that is(are) painful or structurally unsound (Corr et al., 2003). The use of such equipment has been evaluated in other species such as dogs (Evans et al., 2005), chickens (Corr et al., 2003), and dairy (Pastell and Kujala, 2007).

In this study, the results from the visual analog scale indicated that scorers were able to distinguish a difference between the base line sound animals and lame animals up to DPI +7 and +9 for sows injected in the rear and front feet, respectively. Since this was not a blind study, visual scores may be biased by scorer’s knowledge of the sow’s lameness stage. This bias would place an advantage to the visual lameness identification, which suggests that the objective measurements used in this study are performing up to a level comparable to the best possible detection level of the visual identification method. A similar bias would be expected if a numerical rating system had been used. To evaluate this visual lameness scoring method’s ability to distinguish between lame and sound animals in an unbiased manner, a blind study should be conducted. However, the objective of this study was to demonstrate that the objective measurements have the ability to detect lameness similar to the visual analog scale. Based on the bias of the visual scores, the observed differences between the visual scoring and objective measures would be conservative estimates of the true differences. In reality, lameness scorers would have no prior knowledge of the sow
lameness status, likely making the differences between the visual scores and the objective measures as large as or larger than the differences detected in this study.

The objective force plate measures were able to detect a difference between baseline sound animals and lame animals up to DPI +7 and +4 for rear and front toes injected, respectively. While the force plate measurements were not able to detect lameness for a longer time period after lameness induction compared to the visual identification method, the force plate was still able to detect lameness, and implementing a lameness diagnosis algorithm using the force plate would not require the training time and effort that is associated with using visual appraisal. If the visual scorers did not have prior knowledge of the sows’ lameness, the force plate measures may have outperformed the visual identification method when detecting sow lameness.

For the GaitFour, all measurements resolved by DPI +6 with the exception of stride length when injection occurred in the sow’s front toe. This may indicate a need to evaluate the GaitFour measurements for a longer time period post-lameness induction to determine when stride length resolves. Since stride length can be observed by an individual, this may be a measurement that does not resolve before the visual scores indicate that lameness has resolved. Grégoire and co-workers (2013) reported that lame sows tend to have shorter stride lengths, walk slower, and have a longer stance time when compared to non-lame sows which agrees with the current study. Additionally, boars selected for poor front leg structure had significantly short stride length compared to boars selected for desirable front leg structure (Morrow et al., 1991).

The results of this study indicate that the objective tools in this analysis can detect lameness and provide a way to classify lame and sound animals, and the measurements taken
using the GaitFour and force plate change with varying lameness severity. Understanding which measurements are most important when classifying lameness will allow for development of an algorithm to detect lameness based on objective measurements. Using objective measurements will remove the differences between scorers and provide a more uniform method to detect sow lameness.

The decisions tests developed from the weight distribution measurements had a lower error rate when sow lameness was more severe when compared to the decision tests developed from the gait measurements. However, the error rates converged as lameness severity decreased. The classification trees associated with the gait measurements used similar decision tests as the weight distribution trees, but have higher error rates. Since the combined classification tree with both types of measures only used weight distribution measurements as decision variables, there is evidence that the weight distribution measurements are more informative than the walking gait measures.

Comparing day-to-day weight distribution differences could allow for detecting lameness in sows. This could account for the variation between sows. However, if a prediction method is implemented into a new herd, there would not be any prior information at the time of implementation. A baseline value would need to be established for each sow before lameness could be detected, which means sows would need to be sound when introduced to the force plate system.

The results of this study indicate that this measuring system was able to identify sow lameness by separately measuring the weight each sow is willing to bear on each leg. Using this device in conjunction with the lameness detection tree created from this study can aid swine producers in detecting lame sows, and thus, improve management decisions in relation
to lame animals. The force plate could be incorporated into an electronic feeding system where sow lameness could be monitored on a daily basis without human interaction. More timely lameness detection could result in more treatable lameness, less mortality and euthanasia due to lameness, improved productivity, less treatment expense, and greater production efficiency and profitability for the sow operation.

References


Knauer MT 2006. Factors influencing sow longevity. Thesis (M.S.), Iowa State University, Ames, Iowa, USA.


Table 3.1. Subjective visual scores and objective weight distribution measures least squares means (standard error) from a study where lameness was induced in the sow’s rear feet.

| Status | Tool | Measure | Day relative to lameness induction | -1 | +1 | +2 | +3 | +4 | +5 | +6 | +7 | +8 | +9 |
|--------|------|---------|----------------------------------|----|----|----|----|----|----|----|----|----|----|----|
|        | VAS  | Score   |                                  | 0.4| 68.4| 55.0| 33.2| 24.9| 13.8| 8.1 | 6.2 | 4.7 | 3.8 |
| Lame   | FP   | Mean    |                                  | 41.8| 24.3| 26.5| 30.1| 34.7| 36.0| 36.9| 39.1| 39.7| 41.2 |
|        | PK  | Measure |                                  | (2.1)a | (2.1)b | (2.2)c | (2.1)c | (2.2)d | (2.1)de | (2.1)c | (2.3)ef | (2.4)c | (3.4)ef |
|        | QR  |         |                                  | 11.2| 22.3| 16.9| 17.0| 15.4| 13.7| 12.8| 14.4| 12.1| 10.6 |
|        | P5  |         |                                  | 22.9| 0.3| 2.4| 5.3| 10.7| 13.0| 15.5| 17.3| 19.7| 22.6 |
|        | P95 |         |                                  | 60.9| 46.2| 46.2| 51.4| 54.7| 55.4| 55.3| 58.4| 58.1| 59.0 |
|        | SD  |         |                                  | 11.9| 14.8| 13.4| 14.0| 13.4| 13.1| 12.4| 13.0| 12.1| 11.4 |
|        | SKEW|         |                                  | (0.13)a | (0.13)b | (0.13)c | (0.13)d | (0.13)e | (0.13)f | (0.13)g | (0.16)h | (0.16)i | (0.27)j |
| Sound  | FP  | Mean    |                                  | 41.3| 50.0| 50.0| 49.3| 47.9| 46.0| 44.7| 45.0| 43.9| 43.2 |
|        | PK  | Measure |                                  | (2.1)a | (2.1)b | (2.2)c | (2.1)c | (2.2)d | (2.1)de | (2.1)ef | (2.3)fg | (2.4)c | (3.4)fg |
|        | QR  |         |                                  | 11.3| 20.0| 16.8| 15.9| 15.1| 13.0| 12.8| 12.5| 11.9| 10.3 |
|        | P5  |         |                                  | 22.1| 27.9| 28.6| 27.2| 27.4| 27.2| 27.1| 27.7| 26.0| 26.4 |
|        | P95 |         |                                  | 60.0| 74.5| 73.0| 72.6| 70.8| 68.0| 65.4| 66.2| 64.0| 61.1 |
|        | SD  |         |                                  | (0.13)a | (0.13)b | (0.13)c | (0.13)d | (0.13)e | (0.13)f | (0.13)g | (0.16)h | (0.16)i | (0.27)j |
|        | SKEW|         |                                  | (0.13)a | (0.13)b | (0.13)c | (0.13)d | (0.13)e | (0.13)f | (0.13)g | (0.16)h | (0.16)i | (0.27)j |

*The model used for each measurement is described in the text.

*Indicates the status of the foot measured. The 24 sows were injected with 10mg amphotericin B in the distal inter-phalangeal joint according to the methods outlined in Karraker et al. (2013).

*Tool indicates which lameness detection tool was used: VAS (visual analog scale) and FP (force plate).

*Indicates which measurement was used: Mean (mean weight, kg. placed on the foot), QR (interquartile range of the weight, kg. placed on each foot), P5 (5th percentile of weight, kg. placed on each foot), P95 (95th percentile of weight, kg. placed on each foot), SD standard deviation of weight, kg. placed on each foot), SKEW (skewness of weight distribution for each foot), and KURT (kurtosis of weight distribution of each foot).
Table 3.2. Subjective visual scores and objective weight distribution measures least squares means (standard error) from a study where lameness was induced in the sow’s front feet\(^1\)

<table>
<thead>
<tr>
<th>Status(^2)</th>
<th>Tool(^3)</th>
<th>Measure(^4)</th>
<th>Day relative to lameness induction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>-1</td>
<td>+1</td>
</tr>
<tr>
<td>Lame</td>
<td>FP</td>
<td>VAS</td>
<td>Score</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3.1) (^b)</td>
<td>(3.4) (^c)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>58.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.1) (^a)</td>
<td>(2.1) (^b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>QR</td>
<td>12.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.3) (^a)</td>
<td>(1.3) (^b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P5</td>
<td>34.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.0) (^a)</td>
<td>(2.0) (^b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P95</td>
<td>78.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3.0) (^a)</td>
<td>(3.0) (^b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>13.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.8) (^a)</td>
<td>(0.8) (^b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SKEW</td>
<td>-0.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.13) (^a)</td>
<td>(0.13) (^b)</td>
</tr>
<tr>
<td>Sound</td>
<td>FP</td>
<td>VAS</td>
<td>4.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.6) (^a)</td>
<td>(0.6) (^b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>58.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.1) (^a)</td>
<td>(2.1) (^b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>QR</td>
<td>12.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.3) (^a)</td>
<td>(1.3) (^b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P5</td>
<td>38.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3.0) (^a)</td>
<td>(3.0) (^b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>13.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.8) (^a)</td>
<td>(0.8) (^b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SKEW</td>
<td>-0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.13) (^a)</td>
<td>(0.13) (^b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KURT</td>
<td>5.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.6) (^a)</td>
<td>(0.6) (^b)</td>
</tr>
</tbody>
</table>

\(^{1}\)The model used for each measurement is described in the text.

\(^{2}\)Indicates the status of the foot measured. The 24 sows were injected with 10mg amphotericin B in the distal inter-phalangeal joint according to the methods outlined in Karriker et al. (2013).

\(^{3}\)Tool indicates which lameness detection tool was used: VAS (visual analog scale) and FP (force plate).

