Genetic Evaluation of Clinical Mastitis in Dairy Cattle

Emma Carlén
Faculty of Veterinary Medicine and Animal Science
Department of Animal Breeding and Genetics
Uppsala

Doctoral Thesis
Swedish University of Agricultural Sciences
Uppsala 2008

Acta Universitatis agriculturae Sueciae
2008:63

Cover illustration by Hanna Fahlkvist

ISSN 1652-6880
ISBN 978-91-85913-96-1

Tryck: SLU Service/Repro, Uppsala 2008

Genetic Evaluation of Clinical Mastitis in Dairy Cattle

Abstract

This thesis aims to advance our understanding of the genetic background of mastitis resistance in dairy cattle. More particularly, it seeks to improve the potential of genetic evaluation of clinical mastitis (CM) both by utilizing more of the available information and by applying appropriate methodology.

A linear cross-sectional model was used to estimate genetic parameters for binary CM, somatic cell count (SCC) and milk production in the first three lactations of Swedish Holstein cows. The unfavorable genetic correlations found between udder health and production traits emphasize the need to include mastitis resistance in the breeding goal. The higher heritability of SCC (0.10-0.14) than CM (0.01-0.03) and the high genetic correlation between these two traits (average 0.70) imply that SCC is a useful indicator trait in breeding for improved mastitis resistance.

The method of survival analysis was used to analyze time to first CM and was compared with a linear cross-sectional model on field data and with linear and threshold cross-sectional models in a simulation study. Despite the theoretical advantages of survival analysis, there was no difference in the accuracy of genetic evaluation when methods were compared within the same length of opportunity period. The correlation between true and predicted sire breeding values in the simulation study was, however, 8% greater when data from the full lactation rather than the first 150 days were used.

Longitudinal binary CM data in 12 intervals of first lactation were analyzed as repeated observations in a linear random regression model. This method has some appealing features so far as genetic evaluation of CM is concerned, but it was found rather sensitive for parameter estimation in the current setting. Genetic parameters from the chosen random regression model agreed rather well, however, with the corresponding estimates from a linear longitudinal multivariate model in which records of CM in the different intervals were considered as different traits. Both longitudinal models indicated clearly that CM is not the same trait genetically throughout lactation, something that is ignored in a cross-sectional model.

Keywords: dairy cattle, clinical mastitis, genetic evaluation, linear model, survival analysis, random regression model, longitudinal model, heritability, genetic correlation, breeding value prediction

Author's address: Emma Carlén, Department of Animal Breeding and Genetics, Box 7023, 750 07 Uppsala, Sweden
E-mail: Emma.Carlen@hgen.slu.se

The real voyage of discovery
consists not in seeking new landscapes
but in having new eyes

Marcel Proust

Contents

List of publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:

I Carlén, E., Strandberg, E. and Roth, A. 2004. Genetic parameters for clinical mastitis, somatic cell score, and production in the first three lactations of Swedish Holstein cows. *Journal of Dairy Science* 87, 3062-3070.

II Carlén, E., Schneider, M. del P. and Strandberg, E. 2005. Comparison between linear models and survival analysis for genetic evaluation of clinical mastitis in dairy cattle. *Journal of Dairy Science* 88, 797-803.

III Carlén, E., Emanuelson, U. and Strandberg, E. 2006. Genetic evaluation of mastitis in dairy cattle using linear models, threshold models, and survival analysis: A simulation study. *Journal of Dairy Science* 89, 4049-4057.

IV Carlén, E., Grandinson, K., Emanuelson, U. and Strandberg, E. 2008. Random regression models for genetic evaluation of clinical mastitis in dairy cattle. *Submitted*.

Abbreviations

CM	clinical mastitis
CSM	cross-sectional model
DIM	days in milk
GE	genetic evaluation
LM	linear model
LMVM	longitudinal multivariate model
MAST	binary clinical mastitis
NAV	the Nordic cattle genetic evaluation
NMAST	number of clinical mastitis cases
PBV	predicted breeding value
RRM	random regression model
SA	survival analysis
SCC	somatic cell count
TBV	true breeding value
TFM	time to first clinical mastitis
TM	threshold model
TMI	total merit index

Introduction

The main focus in dairy cattle breeding has traditionally been increased production. Intensive selection, together with improvements in environmental factors, such as housing, feeding, health measures, management and use of artificial insemination, has resulted in considerable increases in productivity per animal. As an example, the average milk yield of milk-recorded cows in Sweden increased from about 4500 to 9300 kg energy corrected milk between 1960 and 2006 (Swedish Dairy Association, 2007b). Similar trends can be seen in dairy cattle populations worldwide. It is today generally accepted that genetic selection for high milk yield results in undesirable side effects, with the animals being more prone to metabolic, reproductive and other health problems, including mastitis. One suggested biological explanation of the adverse effects is that a disproportionately large amount of resources available are allocated to the trait selected for, i.e. productivity, leaving the animal lacking in resources to adequately respond to other demands (Rauw *et al.*, 1998).

Sweden, together with the other Nordic countries, has a long tradition of combining productivity and functionality (e.g. health and reproduction) in a broad breeding goal for dairy cattle. Despite this, the number of veterinary treatments and involuntary culling related to health and fertility remains a problem (Swedish Dairy Association, 2007a). Mastitis is a major concern within the dairy cattle industry, as it is associated with animal suffering and substantial economic losses. This thesis aims to advance our understanding of the genetic background to mastitis resistance and, especially, to improve the genetic evaluation of resistance to clinical cases of mastitis. The assignment to animals of breeding values that facilitate accurate selection of superior individuals can contribute to genetic progress on mastitis resistance and should, in turn, improve both the health status of the cow and the economic situation of the farmer.

Background

Mastitis in dairy cattle

Mastitis is an inflammation of the mammary gland. It occurs as a result of the introduction and multiplication of pathogenic microorganisms in the udder (Harmon, 1994). Bacteria are the main cause of disease and the route of infection is usually through the teat canal. It is a highly complex disease, because it has numerous causative pathogens, a wide variety of physiological responses, and a multifactorial background in which several genes and many environmental factors are involved.

Causative pathogens and symptoms

The most common major mastitis pathogens, and those that make a serious impact on the cow, include *Staphylococcus aureus*, streptococci and coliforms. The pathogens can be either contagious (e.g. *S. aureus*) or of environmental origin (e.g. *Escherichia coli*). In the first case, the infected udder is the major reservoir and infections are often spread among cows during the milking process. Environmental pathogens, on the other hand, are commonly found in features of the surrounding environment of the cow, such as bedding, soil and manure (Harmon, 1994; Akers, 2002).

Mastitis can be divided into subclinical or clinical varieties, and into short-term or chronic manifestations, depending on the intensity and duration of infection. Different pathogens are also associated with different infection patterns. A clear distinction between *S. aureus* and *E. coli* has been demonstrated: the former mainly cause subclinical and chronic infections and the latter predominantly lead to isolated clinical cases (Schukken *et al.*, 1997; Vaarst and Enevoldsen, 1997). In subclinical mastitis there are no visible signs of infection, but there is reduced milk production and a

changed milk composition with an increased concentration of somatic cells (white blood cells and epithelial cells) and bacteria present in the milk. Clinical mastitis (CM) is characterized by visible signs, such as clots in, or discoloration of, the milk, and a tender and swollen udder. Fever and loss of appetite may occur. As for subclinical mastitis, the milk production is decreased and the milk composition is considerably altered (Harmon, 1994). Chronic mastitis is described as repeated cases of mastitis where the cow often fails to respond successfully to treatment, although clinical symptoms may disappear temporarily (Akers, 2002).

The relationship with somatic cell count

Milk somatic cells play a protective role against infectious disease in the mammary gland. The elevated somatic cell count (SCC), which is a measure of the number of cells per ml milk, and the change in relative proportions of the different cell types, are necessary and correlated defense mechanisms that are often very effective in eradicating the majority of infections. If defense is not sufficient, bacteria multiply and release toxins with negative effects on the mammary gland (Kehrli and Shuster, 1994). Infection status is the major factor influencing SCC (Harmon, 1994; Schepers *et al.*, 1997). Milk from a healthy udder usually contains less than 100 000 somatic cells per ml; and these cells are mainly macrophages and lymphocytes. In milk from an infected udder, by contrast, the number of somatic cells per ml might exceed 1 000 000 and be predominantly neutrophils (Kehrli and Shuster, 1994). SCC is generally used for identifying cows with subclinical mastitis; a change in SCC from under to over a threshold of 200 000 cells per ml has been reported to be a predictor of intramammary infection (Dohoo and Leslie, 1991; Schepers *et al.*, 1997).