\(^{4}\)Indicates which measurement was used: Mean (mean weight, kg, placed on the foot), QR (interquartile range of the weight, kg, place on each foot), P5 (5\(^{th}\) percentile of weight, kg, placed on each foot), P95 (95\(^{th}\) percentile of weight, kg, placed on each foot), SD standard deviation of weight, kg, placed on each foot), SKEW (skewness of weight distribution for each foot), and KURT (kurtosis of weight distribution of each foot).
### Table 3.3. Subjective visual scores and objective walking gait measures least squares means (standard error) from a study where lameness was induced in the sow’s feet

<table>
<thead>
<tr>
<th>Half Injected</th>
<th>Status</th>
<th>Tool</th>
<th>Measure</th>
<th>Day relative to lameness induction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>-1</td>
<td>+1</td>
</tr>
<tr>
<td>Rear</td>
<td>Lame</td>
<td>VAS</td>
<td>Score</td>
<td>0.4 (3.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GR</td>
<td>STRT</td>
<td>0.38 (0.03)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>STAT</td>
<td>0.22 (0.02)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>STRL</td>
<td>100.8 (2.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MP</td>
<td>51.8 (1.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NS</td>
<td>26.9 (0.8)</td>
</tr>
<tr>
<td></td>
<td>Sound</td>
<td>GR</td>
<td>STRT</td>
<td>0.38 (0.03)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>STAT</td>
<td>0.22 (0.02)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>STAL</td>
<td>100.9 (2.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MP</td>
<td>53.4 (1.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NS</td>
<td>27.4 (0.8)</td>
</tr>
<tr>
<td>Front</td>
<td>Lame</td>
<td>VAS</td>
<td>Score</td>
<td>-0.9 (3.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GR</td>
<td>STRT</td>
<td>0.37 (0.03)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>STAT</td>
<td>0.23 (0.02)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>STRL</td>
<td>105.4 (2.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MP</td>
<td>65.6 (1.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NS</td>
<td>31.6 (0.8)</td>
</tr>
<tr>
<td></td>
<td>Sound</td>
<td>GR</td>
<td>STRT</td>
<td>0.37 (0.03)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>STAT</td>
<td>0.23 (0.02)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>STRL</td>
<td>106.0 (2.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MP</td>
<td>67.7 (1.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NS</td>
<td>32.6 (0.8)</td>
</tr>
</tbody>
</table>

1. The model used for each measurement is described in the text.
2. Indicates which body half was injected.
3. Indicates the status of the foot measured. The 24 sows were injected with 10mg amphotericin B in the distal inter-phalangeal joint according to the methods outlined in Karriker et al. (2013).
4. Indicates which lameness detection tool was used: VAS (visual analog scale) and GR (GaitFour walkway system).
5. Indicates which measurement was used: STRT (stride time, s), STAT (stance time, s), STRL (stride length, cm), MP (maximum pressure, kg), and NS (number of sensors).
Figure 3.1a-b. Classification tree to detect lameness at day 1 post-injection using measurements collected from the microcomputer-based embedded force plate system to detect induced sow lameness

The 24 sows were injected with 10mg amphotericin B in the distal inter-phalangeal joint at one of four injection sites (left front toes, right front toes, left rear toes, and right rear toes) according to the methods outlined in Karriker et al. (2013). Variables included in analysis were Mean (mean total body weight percentage placed on the foot), QR (interquartile range of the body weight percentage place on each foot), P5 (5th percentile of body weight percentage placed on each foot), P95 (95th percentile of body weight percentage placed on each foot), SD (standard deviation of body weight percentage placed on each foot), SKEW (skewness of body weight percentage distribution for each foot), and KURT (kurtosis of weight distribution of each foot). If the statement at the node is true this tree directs to the left branch, otherwise the tree directs to the right branch. The classifications are the leaves at the bottom of the branches.
Figure 3.2a-b. Classification tree to detect lameness at day 1 post-injection using measurements collected from the GaitFour walkway system to detect induced sow lameness$^1$

![Classification tree]

a. Lameness induced in rear foot  

b. Lameness induced in front foot  

$^1$The 24 sows were injected with 10mg amphotericin B in the distal inter-phalangeal joint at one of four injection sites (left front toes, right front toes, left rear toes, and right rear toes) according to the methods outlined in Karriker et al. (2013). Variables included in analysis were STRT (stride time, s), STAT (stance time, s), STRL (stride length, cm), MP (maximum pressure as a percentage of total body weight), and NS (number of sensors). If the statement at the node is true this tree directs to the left branch, otherwise the tree directs to the right branch. The classifications are the leaves at the bottom of the branches.
Table 3.4. Classification tree error rates and mean decrease in accuracy for each weight distribution variable used to detect induced sow lameness

<table>
<thead>
<tr>
<th>Half Injected(^2)</th>
<th>DPI(^3)</th>
<th>Error Rate(^4)</th>
<th>Mean</th>
<th>SD</th>
<th>SKEW</th>
<th>KURT</th>
<th>P5</th>
<th>P95</th>
<th>QR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rear</td>
<td>+1</td>
<td>4.2%</td>
<td>20.2</td>
<td>-0.3</td>
<td>0.3</td>
<td>9.7</td>
<td>2.6</td>
<td>21.9</td>
<td>3.7</td>
</tr>
<tr>
<td></td>
<td>+2</td>
<td>2.5%</td>
<td>25.2</td>
<td>0.2</td>
<td>1.5</td>
<td>5.1</td>
<td>21.0</td>
<td>20.7</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>+3</td>
<td>0.0%</td>
<td>28.3</td>
<td>0.2</td>
<td>3.3</td>
<td>1.8</td>
<td>18.5</td>
<td>23.6</td>
<td>-1.4</td>
</tr>
<tr>
<td></td>
<td>+4</td>
<td>21.1%</td>
<td>21.9</td>
<td>-1.3</td>
<td>8.5</td>
<td>1.1</td>
<td>17.6</td>
<td>19.3</td>
<td>3.8</td>
</tr>
<tr>
<td></td>
<td>+5</td>
<td>41.3%</td>
<td>19.0</td>
<td>0.1</td>
<td>8.4</td>
<td>-4.8</td>
<td>21.5</td>
<td>21.2</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>+6</td>
<td>31.3%</td>
<td>16.6</td>
<td>3.7</td>
<td>13.7</td>
<td>7.7</td>
<td>16.3</td>
<td>17.1</td>
<td>7.0</td>
</tr>
<tr>
<td></td>
<td>+7</td>
<td>56.7%</td>
<td>8.5</td>
<td>-2.1</td>
<td>8.1</td>
<td>0.4</td>
<td>13.6</td>
<td>10.7</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>+8</td>
<td>42.9%</td>
<td>0.2</td>
<td>0.7</td>
<td>0.4</td>
<td>-5.0</td>
<td>0.1</td>
<td>-1.5</td>
<td>-6.7</td>
</tr>
<tr>
<td></td>
<td>+9</td>
<td>100.0%</td>
<td>-1.7</td>
<td>-2.6</td>
<td>-1.8</td>
<td>-5.9</td>
<td>-4.2</td>
<td>-1.1</td>
<td>-3.8</td>
</tr>
<tr>
<td>Front</td>
<td>+1</td>
<td>10.9%</td>
<td>22.8</td>
<td>-0.4</td>
<td>3.1</td>
<td>1.3</td>
<td>2.2</td>
<td>22.9</td>
<td>-2.8</td>
</tr>
<tr>
<td></td>
<td>+2</td>
<td>16.7%</td>
<td>20.4</td>
<td>1.9</td>
<td>17.3</td>
<td>-0.8</td>
<td>16.6</td>
<td>25.7</td>
<td>-0.7</td>
</tr>
<tr>
<td></td>
<td>+3</td>
<td>40.9%</td>
<td>19.3</td>
<td>-3.4</td>
<td>12.6</td>
<td>-3.8</td>
<td>15.9</td>
<td>19.0</td>
<td>-0.4</td>
</tr>
<tr>
<td></td>
<td>+4</td>
<td>71.1%</td>
<td>7.1</td>
<td>-5.5</td>
<td>0.3</td>
<td>-0.2</td>
<td>7.7</td>
<td>7.3</td>
<td>-3.6</td>
</tr>
<tr>
<td></td>
<td>+5</td>
<td>65.2%</td>
<td>9.3</td>
<td>-3.7</td>
<td>-1.4</td>
<td>-3.4</td>
<td>7.2</td>
<td>7.3</td>
<td>-2.7</td>
</tr>
<tr>
<td></td>
<td>+6</td>
<td>29.2%</td>
<td>9.7</td>
<td>0.8</td>
<td>5.9</td>
<td>-3.2</td>
<td>2.4</td>
<td>7.9</td>
<td>-0.4</td>
</tr>
<tr>
<td></td>
<td>+7</td>
<td>46.2%</td>
<td>0.8</td>
<td>-4.5</td>
<td>2.8</td>
<td>-5.2</td>
<td>-3.3</td>
<td>1.1</td>
<td>-1.2</td>
</tr>
<tr>
<td></td>
<td>+8</td>
<td>100.0%</td>
<td>5.7</td>
<td>-5.3</td>
<td>-2.0</td>
<td>-8.6</td>
<td>-4.8</td>
<td>5.0</td>
<td>-2.7</td>
</tr>
<tr>
<td></td>
<td>+9</td>
<td>100.0%</td>
<td>-0.3</td>
<td>-6.1</td>
<td>-4.3</td>
<td>-4.7</td>
<td>-6.1</td>
<td>0.7</td>
<td>-6.1</td>
</tr>
</tbody>
</table>

\(^1\)The 24 sows were injected with 10mg amphotericin B in the distal inter-phalangeal joint at one of four injection sites (left front toes, right front toes, left rear toes, and right rear toes) according to the methods outlined in Karriker et al. (2013). Variables included in analysis were mean – mean total body weight percentage placed on the foot, QR – interquartile range of the body weight percentage placed on each foot, P5 – 5th percentile of body weight percentage placed on each foot, P95 – 95th percentile of body weight percentage placed on each foot, SD – standard deviation of body weight percentage placed on each foot, SKEW – skewness of body weight percentage distribution for each foot, and KURT – kurtosis of weight distribution of each foot.

\(^2\)Indicates in which body half (front or rear) the sow was injected.

\(^3\)Days post-injection

\(^4\)The error rate is the percent of misclassifications when using the decision tree generated.

\(^5\)The mean decreases in accuracy for each variable should be compared within row and ranked from largest to smallest.
### Table 3.5. Classification tree error rates and mean decrease in accuracy for each walking gait measurement used to detect induced sow lameness

<table>
<thead>
<tr>
<th>Half Injected&lt;sup&gt;2&lt;/sup&gt;</th>
<th>DPI&lt;sup&gt;3&lt;/sup&gt;</th>
<th>Error Rate&lt;sup&gt;4&lt;/sup&gt;</th>
<th>Stance Time</th>
<th>Stride Time</th>
<th>Stride Length</th>
<th>Maximum Pressure</th>
<th>Number of Sensors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rear</td>
<td>+1</td>
<td>31.3%</td>
<td>17.8</td>
<td>1.2</td>
<td>-2.2</td>
<td>42.5</td>
<td>7.4</td>
</tr>
<tr>
<td></td>
<td>+6</td>
<td>29.8%</td>
<td>1.0</td>
<td>-6.2</td>
<td>-6.6</td>
<td>4.7</td>
<td>-4.7</td>
</tr>
<tr>
<td>Front</td>
<td>+1</td>
<td>30.4%</td>
<td>11.5</td>
<td>1.5</td>
<td>-6.9</td>
<td>46.3</td>
<td>15.3</td>
</tr>
<tr>
<td></td>
<td>+6</td>
<td>60.9%</td>
<td>-11.5</td>
<td>-9.3</td>
<td>-12.0</td>
<td>-1.8</td>
<td>0.5</td>
</tr>
</tbody>
</table>

<sup>1</sup>The 24 sows were injected with 10mg amphotericin B in the distal inter-phalangeal joint at one of four injection sites (left front toes, right front toes, left rear toes, and right rear toes) according to the methods outlined in Karriker <em>et al.</em> (2013). Variables included in analysis were stride time, stance time, stride length, maximum pressure as a percentage of total body weight, and number of sensors.  
<sup>2</sup>Indicates in which body half (front or rear) the sow was injected.  
<sup>3</sup>Days post-injection  
<sup>4</sup>The error rate is the percent of misclassifications when using the decision tree generated.  
<sup>5</sup>The mean decreases in accuracy for each variable should be compared within row and ranked from largest to smallest.
CHAPTER 4: GENETIC PARAMETER ESTIMATES AND RELATIVE ECONOMIC VALUES FOR PUREBRED AND CROSSBRED SOW LIFETIME PRODUCTIVITY

Modified from a paper to be submitted to *Journal of Animal Breeding and Genetics*

C. E. Abell\(^1\), R. L. Fernando\(^1\), T. V. Serenius\(^2\), M. F. Rothschild\(^1\), K. A. Gray\(^3\), and K. J. Stalder\(^1\)

\(^1\)Department of Animal Science, Iowa State University, Ames, IA 50011  
\(^2\)Figen Ltd., PO Box 319, 60101 Seinäjoki, Finland  
\(^3\)Smithfield Premium Genetics, 316 W. Charity Rd, Rose Hill, NC 28458

**Summary**

The breeding objective for a swine selection program is to improve crossbred commercial performance. Most genetic improvement programs are based on an assumed relationship between purebred performance in a nucleus herd and their relatives’ crossbred performance in a commercial herd. The objective of this study was to examine the relationship between purebred and crossbred sow longevity performance. Sow longevity was defined as a Bernoulli trait with a success occurring if a sow remained in the herd for a certain number of parities and including the cumulative number born alive as a means of reproductive success. Trait heritabilities were estimated using a frequentist and a Bayesian approach. Genetic correlations between the Bernoulli traits were calculated using Bayesian methods. Results indicated little to no genetic correlation between crossbred and purebred longevity traits for this population. Genetic correlations between cumulative number born alive at the nucleus level and the commercial level were significantly different from 0, suggesting that cumulative born alive at the nucleus level could be used as an indicator trait to improve crossbred sow productivity. Utilizing crossbred records in a selection program would be the best way to improve crossbred sow productivity.

**Key Words:** sow longevity, threshold model, categorical traits, economic value
Introduction

In a commercial swine breeding herd, sow longevity is a contributing factor to the operation’s overall success and profitability. A sow must remain in the herd between 3 and 4 parities to produce enough piglets before the investment reaches a positive net present value or “she pays for herself” (Stalder, 2000). Increasing the number of lifetime piglets produced per sow reduces the proportion of the sow’s gilt replacement and development costs that must be recovered by each pig. Because of this, a sow should not be voluntarily culled from the breeding herd as long as she is producing at an acceptable level or level equal to the herd average, in terms of litter size and weight and no animal wellbeing issues are present (too big for penning, injury, etc.).

Additionally, removing females in early parities does not allow a swine operation to benefit from the increased sow and downstream grow-finish litter performance compared to the gilt’s productivity and that from her litter’s performance. Not only do sows tend to have larger litters compared to gilts (Roehe and Kennedy, 1995), piglets from sow litters tend to have decreased mortality and improved or superior performance through the nursery and grow-finish phases when compared to piglets from gilt litters (Carney-Hinkle et al., 2013). The performance difference value between the offspring from gilt litters when compared to the offspring from older parity sows will likely dwarf the value of additional parities based on sow performance only (Lowe, 2013). When comparing the replacement gilt to the older sows in a commercial breeding herd, the gilt’s genetic advantage is not sufficient to cover the variable costs associated with gilt development until the sow has reached at least parity 7 under maximum genetic gain conditions (Abell et al., 2010). Using more realistic genetic
improvement and generation interval values, the optimal culling time is more realistically the 10\textsuperscript{th} or 11\textsuperscript{th} parity.

The breeding objective for a swine nucleus selection program is to improve crossbred commercial performance. Most genetic improvement programs are based on an assumed relationship between purebred performance in a nucleus herd and their relatives’ crossbred performance in a commercial herd. In addition to the sows’ genetic makeup differences (i.e. purebred versus crossbred) in nucleus and commercial herds, there are differences in management practices between the two types of herds (purebred and commercial). Sow longevity is an economically important trait at the commercial level; however, nucleus sows typically produce for fewer parities due to the desire to shorten the generation interval in order to increase the rate of genetic improvement. Removing sows at the nucleus level before they have fully expressed their lifetime potential does not allow for direct selection on longevity in the nucleus herd. Since nucleus animals do not have phenotypic records for lifetime measurements, an indicator trait expressed at an earlier parity would be necessary to select for longevity. If 2\textsuperscript{nd} parity reproductive performance among purebred sows in a breeder’s herds is related to the reproductive performance at later parities among crossbred sows in commercial sow herds, this would suggest that selection on nucleus animals after only 2 parities can have a positive impact on improving crossbred longevity.

The objective of this study was to examine the relationship between purebred and crossbred reproductive performance using whether or not a sow remained in the herd up to a certain number of parities and cumulative born alive, or the sum total number born alive for a sow across parities, as the reproductive traits of interest. Defining the relationship between purebred and crossbred sow longevity will allow genetic suppliers to improve selection
methods for increased longevity and downstream grow-finish performance at the commercial level.

**Materials and Methods**

Performance records for 11,506 purebred sows and 12,897 crossbred sows were evaluated for this study. Purebred sows were from a Landrace pure line, and the crossbred sows were F1 offspring from a cross between a Landrace pure line and a Large White pure line. Purebred sow information was obtained from 5 nucleus herds and crossbred sow performance information was from 2 herds. There were 1,039 and 213 sires for the purebred and crossbred sows, respectively. Of the 213 sires used at the crossbred level, 205 had dams with records at the purebred level, providing genetic ties between the purebred and crossbred herds. Purebred records were from sows first farrowing in February 1993 through May 2011. Crossbred sows first farrowed in November 2004 to May 2011. Sows that were active or that were continuing to produce (not yet completed their lifetime record) in the herd were also included in the analysis, but not treated as censored data. Previous work has shown that if the number of completed records overwhelms the number of censored records, there is little re-ranking among the breeding value estimations (Engblom et al., 2010). Pedigree information for at least three generations was known for each individual with performance records. The removal parity distributions for the purebred and crossbred sows are in Table 4.1. It is clear that crossbred sows remained in the herd for more parities on average compared to purebred sows.

Sow longevity traits were defined based on the sow’s removal parity. At the nucleus level, sows were considered successful if they completed parity 2 (farrowed and weaned their second litter). This trait is called PL2. At the commercial level, four degrees of success
(Parity 2, Parity 3, Parity 4, and Parity 5) were analyzed as separate traits. These traits are called CL2, CL3, CL4, and CL5, respectively. Additionally, cumulative number born alive traits were defined for purebred and crossbred sows. For purebred sows, cumulative number born alive up to parity two (PNBA2) and for a lifetime for purebred animals (PNBALF) were defined. For crossbred sows, cumulative number born alive up to parity 2 (CNBA2), parity 3 (CNBA3), parity 4 (CNBA4), parity 5 (CNBA5) and for a lifetime for crossbred animals (CNBALF) were defined. If results are similar between cumulative born alive measurements at different parities, earlier measurements could be used in a genetic evaluation to obtain a record earlier in an animal’s life and reduce the number of censored records in the analysis. Descriptive statistics for the cumulative born alive traits are shown in Table 4.2. The heritabilities for each trait and genetic correlations between the nucleus and commercial level successes were estimated.

A threshold model assuming an underlying normally distributed latent variable was used to calculate estimates for the longevity traits. Survival methods were not used since the number of censored records was overwhelmed by the number of complete records. Additionally, survival methods are more complex compared to linear models and most other production traits are evaluated using a linear model. A choice was made to continue with the linear model instead of moving to a different method like survival analysis. The estimates for the cumulative number born alive traits were calculated using a model for a normal distribution. The model used for the both the frequentist and Bayesian estimates can be written as:
\[ y = X\mu + Wf + Zu + e \]

where \( y \) is the vector of observations, \( \mu \) is the mean, \( f \) is the vector of random contemporary group effects, \( u \) is the additive genetic animal effects, \( e \) is the vector of residual effects, and \( X, W, \) and \( Z \) are known incidence matrices. Herd, year, and month of last farrowing were used to define the contemporary groups. Contemporary group was fitted as a random effect for all longevity and number born alive traits analyzed. When using a threshold model and Bayesian analysis for categorical traits, the predictive error is reduced if contemporary groups are fitted as random rather than fixed effects (Luo et al., 2001). All models included a random animal effect. To calculate genetic correlations, separate bivariate models were used to reduce computation time.

Heritability estimates were calculated using a frequentist (ASREML, VSN International, Hemel Hempstead, UK) and a Bayesian approach (Korsgaard et al., 2003). Heritability was defined as

\[ h^2 = \frac{\sigma^2_g}{\sigma^2_g + \sigma^2_f + \sigma^2_e} \]

where \( \sigma^2_g \) is the additive genetic variance, \( \sigma^2_f \) is the contemporary group variance, and \( \sigma^2_e \) is the residual variance. A restricted maximum likelihood approach using the average information matrix was utilized to obtain frequentist estimates. Using the average information requires the first derivative for the likelihood, but the second derivative is replaced by the average information matrix (Johnson and Thompson, 1994). Multiple starting values were used to ensure that convergence occurred at a global maximum. The default convergence criteria in ASREML were used for the frequentist method (Gilmour et al., 2005). Standard errors of heritabilities and genetic correlations were estimated using the
methods described in Lynch and Walsh (1998). Genetic correlations cannot be calculated between two threshold traits using ASREML, and thus, were only calculated between the cumulative number born alive traits and between one longevity and one number born alive trait. There are no other publicly available maximum likelihood software packages that can accurately estimate genetic correlations between two threshold traits.

For the Bayesian method, 60,000 iterations were computed with 20,000 used as burn-in. The remaining 40,000 iterations were used to calculate the heritability estimates and associated standard errors. These estimates were based on the empirical distribution of the Markov Chain. A priori, additive genetic animal effects were assumed to be normally distributed, \( N(\mathbf{0}, \mathbf{A} \otimes \mathbf{G}) \) where \( \mathbf{A} \) is the additive genetic relationship matrix and \( \mathbf{G} \) is the additive genetic (co)variance matrix for the two traits in the bivariate analysis. The residual effects were assumed to have an inverted Wishart distribution. The details for a multivariate analysis of categorical traits can be found in Korsgaard et al. (2003). A blocked Gibbs sampler (Garcia-Cortes and Sørensen, 1996) was developed using C++ programs to implement the multivariate analysis as described in Korsgaard et al. (2003). Due to the computation time associated with Bayesian methods, it is impractical to use Bayesian methods for analyses where accurate frequentist methods are available.

Using the genetic variance estimate for PNBALF, the relative economic importance for lifetime pigs produced compared to other economically relevant swine traits was determined. The other reproductive traits included in the analysis were number born alive (NBA), number weaned (NW), litter weaning weight (LW), litter birth weight (LBW), and wean to estrus interval (WTE). The production traits included in the analysis were days to 250 lbs. (D250), backfat (BF), and loin eye area (LEA). The genetic standard deviations for
these traits were obtained from a commercial seedstock supplier. The economic values per unit were $27.93 for a pig born alive, $38.57 for weaned pig, $0.62 per pound of live weight, $1.90 cost for a nonproductive day, and $0.22 cost per pig day in the finisher (NSR, 2010). The economic value for a day spent in the finisher includes facility costs and feed costs associated with the animal’s body maintenance (to allow the body to function, heart to beat, excrete waste products, etc. on a daily basis). The values for BF and LEA were $18.10 and $2.48 based on the change in carcass lean content with 1 unit change in BF and LEA, respectively (NSR, 2010). The standardized economic weight was calculated by multiplying each traits’ value per unit by the genetic standard deviation. The relative economic value for each trait was calculated as the standardized economic value for the trait divided by the sum total of the standardized economic values (Mwansa et al., 2002).