Factors affecting mastitis incidence

Like many other economically important traits in dairy cattle, mastitis (and resistance to it) is multifactorial. The incidence of mastitis in a herd is associated with both the cows' exposure to causative pathogens in the surrounding environment and the cows' resistance, i.e. ability to combat infection. Risk factors associated with mastitis are often related to management practices, including milking technique and equipment, housing, cleanliness of the environment, hygienic quality of feed and water, preventive health measures and stress. Non-management factors such as season, parity, lactation stage, breed, udder conformation, milk production, milking speed and reproductive disorders are also known to be associated with mastitis (Schukken *et al.*, 1990; Barkema *et al.*, 1998; Hagnestam *et al.*,

2007; Nyman, 2007). The incidence of CM increases with increasing parity and is highest in early lactation, especially in first-parity cows (Barkema *et al.*, 1998; Zwald *et al.*, 2006; Hagnestam *et al.*, 2007; Swedish Dairy Association, 2007a). Figure 1 illustrates the relative incidence of CM within the first three lactations of Swedish Holstein cows. Breed differences have been reported in several studies, and of the two major dairy breeds in Sweden, Swedish Red cows have better udder health than Swedish Holstein cows (Emanuelson *et al.*, 1993; Nyman, 2007; Swedish Dairy Association, 2007a). In addition to the factors already mentioned, the genetic constitution and innate immune defense of a cow plays an important role in determining disease resistance in individual cows. There are several anatomical, physiological and immunological defense mechanisms in the cow against mastitis, and a large number of genes operate in these defenses (Shook, 1989). Most of these genes are unknown and are believed to have a relatively small effect. Quantitative trait loci and several candidate genes, mainly alleles at the bovine major histocompatibility complex (BoLA) locus, have, however, been found to be associated with CM and other udder health traits (e.g. review by Rupp and Boichard, 2003; Holmberg, 2007).

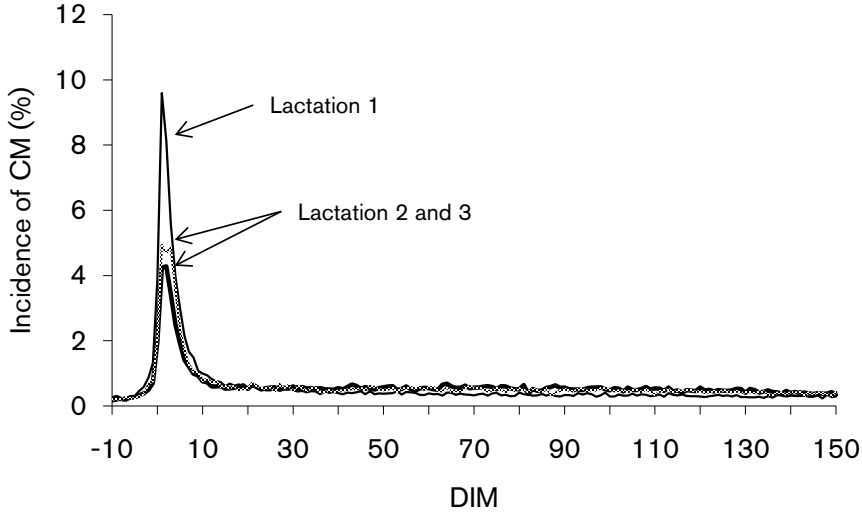

Figure 1. Relative incidence of the total number of CM cases within 150 days of the first three lactations in Swedish Holstein cows. This restricted period included 60-65% of the total number of CM cases within the lactation (results from Paper I).

Reasons for reducing the mastitis occurrence

There are many reasons why it is important to reduce the incidence of mastitis in the dairy cattle population. It is a very common disease associated with serious economic losses, impaired animal welfare and consumer and ethical concerns. In modern agriculture, the decreasing cost of production is often more valuable for the farmer than increasing income. Further, consumers expect their products to come from healthy animals and to be of high quality. Antibiotics are extensively used worldwide for treating CM, implying an increased risk of residues in milk and of the development of antibiotic resistance, which is considered to be a major public health threat. The maintenance of consumer confidence, by satisfactory levels of animal welfare and restricted use of antibiotics, continues to be important.

High incidence and a common reason for culling

Generally, the incidence of CM per cow-year varies between 20-40% (Heringstad *et al.*, 2000). In Sweden during the milk-recording year 2006/2007, CM incidence was 15.7% and treatments for CM represented 47% of all veterinary treated diseases. There was a 4% lower incidence in Swedish Red cows (13.8%) than there was in Swedish Holstein cows (17.8%) (Swedish Dairy Association, 2007a). This incidence only reflects reported veterinary treated cases, which are also affected by the farmer's ability to detect disease and willingness to call the veterinarian as well as the veterinarians' decision to report the disease. The reported incidence therefore underestimates actual incidence.

Furthermore, the risk for a cow of being culled following the occurrence of CM or elevated SCC has been reported to increase by a factor ranging from 1.5-5 (Seegers *et al.*, 2003). That CM affects culling decisions throughout lactation has been shown by Schneider (2006), but the effect here was dependent on pregnancy status and on lactation stage at the time of treatment. In Sweden during the milk-recording year 2006/2007, udder diseases were the second leading reason for culling after fertility problems and accounted for 15.6 and 16.8% of culled Swedish Red and Swedish Holstein cows, respectively. Adding the corresponding figures for high SCC (8.9 and 10.9%) makes udder health problems account for nearly as high a proportion of culled cows as fertility problems in the Swedish Red breed, and a higher proportion than fertility problems in the Swedish Holstein breed (Swedish Dairy Association, 2007a).

Economic losses

Costs related to mastitis are extensive. They are associated with reduced milk production, discarded milk because of antibiotic residues or low quality, veterinary and treatment, increased labor, increased risk of early involuntary culling and thus increased replacement costs, reduced milk price connected with increased SCC in the bulk milk tank as well as an increased disease risk in the future of affected and previously unaffected animals. Estimates of the cost per case of CM vary depending on sources of economic loss included, data and estimation method, and can therefore not be easily compared. In a recent simulation study by Hagnestam-Nielsen and Østergaard (accepted 2008) the economic consequences of CM under current Swedish farming conditions were examined and the cost of CM was estimated at €428 per case. Reduced milk production accounts for the largest part of the economic loss caused by CM, and in a review by Seegers *et al.* (2003) the total reduction in milk production resulting from CM was around 375 kg (5% at the lactation level). Production losses are, however, very variable and substantially influenced by, for instance, when in lactation the cow become diseased. Hagnestam *et al.* (2007) estimated a reduction in 305-day milk production between 0-902 kg (11%) depending on parity and the week of lactation at clinical onset. In the same paper it was also demonstrated that milk yield started to decline two to four weeks before diagnosis and that it was suppressed throughout the lactation. Apart from production losses, mastitis-related involuntary culling involves considerable costs.

Genetic selection for improved mastitis resistance

Generally, mastitis control programs have focused on environmental measures, i.e. improvements in management, as a means of reducing mastitis incidence. The heritability for CM is low, which has often been misinterpreted as meaning that genetic selection to improve the innate resistance has a limited role to play in mastitis control programs. However, the low heritability is mainly due to large environmental variation, which is difficult to control by any means, and considerable genetic differences between bulls exist (e.g. Philipsson *et al.*, 1995; Rupp and Boichard, 2003; Zwald *et al.*, 2004). In addition to CM, there are suitable indicator traits that can be used for genetic selection. Thus, there is a possibility for genetic improvement of mastitis resistance in the dairy cattle population.

Compared with management measures, genetic selection gives slower but accumulating effects involving lower costs and less effort. It is therefore

a permanent and cost-efficient method that should be used as an important complement to various management measures for controlling mastitis. Genetic selection for mastitis resistance is especially important because of the unfavorable genetic correlation with milk production (Shook, 1989; Emanuelson, 1997; Veerkamp and de Haas, 2005). Genetic selection can be performed by direct or indirect selection, or by a combination of both approaches. The choice of approach depends on the breeding goal, the availability and accuracy of records, the population structure and the genetic parameters of goal and indicator traits.

The need for genetic parameters

The genetic parameters for a trait, which are calculated from variances and covariances obtained in statistical analyses of phenotypic records, are essential in animal breeding. The heritability of a trait describes how much of the total phenotypic variation is explained by additive genetic variation, and is thus a measure of the inheritance of the trait. It determines whether genetic selection is possible and, if so, how it would best be performed. It also gives an indication of how much genetic progress to expect. Genetic correlations measure the strength of the association between traits and can, for example, be used to predict unfavorable or favorable correlated responses, and consequently also to decide which traits to include in the breeding goal and whether indirect selection can be applied. Genetic parameters are also needed to predict breeding values to be used in the ranking and selection of superior animals for breeding. These parameters are only valid in a certain population, can change with time, and should therefore be re-estimated regularly.