**Results**

The heritability estimates from the frequentist and Bayesian approaches are presented in Table 4.3. All estimates from either model were significantly different from zero, indicating that longevity is not completely controlled by environmental factors. Based on the animal model estimates, sow longevity would be considered a lowly heritable trait.

Heritability estimates for all of the crossbred traits were similar and do not suggest that any longevity definition would be better to incorporate into a genetic program compared to the other definitions based on heritability alone.

All heritability estimates from the Bayesian approach were significantly different from zero. However, the Bayesian estimates had a larger magnitude compared to the frequentist estimates. One reason this could occur is that Bayesian estimates are based on the mean of the posterior distribution while frequentist methods are based on the mode of the
distribution. Because of the skewness of the posterior distribution for the heritability, the mean and mode would not be the same, resulting in different heritability estimates from the frequentist and Bayesian approaches. The heritability estimates for the crossbred longevity traits increase as the parity they are associated with increases. This suggests that selecting for sows remaining in the commercial herd for at least 5 parities could be more valuable than selection for sows remaining in the commercial herd for fewer parities.

The genetic correlation estimates between purebred and crossbred measures are displayed in Table 4.4. The genetic correlations between the purebred and crossbred cumulative born alive traits were significantly different from 0, suggesting that purebred selection for number born alive can have a desirable impact on crossbred number born alive. However, there was little to no genetic correlation found between the purebred and crossbred longevity traits for this population. The only correlation significantly different from zero was the correlation between PL2 and CL5. These estimates indicate that selecting for crossbred longevity using purebred information alone would be difficult.

The relative economic weights for reproductive traits are shown in Table 4.5. Based on these results, it is clear that lifetime born alive had the greatest relative economic importance compared to the other economically relevant traits. Cumulative born alive had over 13 times greater relative emphasis compared to any other trait. This suggests sow productive lifetime is an economically important production trait for commercial pork producers. Seedstock suppliers should include sow productive lifetime in the commercial level breeding objective and emphasize sow productive lifetime or some other sow longevity related measure in their selection programs. Number born alive at a single litter is the second most relevant trait. In general, a 6 fold increase in relative emphasis was found for
reproductive traits compared to the terminal traits. Number born alive and number weaned were approximately 10 times more important when compared to litter weaning weight, wean to estrus interval, and litter birth weight, and 12 times more important than terminal traits.

**Discussion**

Improving crossbred performance is the goal for most swine breeders. However, most swine genetic companies make selection decisions based on purebred performance records and the genetic correlation between purebred and crossbred performance is not perfect (Cecchinato et al., 2010; McLaren et al., 1985; Wong et al., 1971). While the magnitude of the genetic correlation estimates between the cumulative born alive traits in this study were similar to correlation estimates between purebred and crossbred traits previously reported, the genetic correlations between longevity traits in the current study were lower in magnitude compared to other studies (Cecchinato et al., 2010; McLaren et al., 1985; Wong et al., 1971). The population structure in the current analysis could have prevented the detection of a genetic correlation. A structure where sires were used in both the purebred and crossbred herds would be more ideal to estimate genetic correlations; however, ideal data structures are not always available when analyzing field data. Additionally, the methods utilized in the study may not be sufficient in detecting a genetic association between the Bernoulli traits due to the dataset’s population structure.

The greater genetic correlation estimates between crossbred and purebred performance measures found in previous studies compared to the estimates from this study could be a result of the traits considered in the previous studies being easier to define and quantify. Longevity is based on culling criteria which can vary between the nucleus and commercial levels and from farm-to-farm. Many criteria used to make removal decisions are
subjective. The genetic correlations found in this study may indicate that longevity is a different trait at the purebred and crossbred levels due to the culling criteria used at each production level. Therefore, there may be few genes that have large associations with the different longevity definitions.

Previously reported literature estimates have shown that heritabilities and genetic correlations between traits in a selection index vary between crossbred and purebred populations (Louca and Robison 1967; Lutaaya et al., 2001; Stanislaw et al., 1967). The same is true for the results from this study. Since coefficients for selection indices are derived based on genetic correlations and trait heritabilities, these genetic parameters differences among purebred and crossbred populations can impact the selection index coefficients. If index weights are incorrect, selection would not be optimized and rate of genetic gain for the overall index would decrease. Since the breeding objective of most swine genetics companies is to improve crossbred performance, genetic gain would be optimized if crossbred records were used to make selection decisions at the purebred level.

Genetic gain for crossbred performance can be improved by estimating breeding values using both crossbred and purebred information. Ehlers et al. (2005) showed that heritabilities estimated by pooling crossbred and purebred data are similar to estimates from purebred data alone and crossbred data alone. Provided that there is a genetic correlation between the purebred and crossbred data, pooling data could be used to increase the accuracy of breeding value estimates through the added information. Ehlers et al. (2006) showed that the breeding value accuracy was increased when pooled data were used to estimate breeding values instead of purebred data alone.
In this study, over 40% of the purebred sows remained in the herd for greater than 2 parities. This number is inflated from what would be expected if a nucleus herd was trying to maximize genetic progress. Sows may remain in a nucleus herd due to a reduction in available gilts to enter the herd or fewer lower parity sows in the herd. This decreased supply of gilts and low parity sows could be a result from a health challenge to the system. Keeping animals in the herd that would have otherwise been removed could impact the heritability associated with sow longevity. Changing the selection criteria would result in changing the longevity definition, which affects the genetic parameters associated with the trait. For example, sows that survive a health challenge may not have the same characteristics as sows that have high reproductive performance in a herd with a high health status.

For this analysis, there were approximately equal numbers of purebred and crossbred sows due to the longer data collection period for purebred sows. This would be expected if a portion of the nucleus offspring were used to form a commercial test herd so that crossbred records can be incorporated into a genetic evaluation. Since this analysis presented evidence that there is little to no genetic correlation between purebred and crossbred longevity measures for this population, genetic companies should make use of maternal records from their crossbred sow population through the development of a commercial test herd or develop relationships with customers willing to keep meticulous records. In either case, these records can be used to augment the purebred records for more lowly heritable traits to improve accuracy and speed genetic progress.

This study has shown that the frequentist and Bayesian approaches can be used to estimateheritabilities for Bernoulli traits; however, the Bayesian approach was able to estimate genetic correlations between multiple threshold traits. This added benefit would
allow genetic suppliers to incorporate more than one threshold trait into a genetic evaluation. Bayesian methods require a burn-in period for the Markov chain to converge to the stationary distribution. Because of this and the number of iterations required for accurate estimates, Bayesian methods have a greater computational time compared to frequentist methods. Therefore, a frequentist approach should be used when available.

Cumulative born alive is a combination of length of productive life and litter size. Using the cumulative born alive estimates would be different from selecting for litter size alone, since sows producing for fewer parities would be penalized in the cumulative born alive trait. Thus, improving cumulative born alive would improve sow lifetime productivity rather than productivity at a single parity. The genetic correlations reported in this study suggest that there is a genetic correlation between cumulative born alive at the purebred and crossbred levels and that selection for cumulative born alive at the purebred level could improve crossbred sow lifetime productivity. Based on the results from this study, purebred records for cumulative born alive traits can be used as indicator traits to select for improved crossbred sow productivity; however, the short lifespan of nucleus animals may not allow for genetic companies to directly select for crossbred sow longevity using purebred performance due to the genetic correlation being near 0 as was the case for this population.

Because of the large relative emphasis that should be placed on lifetime born alive, genetic companies must focus on adding sow lifetime productivity to a selection program. One way to do this would be to implement a commercial test herd. With the commercial test herd, genetic companies could collect data from crossbred animals to be used in the genetic evaluations. Additionally, the records could be used to incorporate genome-enabled selection into a breeding program. Genome-enabled selection would improve the accuracy for traits,
such as lifetime productivity, that are measured late in life and difficult to measure on purebred animals.

References


Table 4.1. Removal parity distribution for crossbred and purebred sows in a study to evaluate the relationship between purebred and crossbred sow performance<sup>1</sup>

<table>
<thead>
<tr>
<th>Removal Parity</th>
<th>Crossbred Sows (%)</th>
<th>Purebred Sows (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20.8</td>
<td>36.8</td>
</tr>
<tr>
<td>2</td>
<td>14.1</td>
<td>20.3</td>
</tr>
<tr>
<td>3</td>
<td>10.9</td>
<td>17.1</td>
</tr>
<tr>
<td>4</td>
<td>13.9</td>
<td>14.3</td>
</tr>
<tr>
<td>5</td>
<td>13.0</td>
<td>8.8</td>
</tr>
<tr>
<td>6</td>
<td>14.6</td>
<td>1.9</td>
</tr>
<tr>
<td>7</td>
<td>9.1</td>
<td>0.6</td>
</tr>
<tr>
<td>8</td>
<td>3.3</td>
<td>0.3</td>
</tr>
<tr>
<td>9</td>
<td>0.3</td>
<td>0.1</td>
</tr>
</tbody>
</table>

<sup>1</sup>Purebred sows were from a Landrace pure line, and the crossbred sows were F1 offspring from a cross between a Landrace pure line and a Large White pure line.
Table 4.2. Descriptive statistics for purebred and crossbred cumulative born alive traits in a study to evaluate the relationship between purebred and crossbred sow performance.1

<table>
<thead>
<tr>
<th>Trait</th>
<th>Mean ± SE</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>PNBA2</td>
<td>17.4 ± 7.4</td>
<td>0</td>
<td>37</td>
</tr>
<tr>
<td>PNBALF</td>
<td>27.5 ± 19.5</td>
<td>0</td>
<td>132</td>
</tr>
<tr>
<td>CNBA2</td>
<td>21.6 ± 6.8</td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td>CNBA3</td>
<td>30.6 ± 12.2</td>
<td>0</td>
<td>59</td>
</tr>
<tr>
<td>CNBA4</td>
<td>38.1 ± 17.9</td>
<td>0</td>
<td>81</td>
</tr>
<tr>
<td>CNBA5</td>
<td>44.5 ± 22.9</td>
<td>0</td>
<td>93</td>
</tr>
<tr>
<td>CNBALF</td>
<td>50.2 ± 30.0</td>
<td>0</td>
<td>146</td>
</tr>
</tbody>
</table>

1 Purebred sows were from a Landrace pure line, and the crossbred sows were F1 offspring from a cross between a Landrace pure line and a Large White pure line.
2 PNBA2 and PNBALF represent the cumulative number born alive for purebred sows at parity 2 and for their lifetime. CNBA2, CNBA3, CNBA4, CNBA5, and CNBALF represent the cumulative number born alive for purebred sows at parity 2, parity 3, parity 4, parity 5, and for their lifetime.
Table 4.3. Frequentist and Bayesian heritability estimates (±SE) for purebred and crossbred sow longevity and cumulative born alive traits using an animal model\(^1\)

<table>
<thead>
<tr>
<th>Trait(^2)</th>
<th>Frequentist</th>
<th>Bayesian</th>
</tr>
</thead>
<tbody>
<tr>
<td>PL2</td>
<td>0.08 ± 0.01</td>
<td>0.61 ± 0.03</td>
</tr>
<tr>
<td>CL2</td>
<td>0.08 ± 0.01</td>
<td>0.15 ± 0.01</td>
</tr>
<tr>
<td>CL3</td>
<td>0.10 ± 0.01</td>
<td>0.18 ± 0.01</td>
</tr>
<tr>
<td>CL4</td>
<td>0.09 ± 0.01</td>
<td>0.20 ± 0.01</td>
</tr>
<tr>
<td>CL5</td>
<td>0.09 ± 0.01</td>
<td>0.29 ± 0.01</td>
</tr>
<tr>
<td>PNBA2</td>
<td>0.10 ± 0.01</td>
<td></td>
</tr>
<tr>
<td>PNBALF</td>
<td>0.19 ± 0.01</td>
<td></td>
</tr>
<tr>
<td>CNBA2</td>
<td>0.21 ± 0.02</td>
<td></td>
</tr>
<tr>
<td>CNBA3</td>
<td>0.23 ± 0.03</td>
<td></td>
</tr>
<tr>
<td>CNBA4</td>
<td>0.22 ± 0.03</td>
<td></td>
</tr>
<tr>
<td>CNBA5</td>
<td>0.18 ± 0.02</td>
<td></td>
</tr>
<tr>
<td>CNBALF</td>
<td>0.14 ± 0.02</td>
<td></td>
</tr>
</tbody>
</table>

\(^1\)Purebred sows were from a Landrace pure line, and the crossbred sows were F1 offspring from a cross between a Landrace pure line and a Large White pure line.

\(^2\)PL2, CL2, CL3, CL4, and CL5 are Bernoulli traits. For PL2, a success is considered to be a sow that remains in the nucleus herd long enough to farrow 2 litters. For CL2, CL3, CL4, and CL5, a success is considered to be a sow that remains in the nucleus herd long enough to farrow 2, 3, 4, and 5 litters, respectively. PNBA2 and PNBALF represent the cumulative number born alive for purebred sows at parity 2 and for their lifetime. CNBA2, CNBA3, CNBA4, CNBA5, and CNBALF represent the cumulative number born alive for purebred sows at parity 2, parity 3, parity 4, parity 5, and for their lifetime.
Table 4.4. Genetic correlation estimates (SE) between and crossbred sow longevity and cumulative born alive and a purebred longevity trait using an animal model\textsuperscript{1}

<table>
<thead>
<tr>
<th></th>
<th>PL2</th>
<th>CL2</th>
<th>CL3</th>
<th>CL4</th>
<th>CL5</th>
<th>PNBA2</th>
<th>PNBALF</th>
<th>CNBA2</th>
<th>CNBA3</th>
<th>CNBA4</th>
<th>CNBA5</th>
<th>CNBALF</th>
</tr>
</thead>
<tbody>
<tr>
<td>PL2</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.21</td>
<td></td>
<td>0.38</td>
<td>0.34</td>
<td>0.31</td>
<td>0.31</td>
<td>0.31</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.01)</td>
<td></td>
<td>(0.17)</td>
<td>(0.14)</td>
<td>(0.13)</td>
<td>(0.13)</td>
<td>(0.13)</td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td>CL2</td>
<td></td>
<td></td>
<td></td>
<td>0.24</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.14)</td>
<td>(0.11)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL3</td>
<td></td>
<td></td>
<td></td>
<td>0.49</td>
<td>0.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.16)</td>
<td>(0.13)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL4</td>
<td></td>
<td></td>
<td></td>
<td>0.25</td>
<td>0.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.13)</td>
<td>(0.11)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL5</td>
<td></td>
<td></td>
<td></td>
<td>0.32</td>
<td>0.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.10)</td>
<td>(0.08)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PNBA2</td>
<td></td>
<td></td>
<td></td>
<td>0.89</td>
<td>0.50</td>
<td>0.44</td>
<td>0.40</td>
<td>0.33</td>
<td>0.28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.02)</td>
<td>(0.16)</td>
<td>(0.14)</td>
<td>(0.12)</td>
<td>(0.11)</td>
<td>(0.09)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PNBALF</td>
<td></td>
<td></td>
<td></td>
<td>0.39</td>
<td>0.41</td>
<td>0.39</td>
<td>0.35</td>
<td>0.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.12)</td>
<td>(0.11)</td>
<td>(0.10)</td>
<td>(0.09)</td>
<td>(0.08)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNBA2</td>
<td></td>
<td></td>
<td></td>
<td>0.94</td>
<td>0.94</td>
<td>0.79</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNBA3</td>
<td></td>
<td></td>
<td></td>
<td>0.99</td>
<td>0.92</td>
<td>0.79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.00)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNBA4</td>
<td></td>
<td></td>
<td></td>
<td>0.99</td>
<td>0.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.00)</td>
<td>(0.01)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNBA5</td>
<td></td>
<td></td>
<td></td>
<td>0.99</td>
<td>0.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.00)</td>
</tr>
</tbody>
</table>

\textsuperscript{1}Purebred sows were from a Landrace pure line, and the crossbred sows were F1 offspring from a cross between a Landrace pure line and a Large White pure line.

\textsuperscript{2}PL2, CL2, CL3, CL4, and CL5 are Bernoulli traits. For PL2, a success is considered to be a sow that remains in the nucleus herd long enough to farrow 2 litters. For CL2, CL3, CL4, and CL5, a success is considered to be a sow that remains in the nucleus herd long enough to farrow 2, 3, 4, and 5 litters, respectively. PNBA2 and PNBALF represent the cumulative number born alive for purebred sows at parity 2 and for their lifetime. CNBA2, CNBA3, CNBA4, CNBA5, and CNBALF represent the cumulative number born alive for purebred sows at parity 2, parity 3, parity 4, parity 5, and for their lifetime. The genetic correlations between two Bernoulli traits were estimated using a Bayesian approach, and all other genetic correlations were estimated using ASREML.
Table 4.5. Relative emphasis of economically important swine traits based on genetic standard deviation and economic value per trait unit

<table>
<thead>
<tr>
<th>Trait</th>
<th>Economic Value per Trait Unit ($v_i$)</th>
<th>Genetic Standard Deviation ($\sigma_g$)</th>
<th>Standardized Economic Weight ($E_i$)</th>
<th>Relative Emphasis (E)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NBA</td>
<td>$27.93</td>
<td>0.70</td>
<td>19.55</td>
<td>0.066</td>
</tr>
<tr>
<td>NW</td>
<td>$38.57</td>
<td>0.28</td>
<td>10.91</td>
<td>0.037</td>
</tr>
<tr>
<td>LW</td>
<td>$0.62</td>
<td>8.07</td>
<td>5.00</td>
<td>0.017</td>
</tr>
<tr>
<td>WTE</td>
<td>$1.90</td>
<td>1.05</td>
<td>2.00</td>
<td>0.007</td>
</tr>
<tr>
<td>LBW</td>
<td>$0.62</td>
<td>2.92</td>
<td>1.81</td>
<td>0.006</td>
</tr>
<tr>
<td>NBALF</td>
<td>$27.93</td>
<td>9.19</td>
<td>256.58</td>
<td>0.867</td>
</tr>
<tr>
<td>D250</td>
<td>$0.22</td>
<td>5.67</td>
<td>1.24</td>
<td>0.004</td>
</tr>
<tr>
<td>BF</td>
<td>-$18.10</td>
<td>0.09</td>
<td>1.62</td>
<td>0.005</td>
</tr>
<tr>
<td>LEA</td>
<td>$2.48</td>
<td>0.54</td>
<td>1.33</td>
<td>0.004</td>
</tr>
</tbody>
</table>

1NBA – Number born alive, NW – number weaned, LW – litter weaning weight (lbs), WTE – wean to estrus interval, LBW – litter birth weight (lbs), NBALF – lifetime born alive, D250 – days to 250 lbs, BF – Backfat (in), LEA – loin eye area (in$^2$)
2Economic values per unit were based on values reported by NSR (2010).
3Genetic standard deviations were obtained from a commercial seedstock supplier with the exception of NBALF which was based on results from this study.
4$E_i = v_i * \sigma_g$
5$E = E_i / \Sigma (E_i)$
CHAPTER 5: TOTAL COST ESTIMATION FOR IMPLEMENTING GENOME-ENABLED SELECTION IN A MULTI-LEVEL SWINE PRODUCTION SYSTEM

Modified from a paper submitted to *Genetics Selection Evolution*

C. E. Abell¹, J. C. M. Dekkers¹, M. F. Rothschild¹, J. W. Mabry¹, and K. J. Stalder¹

¹Department of Animal Science, Iowa State University, Ames, IA 50011

Abstract

**Background:** Determining an animal’s genetic merit using genome-enabled selection can improve estimated breeding value (EBV) accuracy. Increasing EBV accuracy will increase the rate of genetic gain within the breeding program; however, the magnitude of the accuracy improvement must be large enough to recover the costs associated with implementing genome-enabled selection. One way to reduce the genome-enabled selection costs is to genotype selection candidates from the nucleus herd using a low density chip and use high density chip genotyping for animals that are used as parents in the nucleus breeding herd.

The objective of this study was to develop a tool to estimate the cost structure associated with incorporating genome-enabled selection into multi-level commercial breeding programs.

**Results:** For the purpose of this deterministic study, it was assumed that the commercial pig is created from a terminal line sire and a dam that is a cross between two maternal lines. It was assumed that a 1000 sow nucleus population formed the basis for both maternal line genetic programs and all male and female selection candidates were genotyped at low density and all animals used for breeding at high density. With the assumptions used in this analysis, it was estimated that genome-enabled selection costs for a maternal line would be approximately US$0.082 per weaned pig in the commercial production system, assuming that the boars produced in the nucleus are used at capacity. Since there are two maternal
lines, a total of US$0.164 per weaned pig would be needed to incorporate genome-enabled selection into the two maternal lines. Similarly, for a 600 sow terminal line nucleus herd and genotyping only male selection candidates with the low density panel, the cost per weaned pig in the commercial herd was estimated to be US$0.044. This means that US$0.21 per weaned pig produced at the commercial level that are sired by boars obtained from the nucleus herd breeding program needs to be added to the genetic merit value in order to break even on the additional cost required when genome-enabled selection is utilized in both maternal lines and the terminal line.

**Conclusions:** By modifying the input values, such as herd size and genotyping strategy, to represent the specific design of any swine breeding and production system, a flexible spreadsheet tool developed from this work can be utilized to estimate the additional costs associated with genome-enabled selection. This tool will aid producers in estimating the economic viability of incorporating genome-enabled selection into their specific breeding program.

**Background**

The common method to estimate breeding values and rank animals based on genetic merit is known as traditional BLUP (best linear unbiased prediction) selection. Traditional BLUP selection relies on phenotypic information measured directly on selection candidates and their relatives to predict the genetic merit for all animals. This method uses recorded performance values for traits measured. If molecular information is included in the selection program, it would be included on a marker by marker basis and treated as a fixed effect.