The unfavorable genetic relationship with milk production

Estimates of the genetic correlation between milk production and mastitis susceptibility based on Nordic data and summarized by Heringstad *et al.* (2000) ranged from 0.24-0.55, with an average of 0.43. Other studies have reported estimated genetic correlations between production traits and CM within the same range (Rupp and Boichard, 1999; Hansen *et al.*, 2002; Hinrichs *et al.*, 2005; Negussie *et al.*, 2008). The genetic correlation between production traits and SCC is also in general positive, although weaker (especially in later lactations). An average estimate of 0.14 was reported for first lactation in a review by Mrode and Swanson (1996). The positive and thus unfavorable genetic correlations between CM or SCC and milk yield emphasize the need to include mastitis resistance in the breeding goal to prevent a decreased genetic level of resistance to mastitis as a

consequence of selection for yield only (e.g. Emanuelson, 1997). Such deterioration in resistance to mastitis was predicted some decades ago, and although they are not very large per year, the effects here could be of considerable concern in a long-term perspective. Selection based on both yield and mastitis traits was shown to counteract (diminish) this deterioration and increase the economic response when compared with selection for yield only (Strandberg and Shook, 1989; Colleau and Le Bihan-Duval, 1995).

Direct selection based on clinical records

In direct selection the actual trait of interest is measured. In the case of mastitis resistance, direct selection could be based on clinical mastitis records or bacteriological test results. One advantage of using bacterial infection is that it gives an indication of both subclinical and clinical cases (Weller *et al.*, 1992), and some studies have recently been performed considering pathogen-specific mastitis (e.g. de Haas, 2003; Holmberg, 2007). Bacteriological testing is, however, not practical on a large scale, and therefore the most common option is to use clinical records (Emanuelson, 1997).

Currently, only the Nordic countries (Sweden, Denmark, Finland and Norway) have well-established national health-recording systems and include CM directly in their breeding programs. In all these countries data from the health-recording system is combined with data from milk-recording and AI records to create a single data-base to be used for both management and selection purposes. In Sweden, approximately 85% of cows are affiliated in the cow data base (Swedish Dairy Association, 2007b). The nation-wide health-recording systems in Norway, Finland, Sweden and Denmark were introduced 1975, 1982, 1984 and 1990, respectively; recordings of treatments are primarily performed by veterinarians, but in Denmark and Finland also, to some extent, by farmers (Heringstad *et al.*, 2000). In some other countries records of CM are currently available on a limited scale (i.e. from research herds or selected commercial herds), and so far they have only been used for research purposes (e.g. Zwald *et al.*, 2004; Hinrichs *et al.*, 2005).

The CM records traditionally used for genetic selection fail to distinguish between both different types of mastitis and degrees of severity. Despite the quantitative genetic background and, probably, an underlying continuous liability to CM, the phenotypic expression is categorized in two distinct classes (diseased if liability exceeds a certain threshold). CM is therefore considered a so-called threshold, or categorical, trait, and this imposes

certain restrictions on statistical analysis (e.g. Gianola, 1982). The heritability for CM is often around 2-4% on the observed scale (Heringstad *et al.*, 2000; Hansen *et al.*, 2002; Interbull, 2008) and between 6-12% on the underlying scale (Heringstad *et al.*, 2000; Zwald *et al.*, 2004; Hinrichs *et al.*, 2005). Despite the low heritability for CM, which could in part be connected with the all-or-none character of the trait, accuracy of selection can be quite high and genetic progress significant provided that progeny group size is large enough (Emanuelson *et al.*, 1988; Shook, 1989; Heringstad *et al.*, 2000).

Indirect selection measures

In indirect selection an indicator trait is measured. In view of the lack of CM records in most countries, this approach is commonly used in efforts to bring about genetic improvements in mastitis resistance. A good indicator trait should have a higher heritability than, and be highly genetically correlated with, the goal trait; ideally data on it ought to be easy to measure and collect (Mrode and Swanson, 1996).

Somatic cell count

SCC has several desirable attributes as an indicator trait for CM, and its use for this purpose is therefore widespread (Interbull, 2008). Estimates of the heritability for lactation-average SCC are higher than those for CM and usually within the range of 0.1-0.2. The genetic correlation between SCC and CM is moderate to high (often around 0.6-0.8), suggesting that genes predisposing cows to a low SCC also result in a lower rate of CM (e.g. Mrode and Swanson, 1996; Heringstad *et al.*, 2000; Rupp and Boichard, 2000; Hinrichs *et al.*, 2005; Negussie *et al.*, 2006). Other advantages are that SCC data are easily and objectively measured on a continuous scale and tend to be normally distributed when transformed to a logarithmic scale; they are also readily available at a low additional cost in most milk-recording schemes, and they reflect both clinical and subclinical mastitis (Philipsson *et al.*, 1995; Mrode and Swanson, 1996; Heringstad *et al.*, 2000).

One concern about genetic selection for reduced SCC has been that it might reduce not only susceptibility to mastitis infection but also the cow's ability to respond to infection (Kehrli and Schuster, 1994; Schukken *et al.*, 1997). Linear relationships between sires' breeding values for SCC and CM have, however, been reported (i.e. the lower the SCC the lower the CM incidence) (Philipsson *et al.*, 1995; Nash *et al.*, 2000; Negussie *et al.*, 2006). Hence SCC should be decreased to the lowest possible value at least within the range covered by the population mean and the genetic variance

(Emanuelson, 1997; Veerkamp and de Haas, 2005). Further, in practice selection schemes aim merely to set aside bulls that transmit the highest SCC values; such a selection will hardly erode genetic resistance to mastitis (Philipsson and Lindhé, 2003). The traditional use of lactation average SCC has some shortcomings when it comes to predicting CM, and thus alternatives measures based on test-day records of SCC have been proposed to better model the dynamics of SCC in connection with CM cases (e.g. de Haas, 2003).

Other indicator traits

Udder conformation is the second most common indirect trait for mastitis resistance used today. The relationship between various udder type traits and CM or SCC has been investigated, but genetic correlations are generally low and results are somewhat inconsistent. Udder depth and fore udder attachment seem to be most frequently associated with mastitis resistance (Mrode and Swanson, 1996; Rupp and Boichard, 1999; Nash *et al.*, 2000; Rupp and Boichard, 2003). Results suggest that selection for a higher, and more tightly attached, udder will improve resistance.

Other traits that have been found to be associated either with CM or with SCC are milking speed, electrical conductivity in milk (Norberg, 2004) and markers of immune response. Faster milking has been reported to be genetically associated with increased SCC (Luttinen and Juga, 1997; Boettcher *et al.*, 1998; Rupp and Boichard, 1999), but several studies have indicated the opposite relationship with CM, and thus that slower milking increases CM (review by Rupp and Boichard, 2003).

Combining direct and indirect measures

In view of the complexity of mastitis resistance, a combination of direct and indirect measures merging different aspects of udder health in an udder health index probably represents the best approach to genetic selection (Schukken *et al.*, 1997). In connection with information on CM and SCC, a combination of both measures has been proven to be most efficient irrespective of daughter group size. Where only one of the traits was considered, SCC was more efficient than CM when daughter groups were small (less than 100), whereas the opposite was true for larger daughter groups (Philipsson *et al.*, 1995). When no information on CM is available, the efficiency of selection can be improved by combining SCC, udder type traits and milking speed (de Jong and Lansbergen, 1996; Boettcher *et al.*, 1998).

Genetic evaluation of clinical mastitis

In the Nordic countries, there is a long tradition of employing genetic evaluation (GE) systems for health and fertility traits, and of using a total merit index (TMI) for bulls in which the most important production, reproduction and health traits are combined together on the basis of relative weights. Selection based on TMI has proven effective in maintaining the functionality of cows while simultaneously increasing production. In Sweden predicted breeding values (PBV) for CM were first published in 1984. They were subsequently included in the TMI, which had been in place since 1975. Similar developments occurred in the other Nordic countries (Heringstad *et al.*, 2000; Philipsson and Lindhé, 2003). Positive responses to selection for mastitis resistance have been reported in all these countries, as summarized by Philipsson and Lindhé (2003).

Current genetic trends for CM differ, however, between countries and breeds. In the Nordic cattle genetic evaluation (NAV), which is a joint evaluation for Sweden, Denmark and Finland, an unfavorable trend can be seen for Holstein, whereas there is no obvious trend for the Red breeds (Johansson *et al.*, 2006). This is illustrated in Figure 2 where the expected increase in CM frequency in the Holstein breed is about 2-3% of daughters from sires born in 1986 compared to 2000. The difference between breeds can be explained by the extensive use, in the Holstein breed, of bulls from North America, where greater emphasis has been put on production than functionality, and on the long tradition of including mastitis resistance in breeding goals adopted in the Nordic countries. Genetic improvement of CM in the Norwegian dairy cow population has been shown for cows born after 1990 (Heringstad *et al.*, 2003a). This is probably the result of the fact that in Norway, as compared with NAV countries, larger daughter groups are used and greater, and increasing, weight has been given to CM in the TMI.