With advances in molecular technology and new biological tools, researchers have established ways to incorporate variation at the DNA level into breeding programs. When
breeders incorporate the variation from genomic information, into a selection strategy, it is known as genome-enabled selection. This information is used to enhance traditional breeding value estimation. Estimating an animal’s genetic merit at the molecular level may improve estimated breeding value (EBV) accuracy\[^1\]. This improved accuracy could further increase the rate of genetic gain for the population. Traits that are lowly heritable, hard to measure, sex-limited, measured later in life, or measured after slaughter have the greatest potential for accuracy improvement when using genome-enabled selection compared to traits that can be directly measured on all the candidates before selection.

Utilizing a low density marker panel can be an effective method to reduce genotyping costs once the initial training data set has been established. A training population is needed to determine population haplotypes so that imputation can be used to infer genotypes from a low density marker panel. Inferring high density genotypes from a low density panel is known as imputation\[^2\]. When imputation is used, selection candidates are genotyped using low density panels and the actual selected animals are re-genotyped using a high density marker panel. To further reduce costs, companies may choose to genotype only males. While this saves costs, genomic EBV accuracy for females may be reduced especially for sex-limited traits or novel traits where phenotypic data are not routinely collected.

The objective of this study was to develop a tool to estimate the cost structure associated with incorporating genome-enabled selection into commercial breeding programs. Estimating an animal’s genetic merit using genome-enabled selection can improve the accuracy of EBVs; however, this improved accuracy should be large enough to recover the costs associated with implementing genome-enabled selection.
Methods

Most commercial animals are the offspring from a mating between a female that is a cross between two maternal lines and a male that is from a terminal line. The three lines can be purebred, synthetic, or some composite of the same type. Therefore, each maternal line contributes 25% of the genetic material to commercial animals, and the terminal line makes up the other 50%. One maternal line is typically derived from a Landrace population, while the other maternal is usually derived from a Large White population. The terminal line is often derived from a Duroc population. Each maternal line nucleus was assumed to have 1,000 sows while there were assumed to be 600 sows in the terminal nucleus. All three lines must be selected for improved performance at the commercial level, and thus, EBVs must be estimated for each line. The terminal line is typically selected based on a terminal sire index that can consist of growth, meat quality, and carcass traits. The maternal lines are often selected based on a maternal line index that is comprised of reproductive traits that are more heavily emphasized in the index and some terminal market traits.

When considering the costs associated with genotyping, one has to consider not only the genotyping costs, but also other ancillary expenses. In the present analysis, genotyping costs were assumed to be US$115.00 and US$55.00 for the high and low density panels, respectively. These costs include the cost of the genotyping and all other costs associated with sample collection, DNA isolation and storage, shipping, etc. Genotyping costs will vary depending on the number of animals genotyped and the company used. Additionally, genotyping costs will change (increase or decrease) as new tests become available. For each of the 3 lines, it was assumed that the initial training data consisted of 2,000 ancestor animals that were genotyped using the high density marker panel. The EBV accuracy from genome-
enabled selection can be predicted using the genomic relationship between the training animals and selection candidates\textsuperscript{[3,4]}. Increasing the average genomic relationship between the training data and selection candidates may improve EBV accuracy\textsuperscript{[3,4]}.

Developing genome-enabled EBVs requires more time and computing power compared to traditional BLUP EBVs. For this analysis, it was assumed that 8 months was required to analyze the training dataset and develop the program that will be implemented for genetic evaluations. Based on personal communications with a large swine genetics company, an additional 8 hours was assumed to be needed with each weekly evaluation to prepare data and to ensure the program runs without errors. It was assumed that the EBV development and additional weekly work would cost US$60.00/hour in employee wages and benefits.

A US$50,000 investment was assumed for computing infrastructure. This cost includes equipment and labor associated with set-up. For this study, the infrastructure cost was distributed across the three lines included in the breeding program based on the relative proportion of nucleus sows in each line. For this analysis, it was assumed that the costs of the investment are expected to be recovered within 3 years. In other words, a 3-year planning horizon was assumed. The infrastructure costs were assumed to have no salvage value at the end of the planning horizon. A 5% discounting rate\textsuperscript{[5]} was used to calculate the present value of the nominal annual costs associated with genome-enabled selection.

When assessing cost structure within a genome-enabled selection program, the infrastructure costs are nominal relative to the annual genotyping costs. Therefore, a longer planning horizon would not greatly impact the total annual costs associated with incorporating genome-enabled selection into the breeding program. A longer planning
horizon would allow accumulated genetic improvement in multiple generations to help offset the additional costs. Typically, genetic companies must recover their investments within a 2-3 year time frame in order for the investment to be sustainable.

Once a training data set has been collected and analyzed, routine genotype collection can be scheduled for selection candidates within each line in the breeding program. The number of selection candidates in a given year that must be genotyped will be determined by the genetic sampling strategy and by the number of offspring produced by the nucleus breeding herd. For this analysis, all male and female selection candidates were expected to be genotyped at low density and then all animals used for actual breeding were re-genotyped at high density in both maternal lines. While genotyping both males and females is optimal for increased annual genetic gain and reduced inbreeding, it is not required to make genetic improvement using genome-enabled selection\[^6\]. Only male selection candidates were genotyped for the terminal line and selected boars were re-genotyped at high density. Only males are genotyped in the terminal line since most terminal line selection indices do not include sex-limited traits, lessening the value of genotyping females. This genotyping scheme is representative of the strategy currently being implemented in the swine industry.

For the maternal lines, it was assumed that, on average, each litter in the nucleus herd consisted of 11 pigs born alive\[^7\] with a 1:1 sex ratio\[^8\]. Each sow was assumed to have 2.3 litters a year, on average\[^7\]. The combined pre-weaning and nursery mortality was assumed to be 12%\[^7\]. These assumptions result in a total of 22,264 animals (1000 sows × 2.3 litters × 11 pigs × 0.88 survival; 11,132 males and 11,132 females) to be low density genotyped annually in each maternal line. With 5% and 20% selection for males and females\[^7\], respectively, 557 males (11,132 × 0.05) and 2,226 females (11,132 × 0.20) will be re-genotyped with the high
density panel annually in each maternal line. These represent the number of animals selected, but not all animals will be successful, reproductive members of the breeding herd.

The average sow production in the terminal line nucleus herd was assumed to be 10 pigs born alive per litter with a 1:1 sex ratio\(^8\) and 2.1 litters per sow per year\(^7\). The total mortality through the nursery was assumed to be 12%, as for the maternal lines. This means that 5,544 male pigs are produced and low density genotyped in the terminal line nucleus herd, annually. Assuming 5% male selection, 277 males would be re-genotyped with a high density panel.

The top 5% of the boars produced in the nucleus herd based on genome-enabled EBVs was assumed to be used as nucleus replacements in the maternal and terminal lines. The top 60% of male candidates in the sire line were used in the commercial production system for the terminal line. Boars were assumed to produce 25 semen doses weekly with 15% of doses being thrown away due to semen quality issues\(^9\). It was assumed that 25% of selected boars would not reach production due to infertility, disease, etc. All boars utilized for reproduction within the nucleus or at the commercial level were assumed to be used to maximum capacity, meaning that all doses that are not discarded are used for insemination. This means that 461,282 and 2,756,754 doses were produced annually from the maternal line boars and terminal line boars, respectively. For example, the annual doses produced in a terminal line would be calculated as 5,544 selection candidates × 75% productive × 60% used × 25 doses per week × 85% viable semen × 52 weeks.

The total number of weaned pigs produced each year from the semen doses collected was calculated as 1,940,845 and 10,544,584 for the maternal and terminal lines, respectively. The number of total weaned pigs produced each year was calculated by multiplying the
expected number of litters by the number of pigs weaned in a litter. This was based on 10% pre-weaning mortality in the commercial herd, 2 doses of semen per sow serviced, and an 85% farrowing rate for commercial sows with 10 and 11 pigs born alive per litter in the terminal and maternal lines, respectively\[^7\]. For example, the number of weaned pigs produced from the terminal boars would calculated as $2,756,754 \text{ doses} / 2 \text{ doses per service} \times 85\% \text{ farrowing rate} \times 10 \text{ pigs/litter} \times 90\% \text{ survival to weaning}$. For the maternal lines, 40% of the gilts developed were assumed to never produce a litter\[^10\], and for the sows that farrowed at least one litter, it was assumed that each sow produced 35 weaned pigs per lifetime, on average\[^7\].

A spreadsheet was developed to calculate the total estimated costs associated with incorporating genome-enabled selection into a swine breeding program. The spreadsheet can calculate the accumulated costs for up to 6 different genetic lines that could be used to create a commercial animal. The nucleus herd size and nucleus production levels (i.e. number born alive, litters per sow per year, mortality, and boar production) can be changed to reflect the production system of the user. Which selection candidates are sampled on an annual basis can be altered to depict the strategy the company anticipates implementing. Similarly, the spreadsheet can account for additional phenotyping costs associated with a novel trait being added to the selection criteria. Additionally, the spreadsheet has the capability to account for a multiplication factor if the company has a multiplication level within their production system.

**Results and Discussion**

The total start-up cost was US$326,031.00 for both maternal lines and US$318,338.00 for the terminal line. This includes the costs associated with developing the
genome-enabled EBVs and the genotyping of the initial training data set. The annual costs of routine genotyping were US$1,545,045.00 and US$337,278.00 for the maternal and terminal lines, respectively. Assuming a 3-year planning horizon and a 5% discounting rate\textsuperscript{[5]}, the increased revenue for the maternal and terminal lines must be US$1,674,334.68 and US$463,516.97, respectively, in order to break even on the increased costs associated with incorporating genomic information into the selection program. Infrastructure costs are relatively small compared to genotyping costs; therefore, the annual total cost and cost per weaned pig are relatively insensitive to changes in the infrastructure costs.

Investments are routinely evaluated for their return on investment, which in turn is often annualized in order to evaluate yearly expenses relative to income. In the swine industry, genetic companies will incur the costs associated with genome-enabled selection. How genetic suppliers market their animals will govern how the costs are recovered. Some companies may sell boars or replacements while other companies may receive royalties based on the number of pigs weaned by their customers. For this study, it was assumed that the genetic company making the investment in genome-enabled selection receives royalties from weaned pigs at the commercial level.

Dividing the annualized costs by the number of semen doses per year, results in costs of US$3.63 and US$0.17 per dose in the maternal and terminal lines, respectively. To calculate the cost per nucleus boar, the annualized cost was divided by the number of boars used for commercial production (male selection candidates × percent used). The cost per nucleus boar was US$3,008.15 and US$139.34 for the maternal and terminal lines, respectively. Using the total number of female pigs produced annually from the maternal nucleus herds (1,940,845 / 2), it was determined that genome-enabled selection costs would
be approximately US$2.88 per nucleus daughter (F1 cross between two maternal lines). This number was calculated from dividing the annualized costs by the number of productive sows produced from the nucleus boars, accounting for the gilt drop out percent. Assuming that each sow produced 35 weaned pigs in her lifetime, there is a US$0.082 cost per weaned pig in the commercial production system for each maternal line (US$2.88 / 35). The cost per weaned pig in the commercial herd was determined to be US$0.044 for the terminal line, assuming that 10,544,584 weaned pigs are produced annually. This means that US$0.21 per weaned pig from boars produced in the nucleus would need to be added to the genetic merit for each market pig in order to break even on the additional cost associated with genome-enabled selection for all 3 lines. The marketing structure for the genetic supplier will determine how the additional costs associated with genome-enabled selection can or will be recovered from commercial sales.