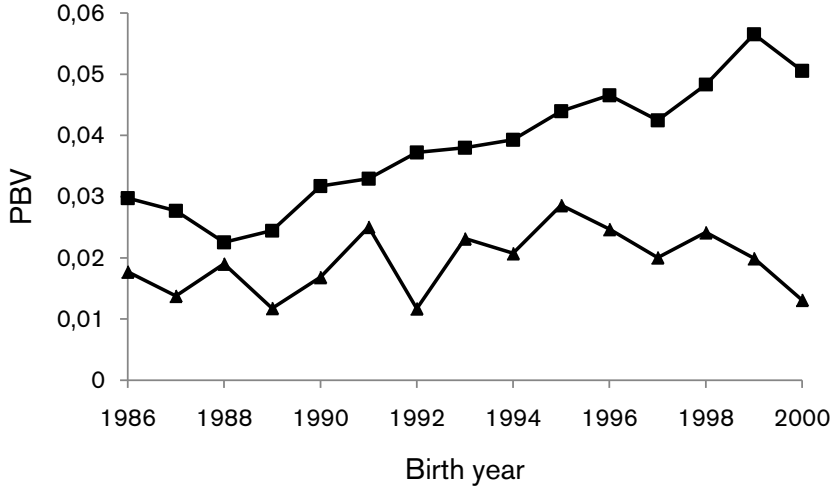

Figure 2. The genetic trend for binary CM in NAV shown as average predicted breeding values (PBV) estimated from linear cross-sectional models for sires with different birth years of the Holstein (■) and the Red breeds (▲). Lower values are favorable, thus indicating less susceptibility to CM (reproduced from Johansson *et al.*, 2006).

Today many countries recognize the benefits of broad and more sustainable breeding goals that incorporate mastitis resistance for their dairy cattle populations, and GE for udder health traits are carried out on both national and international level. In the most recent Interbull routine international genetic evaluation for SCC and CM (August 2008), data were provided by 19 and 2 (groups of) countries, respectively (Interbull, 2008). Nowadays some aspect of udder health is often also included in the TMIs of countries other than the Nordic countries.

Current practice and its limitations

The trait definition of CM for the national GE in Sweden until 2006 was to include veterinary treatments, as well as culling, resulting from mastitis, and to define CM as a binary trait in the period from 10 days before to 150 days after first calving. To increase heritability it is important to include information on culling for mastitis (Koenen *et al.*, 1994) as well as mastitis occurring before calving (Heringstad *et al.*, 2001). The restricted time period was introduced to ensure that cows had a more equal opportunity period and to reduce bias resulting from culling in the later part of the lactation. Lactation average SCC in first lactation was used as an additional source of information to increase the accuracy of CM (Koenen *et al.*, 1994).

Routine GE for mastitis in NAV started in 2006. The approach is to define CM as a binary trait within two defined periods of first lactation (-15 to 50 and 51 to 300 days in milk (DIM), respectively) and one defined period of second and third lactation (-15 to 150 DIM). A multiple-lactation multiple-trait linear sire model is applied in connection with the CM traits in combination with lactation average SCC (5 to 170 DIM) from the first three lactations and udder depth and fore udder attachment from the first lactation. A similar practice for GE of CM exists in Norway, although the trait definition and model are somewhat different to those in NAV (Interbull, 2008).

The traditionally used definition, where the cow is treated as healthy or sick depending upon whether a single CM case occurs over a rather long period, is a so-called cross-sectional approach, and therefore the LM used can be called a linear cross-sectional model (CSM). This method has some obvious disadvantages for GE of CM, mainly connected with insufficient use of the available information and improper handling of ongoing and incomplete records. Therefore alternative, and theoretically better, methods such as threshold CSM, survival analysis (SA) and longitudinal models have been suggested as potentially improving on the GE of CM, as summarized by Mark (2004).

Aims of the thesis

The overall aim was to advance our understanding of the genetic background of mastitis resistance in dairy cattle and, especially, to improve the genetic evaluation of CM both by utilizing more of the available information and by applying appropriate methodology. The assignment to animals of precise breeding values can contribute to genetic progress on mastitis resistance and thus improve both the health status of the cow and the economic situation of the farmer. The more specific objectives were to:

- Estimate heritabilities for CM and SCC, and genetic correlations between these udder health traits and production traits, within and between lactations with the currently used linear model methodology
- Investigate with both field data and a simulation study whether the method of survival analysis would result in a more precise genetic evaluation of CM than the current method
- Explore the feasibility of using a linear random regression model for the genetic evaluation of longitudinal CM data

Summary of the investigations

Materials

Data

Field data on Swedish Holstein cows from the Swedish official milk-recording scheme was used in Papers I, II and IV. The same data set was used for Papers I and II; it included the first three lactations of cows having their first calving between 1995 and 2000, which corresponded to about 220 000 cows. In Paper IV, only the first lactation from cows with first calving between 1998 and 2000 (about 90 000 cows) was considered in order to limit the material to a manageable size. Pedigree was traced back as far as possible, resulting in pedigree files on about 540 000 animals for Paper I, 1100 sires for Paper II and 800 sires for Paper IV.

Simulated data on first-parity cows were used in Paper III, and two different simulation structures were considered. In the first, each replicate consisted of 60 000 cows that were the daughters of 400 unrelated sires and distributed over 1200 herds. This structure, with an average daughter group size of 150 and a fixed herd size of 50, was similar to the current situation in Swedish field data. In the second structure, the same numbers of sires and herds were considered, but in connection with a herd size of 20, which resulted in an average daughter group of 60 and a total of 24 000 cows. Fifty replicates were performed for each population structure.

Trait definitions

CM was the only trait analyzed in Papers II-IV, but with some different trait definitions. Paper I, on the other hand, included the traits CM and

lactation average SCC, as well as 305-day milk, fat and protein yield. The following definitions of CM were used:

MAST = a binary trait distinguishing between cows with at least one reported case of mastitis (1) and cows with no cases (0) during defined periods of the lactation. In the field data, cases included veterinary-treated CM or culling for mastitis. For Papers I-III, the time periods used covered a larger part of the lactation, namely: 10 days before to 150 days after calving (Papers I and II) and the day of calving to either 150 days after calving or the end of the lactation (Paper III). For Paper IV, a longitudinal approach was considered and MAST in shorter intervals (4 1-week followed by 8 4-week intervals) from 10 days before to 241 days after calving were considered either as 12 separate traits to be analyzed with linear longitudinal multivariate models (LMVM) or as 12 repeated observations of the same phenotypic trait creating a series of binary responses to be analyzed with a linear random regression model (RRM).

TFM = time to first clinical mastitis, which was treated as the number of days from a starting point to the first case of mastitis in lactation. In Paper II, it was defined as the number of days from 10 days before calving to either the day of first treatment of CM or culling because of mastitis. In Paper III, it was the number of days from the day of calving to first mastitis case occurring within either 150 days or the complete lactation. Observations from cows without cases of mastitis were considered as censored and the time was defined as the number of days from the starting point to either the day of culling for other reasons than mastitis, lactation day 150 (Paper III; shorter opportunity period) or the day of next calving. Cows in Paper II without cases and no information on a second calving or culling date received a stop time at lactation day 240 based on the assumption that the risk of a cow being culled increases at around that point in lactation.

Figure 3 shows an example of the definition of CM in CSM and SA (as defined in Paper II) and in longitudinal models (as defined in Paper IV) for a cow with two cases as well as for two cows without cases but either next calving or culling.

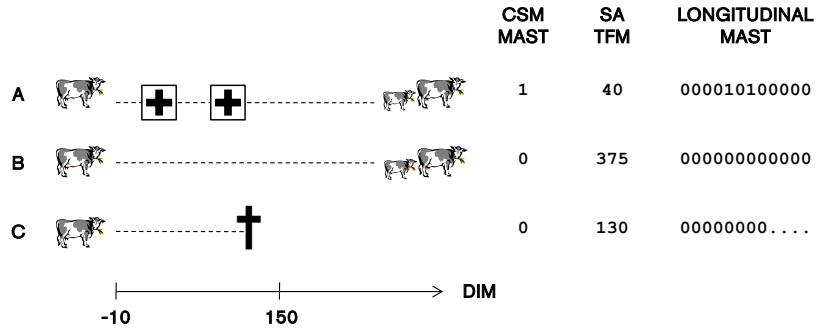

Figure 3. Illustration of the trait definition used for CM when treated as MAST in CSM or longitudinal models and as TFM in SA for a diseased cow with CM cases at lactation day 30 and 100 (A), and for cows without CM with either next calving at lactation day 365 (B) or culling at lactation day 120 (C).