The current rate of genetic gain in the nucleus will determine the proportional increase in rate of genetic gain needed to recover genome-enabled selection costs. According to the National Swine Registry, the current rate of annual genetic gain for their terminal line index is US$0.30, US$0.40, and US$0.30 per weaned pig for the Duroc, Landrace, and Large White populations, respectively[12]. The rate of genetic gain for number born alive is 0.08 and 0.07 pigs per year for Landrace and Large White, respectively. Thus, the total annual genetic improvement for both maternal lines would be US$0.432 and US$0.321 for the Landrace and Large White lines, respectively. For each maternal line, the total genetic improvement value is calculated by multiplying the terminal improvement by 1 plus the increase in number born alive ($0.40 \times 1.08$, for Landrace and $0.30 \times 1.07$ for Large White). This rate of genetic gain may be lower than gains expected from most swine genetic companies.
The current expected improvement in genetic merit at the commercial level and the relative additional improvement needed to pay for incorporating genome-enabled selection into the breeding program are illustrated in Figure 5.1. Considering that 50% of the terminal line index improvement (US$0.15/weaned pig) in the Duroc line would be passed on to the commercial herd and that the cost estimate for genome-enabled selection in the terminal line was US$0.044 per weaned pig, this would mean that a 29% improvement (US$0.044 / US$0.15) in the genetic gain rate would be needed to recover the genome-enabled selection costs. Considering this and the fact that 25% of the improvement occurring in each of the 2 maternal lines will be passed on to the commercial animals, 76% (US$0.082 / US$0.108) and 103% (US$0.082 / US$0.080) improvement in genetic gain rate is needed for the Landrace and Large White populations, respectively. Because of the planning horizon used in this study and by most swine genetic companies, these returns are based on one round of genetic improvement; however, differential genetic improvements made with genome-enabled selection will be accumulated over time.

These results reflect the necessary increase in genetic improvement at the nucleus level to break even on the estimated genome-enabled selection costs provided that the realized genetic improvement in commercial animals is equivalent to the expected genetic improvement. Environmental factors may reduce the realized genetic improvement at the commercial levels compared to the expected improvement. Greater nucleus level genetic improvement would be required if all genetic improvement occurring at the nucleus population is not realized or does not actually occur at the commercial production level.

Another way to recover the costs associated with genome-enabled selection is through maintaining or capturing increased market share. If a company must invest in
genome-enabled selection to maintain their current market share, there must be some perceived marketing value associated with incorporating genome-enabled selection into the breeding program. If a marketing value exists, the difference between the total costs associated with genome-enabled selection per weaned pig and the marketing value, is the increased value in genetic improvement that must be achieved to break even on the investment in genome-enabled selection.

Without increasing the nucleus herd size, the only way to increase the market share is to increase the proportion of male offspring produced in the nucleus herd that are used to produce commercial piglets. The relationship between the proportion of nucleus boars used and proportion of improvement in genetic gain rate for the terminal and maternal lines are shown in Figures 5.2 and 5.3, respectively. For example, if 80% of the male candidates in the terminal line were utilized compared to the 60% used in this study, the proportional improvement in rate of genetic gain needed to break even on the investment in genome-enabled selection would be 22% compared to 29%. The feasible region, or the region where the costs of genome-enabled selection are recovered, is shaded in gray. The gray area not on the break even line indicates a profit is made. For the results in Figure 5.2, it was assumed that the current rate of genetic gain in the commercial animals was US$0.15 per pig from the terminal line. The current rate of genetic gain in the commercial animals was assumed to be US$0.10 per pig from the maternal line for the results in Figure 5.3.

Annual genetic gain is calculated as:
\[ \Delta G = \frac{r \times i \times \sigma_g}{L} \]

where \( r \) is the accuracy, \( i \) is the selection intensity, \( \sigma_g \) is the genetic standard deviation, and \( L \) is the generation interval\(^8\). Using genomic information to make selection decisions will only impact accuracy. Selection intensity is based on production flow and scheduling, and therefore, would not change based on changing the selection method. For many terminal traits, animals will have their own phenotypic value at the time of selection; therefore, litter mates will not have the same breeding value if a selection index that incorporates terminal traits is being utilized in the herd. However, if a parental average is used at selection, all animals from the same litter would have the same breeding value. Thus, selection intensity could be impacted through controlling inbreeding accumulation by ensuring that not all animals from a family are selected. Depending on the mating program used, inbreeding could be higher or lower when using genome-enabled selection compared to traditional BLUP selection. Having the exact genomic-based relationship might allow breeders to avoid matings that are closely related. Conversely, these genomic-based relationship values could be used such that animals more closely related than the population average are mated allowing inbreeding rates to rise.

Additionally, the underlying genetic variance would not be changed due to a change in selection methods. Since pigs are selected and mated relatively soon after sexual maturity, the generation interval would not be impacted by incorporating genomic information into the selection program. This is different compared to other livestock industries, such as dairy cattle, where animals are not selected and mated at sexual maturity. Other industries may
have the opportunity to improve another component of the genetic gain equation, such as generation interval, rather than accuracy alone as in the swine industry.

The primary expected benefit from genome-enabled selection is improved EBV accuracy\textsuperscript{[1]}. Due to the direct relationship between accuracy and rate of genetic gain, increasing EBV accuracy will proportionally increase the rate of genetic gain expected given that selection intensity and generation interval remain constant. This increase in accuracy will have to be sufficient enough to recover the added costs associated with using genomic information in the selection program.

Muir\textsuperscript{[13]} showed that there is potential for increased genetic improvement when using genomic selection in addition to traditional BLUP selection\textsuperscript{[13]}. The added data from genomic information increases breeding value accuracies and therefore, increases response to selection. The study also showed that genomic selection could have an advantage over traditional methods when applied to traits that have low heritability and are difficult to measure due to cost of measurement or only being able to measure the trait on certain animals, such as number born alive.

The improved EBV accuracy in genomic selection compared to traditional BLUP selection results from being able to estimate the genetic merit of an animal at an earlier age, potentially before the animal has performance information from the individual itself or its progeny when genomic selection is practiced. When single-step selection is practiced, the improved accuracy comes from better genetic relationship estimates among animals rather than using expected genetic relationships based on pedigree information\textsuperscript{[14]}. Genomic information can also be utilized to correct pedigree errors. Using genomic information in this way, not only leads to increased accuracy through genotype data, but traditional BLUP EBVs
are improved due to a more correct or more accurate pedigree. Correcting pedigree errors will improve the connectedness for individuals in the population and properly associate relatives’ records.

If genomic information is used alone without phenotypic information, the genetic improvement resulting from selection may not exceed the genetic improvement based on BLUP breeding values\[13,15\]. All available phenotypic and genomic information should be incorporated into the EBVs to ensure the most accurate EBV for selection. Genome-enabled selection will not eliminate the need for phenotype collection. Earlier theories associated with genome-enabled selection suggested that one benefit would be the cost savings associated with reducing or eliminating the collection of phenotypic data\[16\]; however, due to the decay in accuracy associated with genomic breeding values over generations, genomic breeding values must be re-estimated periodically and phenotypic records will be needed\[15\].

The application of genomic breeding values has been investigated for pig populations. Nielsen and coworkers showed the correlation between genomic breeding values and traditional BLUP breeding values to be 0.62 for the 170 boars used in their data set\[17\]. Cleveland and collaborators reported the accuracy of the genomic breeding value for total pigs born per litter to be between 0.64 and 0.82, depending on the type of training data set used\[18\]. The authors reported the accuracy of stillborns per litter to range from 0.33 to 0.68.

Dekkers\[15\] developed a method using selection index theory to calculate the genetic response expected from incorporating genomic information into a selection index. The method deterministically calculated the genetic response anticipated from using genomic selection with defined genetic parameters. The study showed that, for a trait recorded on both sexes prior to selection, selection based on markers alone can improve response by 8.5%
compared to selection based on solely on phenotypic information. Based on stochastic simulations, annual genetic gain could be increased by 23 to 91% for a maternal line\textsuperscript{[16]} and 27 to 33% for a terminal line\textsuperscript{[19]}. Based on these increases in annual genetic gain, there is potential for genome-enabled selection to be profitable for both maternal and terminal line selection programs.

Marker assisted selection has increased genetic improvement from selection for meat quality, net or residual feed intake, and pigs born alive compared to the response from traditional BLUP with the largest gap between marker assisted selection and traditional BLUP being for meat quality; however, no genetic improvement in rate of genetic gain was observed for growth when comparing marker assisted and traditional BLUP selection methods\textsuperscript{[20]}. This suggests that the greatest potential for improvement in the rate of genetic gain from genome-enabled selection will be for lowly heritable traits that are hard to measure, sex limited, measured later in life or measured after slaughter, such as meat quality, disease resistance, feed efficiency, and longevity.

Disease resistance is not easily defined and not systematically measured. Feed efficiency is expensive to measure directly, especially on an individual animal basis. Sow longevity is not recorded until the sow is culled from the herd and is a trait that is only measured on females. If traits are not currently measured and recorded, additional costs associated with measuring the novel traits will be connected with genome-enabled selection if these traits are targeted in a selection program. For a novel trait to be incorporated into a selection program, a measurable phenotype associated with the trait must be clearly defined. Depending on the phenotype, there may be a significant cost associated with the infrastructure needed to collect the data.
Conclusions

Using genomic information to estimate an animal’s genetic merit at the molecular level can improve estimated breeding value (EBV) accuracy when compared to an EBV based solely on phenotypic records. However, genome-enabled selection is expensive and the increase in rate of genetic gain must be large enough to offset the costs associated with incorporating genome-enabled selection into a breeding program. A flexible spreadsheet\(^a\) tool developed from this work can be utilized to estimate the returns needed to recover additional costs associated with genome-enabled selection by modifying the input values such as herd size and genotyping strategy to represent the specific design of any production system.

Endnotes

\(^a\)The spreadsheet can be found at the Iowa Pork Industry Center website (http://www.ipic.iastate.edu/).