Methods

Mixed linear models (LM) were used to analyze MAST (Papers I-IV) as well as SCC and production (Paper I). Estimations of (co)variance components and predictions of breeding values were performed with the software package DMU (Madsen and Jensen, 2008) using the average information algorithm for restricted maximum likelihood. Within the framework of LM, different approaches for the MAST traits were considered in the different papers. An animal model was used in Paper I, whereas sire models were used in Papers II-IV. Further, either CSM (Papers I-IV) or the longitudinal models LMVM and RRM (Paper IV) were applied. SA was used to analyze TFM (Papers II and III) with Weibull proportional hazards sire models. Here the estimates and predictions were obtained using the package of Survival Kit (Ducrocq and Sölkner, 1998). In Paper III, MAST was also analyzed with a cross-sectional threshold model (TM) in the software program ASREML (Gilmour *et al.*, 2002).

All the models for CM contained similar effects, although the build-up of the SA and RRM are somewhat different from the traditional CSM. Apart from the random genetic effect of animal (Paper I) or sire (Papers II-IV), all models for field data contained the fixed effects of age and (year-) month at calving as well as the effect of herd-year at calving, which was either treated as fixed (Paper I) or random (Papers II and IV). In addition, fixed regressions on the proportions of heterosis and North American Holstein genes were included in some models (Papers I and II) to account for the impact of foreign Holstein sires on the Swedish Holstein dairy cow

population. The model used for simulated data was simpler and contained only the random effects of sire and herd.

In the Weibull proportional hazards model, the dependent variable analyzed, $\lambda(t)$, was the hazard of a cow developing CM at time t given that it had not occurred prior to t. This was modeled by $\lambda_0(t)\exp\{model\}$, where $\lambda_0(t)$ is the Weibull baseline hazard function $(\lambda\rho(\lambda t)^{\rho-1})$ with scale parameter λ and shape parameter ρ. The value of ρ indicates whether the hazard is constant (=1), or increasing (>1) or decreasing (<1) with time.

An RRM used for longitudinal observations requires some additional effects. In our RRM (Paper IV) we modeled the phenotypic trajectory of CM over time with a fixed lactation stage effect and deviations around this trajectory with a random regression function for each sire using orthogonal Legendre polynomials as time covariables. In recognition of the repeated observations for a cow, a random effect of permanent environmental effect within cow was added.

For a comparison of the methods of genetic evaluation in Papers II-IV, estimates of heritability, accuracy of selection, and genetic correlations, as well as correlations between breeding values, were used. Pearson product-moment correlations (SAS, 2002) between PBVs for CM from different methods (Papers II-IV) and between PBVs for CM and true breeding values (TBV) for mastitis liability (Paper III) were calculated for different time periods. A correlation less than unity implies re-ranking of the evaluated sires and indicates a difference in PBVs estimated from the different methods.

Main findings

Genetic parameters

Heritability of CM and the relationship to other traits

Estimates of the heritability for MAST from linear CSM in Papers I-III were low and varied between 1-4%, mainly depending on which lactation and opportunity period was being considered. Higher estimates were obtained for first lactation than later lactations (Papers I and II) and for a longer opportunity period (Paper III). The heritability estimates from SA were of similar magnitude to those from linear CSM (Paper II). Somewhat higher (7-8%) and lower (0.1-2%) estimates were found from TM (Paper III) and linear longitudinal models (LMVM and RRM; Paper IV), respectively. These estimates cannot be compared, however, with those

from linear CSM and SA, because of the different scales or trait definitions. It was confirmed in Paper I that a rather strong and unfavorable genetic correlation (0.3-0.5) exists between CM and milk production. This emphasizes the need to include udder health traits in the breeding goal to prevent deterioration of mastitis resistance as a consequence of selection for production only. The genetic correlation between CM and protein or fat production was also unfavorable, but to a smaller degree, especially for fat production. Further, it was demonstrated that the higher heritability for lactation average SCC than for CM, and its high genetic correlation to CM, makes it a suitable trait to use for indirect selection to improve mastitis resistance. The accuracy of selection when the true breeding goal was defined as freedom from clinical cases of mastitis was naturally highest when both measures, CM and SCC, were combined, irrespective of daughter group size. Figure 4 summarizes the genetic parameters estimated for udder health and production traits in Paper I.

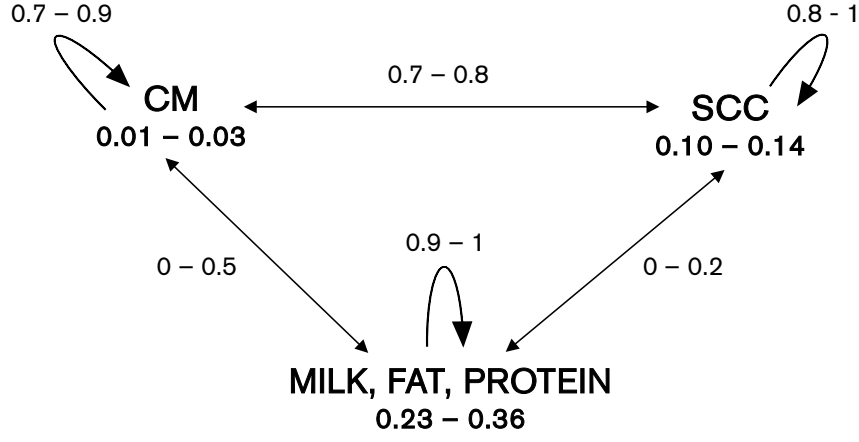

Figure 4. A summary of the heritabilities (in bold) for udder health and production traits and the genetic correlations between traits and within each trait across lactations for the first three lactations of Swedish Holstein cows, as estimated in Paper I with a linear CSM.

Genetic relationships of CM across and within lactations

CM in the first three lactations (up to 150 days) in Paper I had a frequency that increased from 10-15%, and was proven to represent somewhat different traits genetically, with genetic correlations across the lactations varying between 0.7 (lactations 1 and 3) and 0.9 (lactations 2 and 3). Genetic evaluation of mastitis resistance with the aim of improving the

situation in all parities should therefore include all available information rather than information from the first lactation only.

Within lactation, as well, CM can be considered as a series of genetically distinct traits appearing in different time periods. Genetic correlations between selected days occurring between day of calving and lactation day 200 in first lactation varied considerably in both linear RRM and linear LMVM, indicating that early (first 30 days) and late (after 140 days) CM are more highly genetically correlated (>0.6) with each other than they are with the period in between (Paper IV). The correlations between sire PBVs for the time periods 10 days before to 50 days after calving and 51 to 150 days after calving in both linear RRM and linear CSM were far from unity (0.65 and 0.24, respectively), which supported the speculation that CM is not the same trait genetically throughout the lactation.

Comparison of methods for genetic evaluation

The methods of SA and LM for genetic evaluation of CM were compared in Paper II mainly on the basis of the accuracy of selection and correlations between PBVs from the two methods. It was possible to compare the accuracy of selection because the proportion of uncensored records in SA was accounted for in the calculations of the heritability, and it was found to be slightly higher in first lactation and considerably higher in second and third lactation for SA compared with LM (3% and 25% higher, respectively). Correlations between sire PBVs from SA and LM were 0.93, 0.89 and 0.88 for lactations one to three, respectively; the correlations of less than unity indicated that some re-ranking of sires occurred when the different methods were used.

The comparison of SA and LM in Paper II was complemented by a simulation study in Paper III; the latter was also extended to include TM and two different lengths of the opportunity period. The main measure of comparison in this study was the correlation between simulated sire TBVs for mastitis liability and sire PBVs from LM, SA and TM, respectively, which can be seen as the true accuracy. Given the simulated conditions, the method used for GE had no effect on accuracy, when comparisons were made within the opportunity period. The correlation between TBVs and PBVs was 8% greater when the full lactation data were used (0.76) than it was when data were obtained from the first 150 days (0.70) with an average of 150 daughters per sire. The corresponding results were 0.60 and 0.53, respectively, with an average daughter group size of 60. The best sires ranked on PBVs from the full lactation data had an average true genetic merit that was lower (implying less mastitis) than the best sires ranked on

data from the first 150 days, and vice versa for the worst sires. Further, a greater proportion of sires were correctly ranked with the full lactation data. The results also showed that the worst sires, with a large proportion of daughters with mastitis, obtained more precise breeding values, and were ranked more correctly, than the best sires, regardless of method and opportunity period used.

The method of linear RRM was found to be rather unstable and sensitive when used for parameter estimation of binary CM data (Paper IV). However, the chosen model worked satisfactorily with the current data. The heritability curve from RRM for CM data up to 241 days after calving corresponded well with the point estimates for the separate intervals from linear LMVM, and the highest heritability (2%) was found at the beginning of the lactation, where the frequency of mastitis was highest. Also the patterns of genetic correlation for CM between selected days and the remaining part of the lactation from RRM and LMVM were in rather good agreement, especially for time periods with higher genetic variation. Another informal validation of the RRM was obtained from the correlations between summarized PBVs from RRM and PBVs from linear CSM for the time periods 10 days before to 150 days after calving, 10 days before to 50 days after calving and 51 to 150 days after calving: these were 0.96, 0.92 and 0.74, respectively. Thus, only some re-ranking among sires occurred, and the re-ranking that did occur happened more when the late time period was considered on its own. That the best agreement between methods in predicting breeding values was found for the full or early time periods was also demonstrated by a greater number of common sires and higher rank correlations among the best and among the worst sires, respectively, for these time periods compared to the late time period. The agreement was better for the worst sires. This was also found to be the case in Paper III.

General discussion

The general, and nowadays widely accepted, view in the dairy cattle industry is that productivity and functionality must go hand in hand to provide us with more robust cows. It is necessary to include functional traits, such as mastitis resistance, in the breeding goal if we are to prevent deterioration of the kind consequential upon selection for production only, as well as to meet the economic, animal welfare and ethical concerns of farmers, consumers and the society. The results presented in this thesis confirm the need for, and the potential of GE of CM through the following main findings: 1) there is an unfavorable genetic correlation between CM and milk production which emphasizes the need to include CM in the breeding goal; 2) heritability for CM is low, but considerable genetic variation exists and therefore genetic selection is possible; 3) SCC is more heritable and is strongly genetically correlated to CM, making it a suitable trait for indirect selection; 4) the accuracy of selection is greatest when information on CM and SCC is combined; 5) CM is a different trait genetically across and within the lactations which should be considered in GE; 6) the length of the total opportunity period for CM influences the amount of information and thus the accuracy of selection; 7) longitudinal models for GE of CM seem more biologically sensible than cross-sectional models although they might have practical limitations; and 8) fair comparison of the methods of GE is not easy to accomplish, and decisions about which method is the "best" for CM are therefore somewhat arbitrary. The first four findings were expected and confirm results from several previous studies. The last four findings, together with related issues, will be discussed in more detail below.

Prospects for improved genetic evaluation of clinical mastitis

The first prerequisite of an efficient GE of CM is to have good quality data from clinical records available on a large-scale basis and to use large daughter groups. Further, it is important to use the most appropriate methodology and trait definition, and preferably a method that can be applied easily in practice in connection with large data sets and in a multiple trait setting. Although the currently used linear CSM is easily applicable, it has some obvious disadvantages. These have been discussed thoroughly in the literature (e.g. Heringstad *et al.*, 2000 and 2003b; Papers II-IV). Several alternative, and theoretically better, methods of GE of CM have therefore been investigated.

Disadvantages with the current method

The main problem with a CSM is that the information on multiple cases and the timing of a case is ignored, although it is readily available from health-recording data. Another problem is that ongoing and incomplete records cannot be treated properly, and this either further reduces the amount of information or can potentially introduce bias (e.g. Heringstad *et al.*, 2001). Loss of information can occur when such observations are treated as missing. The other option is to include ongoing and incomplete observations with no CM in the analyses as "healthy". This, however, might overestimate certain animals, because a shorter opportunity period does not give the cow the same opportunity to express disease and a cow could be culled for something correlated with CM, such as high SCC. In a CSM, CM is taken to be the same trait genetically for the whole period, and thus a constant genetic value over time is assumed. It has been shown that this is a simplification of this complex trait (e.g. Paper IV). Another disadvantage of the linear CSM is that the assumption of normally distributed data is not fulfilled. This has, however, been shown to be of less importance, at least in the GE of sires (Meijering and Gianola, 1985).

Comparison with alternative methods

Threshold cross-sectional models (Paper III)

A TM takes the binary character of a categorical trait into account and would therefore be preferable to an LM, at least in theory (Gianola, 1982; Gianola and Foulley, 1983). For GE of CM, a threshold CSM in which the unobserved and underlying continuous liability to CM is modeled rather than actual phenotypic outcomes has been used rather extensively in research (e.g. Kadarmideen *et al.*, 2000; Hinrichs *et al.*, 2005; Paper III;

Zwald *et al.*, 2006). Compared to the linear CSM, however, this method has some computational disadvantages, does not use more of the available information and is not in routine use for CM (Interbull, 2008).

Despite the greater heritability obtained when CM is defined on the underlying scale, as happens in TM, selection based on PBVs from the TM may not yield higher genetic progress on the observed scale than selection on PBVs from LM. High correlations (>0.98) between sire PBVs from the two models have been reported for categorical fertility and survival traits (e.g. Weller and Ron, 1992; Boettcher *et al.*, 1999). This implies that very little re-ranking was expected to occur when replacing one of the models with the other. This was confirmed in connection with CM in Paper III, where there was a nearly unity correlation between PBVs from linear CSM and threshold CSM, and where the true accuracy (correlation between TBVs and PBVs) was very similar to the theoretical accuracy (based on heritabilities and number of daughters per sire) for the linear CSM but not for the threshold CSM. Thus, the calculation of the accuracy of selection on the basis of heritability on the underlying scale from a TM overestimates the true accuracy of selection. This has also been concluded by Foulley (1992).

Survival analysis (Papers II and III)

Survival analysis is a statistical method used to study the occurrence and timing of specific events. Some advantages of SA over CSM are that it utilizes more of the available information, and that time-dependent effects as well as censored observations, when a competing event occurs before the event of interest, can be included in it (Ducrocq, 1987). In the field of dairy cattle breeding it is routinely used for GE of longevity in many countries (Interbull, 2008). It has also been used successfully for other traits with a longitudinal character, including interval fertility traits (Schneider, 2006). The use of SA to analyze TFM has also been reported (Saebø and Frigessi, 2004; Saebø *et al.*, 2005; Papers II and III), and although SA deals with some of the problems connected with the CSM, such as the timing of the first CM in lactation and the handling of incomplete records, it ignores multiple cases.

In Paper II, it was concluded that the accuracy of selection was higher for the trait TFM analyzed with SA than for MAST analyzed with linear CSM, especially in later lactations. This may possibly translate into an increase in genetic progress. However, the trait definition that was used implied that the opportunity period used for MAST was restricted to lactation day 150 (as was done in the national GE at that time), whereas this was not the case for TFM, where the full lactation period was used. The

justification for imposing the restricted opportunity period was based on the notion that the linear CSM cannot properly deal with ongoing and incomplete records, and that a restricted time period was believed necessary to give cows more equal lengths of opportunity periods and to reduce any bias resulting from culling.

The use of the same trait definitions in a simulation study (Paper III) in which the correlation could be calculated between TBVs for mastitis liability and PBVs from linear CSM, threshold CSM and SA, gave a similar picture favoring SA. In the simulation study, however, two additional trait definitions were added - namely, TFM restricted at day 150 for SA and MAST in the full lactation for the two CSM. This led to a modified conclusion that within the opportunity period, the method used had no effect on accuracy, but the full opportunity period resulted in a higher true accuracy (8%) than the 150-day period regardless of method used. The reason for this is probably that the longer opportunity period gave a better opportunity for TBVs to be expressed, indicating that the difference in average incidence of CM in the restricted (10.7%) and full period (16.7%) caused the difference in the results.

We had expected SA to be favorable over CSM also when the same opportunity period was considered as a result of the better use of available information (something that increased the observed variation among diseased and among healthy cows). The rather low CM frequency, resulting in a high proportion of censored observations, and the fact that most CM cases occur around calving (Figure 1), giving TFM a close to binary nature, probably contributed to the negligible difference between SA and CSM when equal opportunity periods were considered. Despite the longer opportunity period for SA than for linear CSM in Paper II, a rather small difference was observed between methods in the first lactation. The finding that the accuracy of selection was considerably higher (25% higher) for SA in later lactations could perhaps be related to the fact that both CM and SCC, as well as the risk of culling because of udder health problems (Roxström and Strandberg, 2002; Schneider *et al.*, 2007), increase with increasing lactation. An increased proportion of cows culled because of high SCC but without records of CM could introduce bias in CSM, whereas SA would be less affected. The simulation study presented in Paper III did not consider culling connected with high SCC or other factors related to mastitis.

An attempt to test this hypothesis in a simulation was made by adding the trait SCC, which had a positive genetic correlation with mastitis liability, and by building the assumption that cows could be culled because

of high SCC in the later part of the lactation. If these cows did not have CM before the time of culling, observations were included as "healthy" in the linear CSM and as censored in the SA. The correlations between TBVs for mastitis liability, on the one hand, and PBVs, from either linear CSM or SA, on the other, did not differ much when compared within opportunity period (restricted and full, respectively). Some of the differing simulated settings and the outcomes for the full opportunity period are shown in Table 1. Judging by this result, it seems either that no bias is introduced in the CSM or that any that is has no effect on the PBVs - at least, under the simulated settings. If this is the true scenario, the only reason for using a restricted time period for a CSM would be to give cows more equal opportunity periods.