References


Authors’ Contributions

JD, MR, and JM provided technical information and support for the project. CA and KS conceived the project, participated in spreadsheet development, and drafted the manuscript. All authors read and approved the final manuscript.
**Figure 5.1. Value of expected improvement in genetic merit at commercial level\(^a\)**

<table>
<thead>
<tr>
<th>Maternal Line – Landrace</th>
<th>Maternal Line – Large White</th>
<th>Terminal Line – Duroc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Terminal Line Index – $0.40/\text{year/pig}$</td>
<td>Terminal Line Index – $0.30/\text{year/pig}$</td>
<td>Terminal Line Index – $0.30/\text{year/pig}$</td>
</tr>
<tr>
<td>Number Born Alive – 0.08 pigs/\text{year}</td>
<td>Number Born Alive – 0.07 pigs/\text{year}</td>
<td>$0.15 \text{ transferred to commercial pigs}$</td>
</tr>
<tr>
<td>$0.432 \text{ overall improvement}$</td>
<td>$0.321 \text{ overall improvement}$</td>
<td>$0.044/0.15 = 29%$</td>
</tr>
<tr>
<td>$0.108 \text{ transferred to commercial pigs}$</td>
<td>$0.080 \text{ transferred to commercial pigs}$</td>
<td></td>
</tr>
<tr>
<td>$0.082/0.108 = 76%$</td>
<td>$0.082/0.080 = 103%$</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Current rates of genetic improvement were based on genetic improvement published by the National Swine Registry for each of the three lines. The cost of genome-enabled selection was estimated using the values reported in this paper.
Figure 5.2. Feasible region for profitability when incorporating genome-enabled selection in a terminal line selection program.

The current annual rate of genetic improvement was assumed to be US$0.15 per commercial pig. The number of commercial pigs produced from the nucleus boars was calculated using the values reported in the paper. The estimated cost of genome-enabled selection in the terminal line was US$463,517. This number was estimated using the strategy in this paper.
**Figure 5.3:** Feasible region for profitability when incorporating genome-enabled selection in a maternal line selection program. 

---

The current annual rate of genetic improvement was assumed to be US$0.10 per commercial pig. The number of commercial pigs produced from the nucleus boars was calculated using the values reported in the paper. The estimated cost of genome-enabled selection in the terminal line was US$1,674,335. This number was estimated using the strategy in this paper.
CHAPTER 6: GENERAL DISCUSSION AND CONCLUSIONS

Sow longevity is an economically relevant trait for commercial swine operations. Sow longevity at the nucleus level is not fully expressed since rapid herd turnover is needed to maximize genetic improvement. Therefore, traditional direct selection methods can prove inadequate or insufficient for incorporating longevity into a swine selection program where the breeding objective is to improve the profitability of crossbred pork production at the commercial level. Selection for sow longevity is often based on indicator traits, such as litter size, feet and leg soundness, and backfat, growth rate, and loin muscle area measured at the end of the gilts’ grow-finish phase of production and before she enters the isolation/acclimation areas. Indicator traits must be utilized in the selection program since sows at the nucleus level are not allowed to fully express the trait before they are removed from the herd. The studies in this dissertation evaluated ways to improve sow longevity using multiple selection methods.

The first study analyzed objective measures to detect sow lameness. Developing an automatic lameness diagnosis algorithm will benefit animals, producers, and veterinarians in timely and effective lameness identification among individual sows before clinical signs are apparent. Additionally, this tool could aid producers in their efforts to decrease herd lameness by identifying animals that are less prone to become lame and selecting their offspring to become parents of future generations. Being able to predict sow lameness can aid in delivering maximum animal health and welfare benefits, improving sow lifetime productivity, and optimizing sow farm labor. The results of the first study in this dissertation indicated that the force plate measuring system was able to identify induced sow lameness by separately measuring the weight each sow is willing to bear on each leg on each lameness
trial day. Using this device in conjunction with the lameness detection tree created from this study can aid swine producers in detecting lame sows, and thus, improve management decisions in relation to lame animals. The force plate could be incorporated into an electronic feeding system where sow lameness could be monitored on a daily basis without human interaction. More timely lameness detection could result in more treatable lameness, less mortality and euthanasia due to lameness, improved productivity, less treatment expense, greater production efficiency and profitability for the sow operation.

The force plate system could be implemented in a commercial setting by adding the device to a stall in an electronic feeding system. Each force plate costs approximately $5,000. No extra labor would be needed to operate the force plate on a day-to-day basis, so materials and maintenance (recalibration, replacing load cells, etc.) costs would be the only expense associated with adding the force plate into a commercial swine facility. In order to justify the expense, producers must reduce the labor and treatment costs or retain the salvage value associated with lame sows. Assuming that labor costs are $12 or $20 per hour (depending if the laborer is a department manager or unit manager) and treatment costs are $10-$30/sow/day (depending on the treatment), a company would need to save 250 or 417 hours of labor or 167-500 treatments to break even on the expense of a single force plate. If the salvage value per cull sow is $150, then a company would need to retain the salvage value for 34 sows to justify implementing a force plate into a commercial barn. These savings would have to occur over the lifetime of the force plate system which would typically be depreciated within a 5 year planning horizon.

The force plate system would measure sows on a daily basis, provided they are eating. This would allow producers to monitor sows regularly without having to devote labor
to visually assessing sows for lameness. Without visually evaluating sows daily, early lameness signs can be missed and only sows with severe lameness problems could be detected. Furthermore, the force plate would remove any variation between or within workers performing visual lameness assessments on the sow herd, and time and money could be saved by not having to train new employees to visually assess sows for lameness signs.

Additionally, this work can be used to define lameness in an objective manner that is easily quantifiable with higher repeatability compared to subjective measures. Once a trait is clearly defined and easily measured, it can be incorporated into a genetic evaluation. Selecting for decreased lameness incidences can indirectly improve sow longevity by removing sows that are more likely to become lame and be removed from the herd due to non-reproductive reasons. Removing sows at early parities results in lost opportunity at the grow-finish production phase because of the increased performance of pigs with dams in later parities compared to pigs from gilt litters.

Improving crossbred performance is the goal of most swine breeders. The second part of this dissertation evaluated the relationship between sow longevity at the nucleus level and sow longevity at the commercial level. The results indicated that there is a genetic component associated with sow longevity at both production stages; however, there is little to no genetic correlation between purebred and crossbred longevity from the particular population evaluated. This could be a result of different culling criteria at the nucleus and commercial herds, resulting in different definitions for sow longevity. Because of this, genetic companies need to determine if similar relationships exist between their pure and commercial lines. If so, crossbred records are needed to select for sow longevity. However, many genetic companies do not maintain commercial herds with crossbred animals, and do
not receive records from their commercial customers regarding the performance of their replacement gilts. They will have to identify commercial customers that are willing to keep meticulous records in order to augment the current pure line performance information.

If genetic suppliers cannot rely on accurate record keeping and reporting from their commercial customers, they must decide if maintaining a commercial herd to monitor the performance of their crossbred females would be economical based on the added information it would provide to the selection program. A commercial test herd would allow genetic companies to increase the rate of genetic gain for traits, such as longevity, that are difficult to improve without crossbred records. Genetic progress would be improved because crossbred performance can be used to make selection decisions rather than purebred performance, which is not perfectly correlated with crossbred performance on a genetic level.

The expense associated with developing and maintaining a commercial test herd would vary depending on the relationship a swine genetic company has with its customers. If the company already owns a commercial facility implementing new recording and management practices may be a minimal expense. If a company can develop a relationship with a customer to report commercial performance, the company must still implement new recording and management practices and additionally, the company will have to reimburse the commercial producer. The most expensive way to implement a commercial test herd would be for a genetic company to purchase land, build a facility, and start a commercial herd. This would not be necessary for most genetic companies, but would allow the genetic company complete control over the management decisions made at the herd.

The last study in this dissertation determined the total costs associated with incorporating genome-enabled selection into a swine breeding program. A flexible
spreadsheet tool developed from this work can be utilized to estimate the returns needed to recover additional costs associated with genome-enabled selection by modifying the input values such as herd size and genotyping strategy to represent the specific design of any production system.

Since sow longevity is sex-limited and measured late in life, it is a trait with opportunity to be improved through adding genotypic information into a swine breeding program. However, before incorporating genome-enabled selection into a swine breeding program, the economic viability of the investment must be examined. The spreadsheet developed as part of this dissertation can be utilized by genetic suppliers to determine if incorporating genome-enabled selection for longevity into their breeding program would be viable. If genome-enabled selection is utilized to select for longevity, crossbred information would still be needed to establish phenotypic records for the training population. The crossbred information would be needed since there may be little to no genetic correlation between longevity at the nucleus level and longevity at the commercial level as shown in this dissertation.

Improving sow longevity using all three methods discussed in this dissertation may not be economically feasible for a swine genetic company. Of the methods evaluated, the most important and effective way to improve sow longevity would be to implement information from a commercial test herd into a selection program. This would allow companies to directly measure and select for improve crossbred performance, which is the breeding objective. This would aid in selecting to improve sow productive lifetime as well as other traits with low genetic correlation between purebred and crossbred performance, such as health and disease resistance. Additionally, a commercial test herd would be one of the
first steps to implementing genome-enabled selection into a breeding program, and the force plate system could be incorporated into the facility once the herd has been established.

Extending a selection program to include a commercial test herd would allow genetic companies the opportunity to select for multiple traits using crossbred information and to evaluate other novel approaches to crossbred genetic improvement.

**Implications**

These projects together have assessed multiple approaches to improving sow longevity. Defining lameness in an objective manner can lead to incorporating lameness into a selection program. By removing the indirect causes of decreased longevity, such as lameness, producers can focus on sow performance to make removal decisions and to select for longevity. With sows being removed for performance reasons, it will become easier to select for longevity. Implementing a commercial test herd can allow genetic suppliers to select for crossbred performance based on records from crossbred and purebred animals rather than relying solely on purebred information which may not have an adequate genetic correlation with crossbred performance in order to maintain the desired genetic progress. Additionally, a commercial test herd would benefit genetic companies in selecting for other crossbred traits that are difficult to measure on purebred animals, such as health and disease resistance. With a commercial test herd, genetic companies could collect the necessary performance records needed to incorporate genome-enabled selection into a selection program. Relative genetic gain will increase the most for traits that are hard to measure, measured late in life, and/or sex-limited when using genome-enabled selection compared to traditional selection methods. It is important for producers and genetic companies to assess the economic viability of incorporating genome-enabled selection into a breeding program.
Genetic companies cannot expect customers to pay a premium for animals selected using genomic information without actually improving the animals’ genetic potential relative to animals selected using traditional selection methods. Longevity is one trait where genome-enabled selection may be economically viable. If a genetic company can succeed in improving sow longevity through an effective breeding program, production efficiency and profitability can be improved for commercial swine operations.
REFERENCES

Abell CE, Jones GF, Stalder KJ and Johnson AK 2010. Using the genetic lag value to determine the optimal maximum parity for culling in commercial swine breeding herds. Professional Animal Scientist. 26(4), 404-411.


Huang Y, Hickey JM, Cleveland MA and Maltecca C 2012. Assessment of alternative genotyping strategies to maximize imputation accuracy at minimal cost. Genetics Selection Evolution 44, 25.


Johnson RK 1980. Heterosis and breed effects in swine. Agricultural Experiment Station, Institute of Agriculture and Natural Resources, University of Nebraska, Lincoln, NE, North Central Regional Publication No. 262.


Knauer MT 2006. Factors influencing sow longevity. Thesis (M.S.), Iowa State University, Ames, Iowa, USA.


Koivula M, Strandén I, Su G and Mäntysaari EA 2012. Different methods to calculate genomic predictions – comparisons of BLUP at the single nucleotide polymorphism level (SNP_BLUP), BLUP at the individual level (G_BLUP), and the one-step approach (H-BLUP). Journal of Dairy Science 95(7), 4065-4073.