Table 1. *The correlation between true breeding values (TBV) for mastitis liability (ML) and predicted breeding values for clinical mastitis from linear cross-sectional model (CSM) and survival analysis (SA) in a simulation where culling for high SCC was either excluded (-) or included with different values of the genetic correlation with ML, the threshold over which cows could be culled and the mean time of culling for SCC*

SCC parameters			Correlation with TBV	
Genetic correlation with ML	Threshold (no of cells)	Mean time of culling (SD)	CSM	SA
-	-	-	0.755	0.757
0.5	730 000	250 (110)	0.764	0.768
0.7	730 000	250 (110)	0.767	0.770
0.7	500 000	250 (110)	0.769	0.775
0.7	300 000	250 (110)	0.752	0.759
0.7	300 000	200 (100)	0.748	0.759

Longitudinal models (Paper IV)

Longitudinal models such as repeatability models, LMVM and RRM can deal with repeated observations of individuals over time. Test-day repeatability models for GE of CM have been reported (e.g. Hinrichs *et al.*, 2005), but, as similarly happens in a CSM, a single genetic value of animals over time is assumed. The two latter longitudinal models have the advantage that they can deal with different gene expressions over time, and are therefore preferable where such a difference is believed to be the case. RRM have become very popular for GE of test-day data such as milk production and SCC in dairy cattle (Schaeffer, 2004; Interbull, 2008). They have also, though to a lesser extent, been used for longitudinal binary data. Both LMVM and RRM have been reported for GE of CM, and these

models deal with many of the disadvantages connected with the CSM by including information on multiple cases and the timing of the cases, and by offering a more proper handling of incomplete records. Another advantage is the ability to describe genetic and environmental effects over time. However, these models are more computationally intensive because of the large number of records and increased number of parameters it is necessary to estimate, especially in LMVM (Jensen, 2001). In longitudinal studies, CM in defined shorter (e.g. monthly) intervals of lactation has been treated as separate but correlated traits in LMVM (e.g. Chang, 2002; Chang *et al.*, 2004a; Heringstad *et al.*, 2004), and as repeated observations over time in RRM (e.g. Heringstad *et al.*, 2003b; Rekaya *et al.*, 2003; Chang *et al.*, 2004b). All of these studies analyzed the liability to CM with a TM.

The feasibility of linear RRM for CM was investigated in Paper IV. Were this approach successful, it would be easier to implement for GE in practice - for example, in a bivariate setting with test-day SCC. There was a fairly good agreement between the time-dependent genetic parameters estimated from the chosen RRM and linear LMVM, and strong correlations were found between summarized PBVs from the RRM and PBVs from linear CSM, especially when the early part of lactation was considered. These results were used as an informal validation; they indicated that the chosen linear RRM worked satisfactorily for GE of CM. Although we initially expected the linear LMVM to give results that were closer to the "true" picture, it turned out that this method gave rise to more fluctuating results than the linear RRM, with point estimates for some intervals being biologically unrealistic. This phenomenon of jumpy estimates is possible because no structure is assumed for the (co)variances over time in the LMVM, as is done in an RRM (Jensen, 2001).

Although our informal validation of the finally chosen linear RRM worked, this method seemed very sensitive and unstable for parameter estimation, because preliminary analysis of different linear RRM in combination with different trait definitions for CM (i.e. length of intervals) gave rather different results and often erratic estimates or convergence problems. The problems were probably mainly a consequence of the very low overall CM incidence in our data, which becomes an even more prominent problem in a longitudinal analysis where more zeros than ones are added when shorter intervals are being created. The low incidence, in most intervals except the one around the day of calving, contributes to very low genetic variances and heritabilities that generate problems in the statistical analyses. Attempts to solve the puzzle by creating longer intervals, running bivariate linear RRM with longitudinal CM data and test-day

SCC, or by taking heterogeneous residual variance into account, were not successful. It is possible that a threshold RRM would be a better option for binary CM data, since heritability estimates from a TM are less frequency-dependent than those from an LM. In the studies using threshold RRM (Heringstad *et al.*, 2003b; Rekaya *et al.*, 2003; Chang *et al.*, 2004b) considerably larger incidences of CM than those in Paper IV were also reported, however – something that might have been beneficial in those analyses. The results presented in Paper IV suggest that a linear RRM can be expected to work well for GE, i.e. the prediction of breeding values, when the genetic parameters are known.

Issues related to comparing methods

The criteria available for comparing different methods suffer from some limitations, and choosing the "best" method – i.e. the method that gives the most accurate GE of CM - is further complicated by the many different biological, statistical and practical aspects involved. Another issue is that if models that are as similar as possible are used to achieve fair comparisons, the potential of a theoretically better model might not be fully utilized. An example here is that the opportunity to include time-dependent effects in SA was ignored in Paper II in order to generate a model that was more similar to the linear CSM, which meant that one of the distinct advantages of SA was discarded.

A comparison can be based on either field or simulated data. The advantage of a simulation study is that TBVs for the trait of interest can be simulated and correlated with PBVs from different methods, which will then reflect the true accuracy of selection of the alternative methods. The simulated trait should not favor any of the methods of comparison, which makes simulation a less than trivial process. This was attempted in Paper III, where the underlying mastitis liability was simulated from a Normal distribution rather than TFM directly from a Weibull distribution, thus creating a simulated process that was as separate from time-to-event and binary traits as possible. The disadvantage with a simulation study, however, lies in its simplification of reality; the relevant results should therefore preferably be supported with results from field data.

With field data, methods could be compared by the theoretical accuracy of selection (Paper II) or correlations between PBVs (Papers II-IV). However, as was concluded above, this accuracy is not a good measure to use when comparing LM and TM if the calculations are based on heritabilities on different scales. It is also difficult to compare heritabilities from different studies of CM using LM, because the estimates here are

frequency dependent (e.g. Emanuelson *et al.*, 1988). In Paper III the heritability from the linear CSM was more than 30% higher when the full (3.6%) rather than the restricted (2.7%) opportunity period was used. The same was true for SA. The heritability from the threshold CSM increased as well, but only by about 10% (rising from 7.4-8.2%). When the proportion of uncensored daughters is taken into account in calculating the heritability from SA, the resulting accuracy should be comparable to the accuracy obtained with LM (Yazdi *et al.*, 2002). Size of daughter group also affects the results, and with large daughter groups sires are expected to get rather accurate PBVs regardless of method used. This might be the reason why small differences between the linear CSM, threshold CSM and SA were observed (Papers II and III). It has been suggested that smaller daughter groups favor TM over LM. However, reducing the daughter group size from 150 to 60 in Paper III did not alter the conclusion that the method used (linear CSM, threshold CSM or SA) within opportunity period has no effect on the accuracy.

A correlation between PBVs from different methods less than unity indicates that re-ranking of the evaluated animals occurs when the different methods are used, but it does not tell us which of the methods best predicts the TBVs. The correlations between PBVs from linear CSM and SA varied between 0.88-0.93 in Paper II, but then different lengths of opportunity periods were considered. In the simulation study (Paper III) the corresponding correlation was 0.9, but using the same opportunity period for linear CSM and SA (either 150 days or full lactation) the correlations were above 0.99. In Paper IV, correlations between summarized PBVs from linear RRM and linear CSM were 0.96, 0.92 and 0.74 for the time periods 10 days before to 150 days after calving, 10 days before to 50 days after calving and 51 to 150 days after calving, respectively. Thus, it is clear that comparison of methods for GE of CM is influenced by the trait definition and opportunity periods used for CM.

Time aspects of great importance

In obtaining a precise GE of CM, time aspects seem to be of great importance. A longer total opportunity period, regardless of method, and a longitudinal rather than a cross-sectional approach, increased the amount of the available information used and had an impact on genetic parameters, accuracy of selection and PBVs.

Length of opportunity period

The results presented in this thesis suggest that length of the opportunity period is more important for GE of CM than the method used. An increased heritability with a longer sampling period for CM was also shown in a study by Heringstad *et al.* (2001), in which the heritability on the observed scale was 0.03 and 0.04 for the time periods from 15 days before to either 30 or 210 days after first calving, respectively. The favorable effect of a longer period could, however, be counteracted by a longer generation interval, and waiting for additional results is probably not warranted. Of the methods compared in Papers II and III, only SA can theoretically account for the more variable opportunity periods for cows (more ongoing and incomplete records) arising when a longer time period is considered. Longitudinal models would have the same advantage as SA in this aspect, whereas for CSM there is greater concern about unequal opportunity periods and the risk of introducing bias as a result of any culling that occurs in the later part of the lactation. Different strategies for handling cows culled before the end of the sampling period have been shown to affect heritability estimates for CM (Heringstad *et al.*, 2001). The speculation that culling would introduce bias in GE for CM was not, however, confirmed by the results presented in this thesis.

Further work, extending that undertaken in Papers II and III, was done in an effort to clarify the importance of method versus opportunity period used in the comparison of linear CSM and SA. With field data, additional analyses of SA with TFM restricted at day 150 were performed and, based on estimated heritabilities and accuracies in selection, there was no or only a marginal advantage of SA over linear CSM using equal opportunity periods in first and later lactations, respectively. The correlations between PBVs from the two methods were all above 0.99. For purposes of GE it could be beneficial to get the CM information earlier in the lactation, when milk results are available, as this would decrease the generation interval. This would, however, adversely affect the accuracy of selection. In this scenario, the feature of censoring in SA is believed to be even more important. In the original simulation (without culling because of high SCC) an opportunity period up to day 50 was therefore considered, but SA still had no advantage over the linear CSM. These additional results confirmed the conclusion that opportunity period seems to be more important than the choice of method.

Number of clinical mastitis cases

The appeal of an increased length of opportunity period lies in the fact that it allows more information, i.e. a larger number of CM cases, to be

gathered. If the trait MAST is being considered, the additional information comes from a higher proportion of cows with at least one case of CM. Although there is a peak of CM cases in early lactation, the period up to 150 days of lactation actually only captured 60-65% of the total number of cases in the first three complete lactations (Paper I). The total number of CM cases (- NMAST, as opposed to the binary MAST) in the complete lactation has been analyzed in a few studies (Emanuelson *et al.*, 1988; Sander Nielsen *et al.*, 1997; de Haas, 1998; Heringstad *et al.*, 2006). NMAST has been limited in this way in response to the problem of censoring of information related to the fact that the risk of culling increases as NMAST increases (Heringstad *et al.*, 2006). De Haas (1998) achieved heritabilities of 4 and 3% for NMAST and MAST, respectively, with linear CSM. However, since the corresponding opportunity periods differed, being either the complete or 150 days of lactation, it is not possible to determine whether the increase in heritability for NMAST resulted from extra information relating to additional cases per cow or additional cows with CM. The latter might contribute more to the increase in information, since a rather small proportion of cows with CM have more than one case (de Haas, 1998; Heringstad *et al.*, 2006; Paper IV) and an increase in heritability of the same size was achieved when extending the length of the opportunity period for MAST (Heringstad *et al.*, 2001; Paper III). Moreover, the genetic correlation between MAST and NMAST has been reported to be as high as 0.96, although a longer opportunity period was considered for NMAST (Andersen-Ranberg and Heringstad, 2006).

If it turns out that multiple cases should be included in a CSM approach, an ordinal TM will be preferable over an LM because the latter assumes equal steps between levels, i.e. that it is as bad to go from 0 to 1 as it is to go from 1 to 2 (and so on), and this assumption is questionable. Heringstad *et al.* (2006) analyzed NMAST with linear CSM and ordinal threshold CSM where censoring was either ignored or taken into account. Censoring affected both parameter estimates and PBVs, and although no formal model comparison was conducted, it was concluded that the model including censoring was the better one, since it used more of the available information in a proper way. Rodrigues-Motta *et al.* (2007) suggested an alternative modeling of NMAST using a zero-inflated Poisson model, which is suitable for count data where there is an excess of zeros relative to what is expected from a Poisson distribution.

Between and within lactations (Papers I and IV)

A breeding program should seek to improve mastitis resistance in more than just the first lactation, especially in view of the increasing udder health problems in later parities. CM as well as SCC are somewhat different traits, genetically, between the first three parities (Paper I, Heringstad *et al.*, 2005; Negussie *et al.*, 2006), and inclusion of additional information from later parities should preferably be included if available at time of GE, to increase the accuracy of selection. A multiple-trait model treating CM in the first and later lactations as different traits would be appropriate, given that the genetic correlations reported between second and third parity (>0.9) are higher than those reported between first and either second or third parity (Paper I, Negussie *et al.*, 2006). In the study by Heringstad *et al.* (2005) the estimated genetic correlations between the first two parities, and between second and third parity, were, however, similar. Current GE both in NAV and in Norway includes the first three lactations as three separate traits in a multiple-trait setting.

The genes involved in CM resistance mechanisms seem to differ even more between early (up to lactation day 30) and late lactation stages within a single lactation than they do between the same lactation stage in different lactations (Heringstad *et al.*, 2004). Other longitudinal studies have also revealed considerable differences in heritabilities, as well as genetic correlations, between selected days within the first lactation (e.g. Chang, 2002; Heringstad *et al.*, 2003b; Paper IV). Our expectation in Paper IV was that the genetic correlation between CM on different days would be reduced the further apart the days were, and that the correlation pattern might reveal that CM in early and late lactation was, in fact, genetically distinct traits; this had been concluded in previous studies (Lund *et al.*, 1999; Heringstad *et al.*, 2004). According to our results from both linear RRM and linear LMVM, however, CM up to day 30 in first lactation was strongly genetically correlated with CM between day 140 and 200, whereas CM in the time period between seemed to be a more separate trait genetically. (This is illustrated in Figure 5.) A similar pattern for early CM has been found in some other studies using either threshold RRM or threshold LMVM, although the findings are not as distinct (Chang, 2002; Heringstad *et al.*, 2003b; Chang *et al.*, 2004a; Chang *et al.*, 2004b).

Genetic correlations between traits can arise either if partly the same genes or different but genetically linked genes affect the different traits. One biological explanation of the pattern of less-than-unity genetic correlations

for CM over time would be that pathogen-specific differences in the course of the lactation occur (Hogan *et al.*, 1989) and different pathogens might initiate different genetically determined defense mechanisms. If this explains the pattern in Paper IV, the same or similar pathogens would have to be important in the beginning and at the end of lactation. Another possible explanation would introduce resource allocation theory. This is imaginable if the resources available for immune defense are reduced because an increased proportion of them (the resources) are allocated to milk production (early in lactation) or to the fetus development (later in lactation). Different genes involved in disease resistance might be expressed depending on whether or not the cow is experiencing stress.

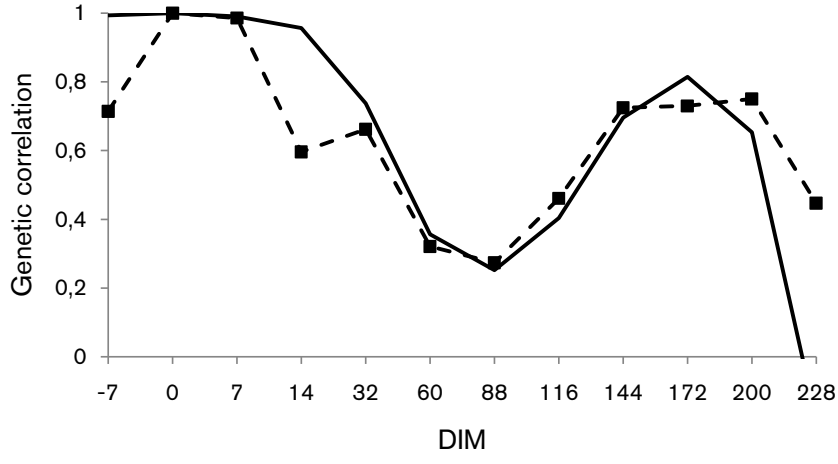

Figure 5. Genetic correlations for clinical mastitis between the day of first calving and the remaining part of the lactation from a linear random regression model (solid line) and the corresponding point estimates (■) from separate intervals from a linear longitudinal multivariate model (Paper IV).

The usefulness of longitudinal models

If it is practically feasible, longitudinal models should be favored over CSM and SA. The reason for this is that a CSM (at least, where only one interval is being considered) analyzing MAST or NMAST summarizes possible repeated CM cases into a single response class as well as assuming a constant heritability, unity genetic correlation and a constant genetic value of sires over time, which can be biologically questioned. A constant genetic value is also assumed in SA, although here it might be possible to model a time-dependent random effect of sire. The routine GE for CM has been

improved over the last few years, so that several lactations and at least two time periods are considered in the first lactation in NAV and in the first three lactations in Norway. For example, in the first lactation, CM before and after day 50 are treated as two separate traits in NAV, whereas CM before day 30, between day 31 and 120, and after day 121 are treated as three separate traits in Norway (Interbull, 2008). The definition used in Norway is attractive if the results in Paper IV reveal the true genetic picture of CM. The development in routine GE for CM towards a more longitudinal approach seems to be a step in the right direction. Important aims in the further development of these models might include investigation of the most suitable number of intervals as well as cut points between intervals. There might be a trade-off between the number of intervals that are biologically desirable and the number that are practically feasible. Whether these intervals are best analyzed with LMVM or RRM is another issue that will require investigation.

One possibility with longitudinal models is to develop new criteria for the ranking and selection of sires. In an RRM, these are based on the sire regression coefficients and could, for example, be the expected fraction of days without CM (Heringstad *et al.*, 2003b) or the summarized PBVs calculated as area under the curve (Paper IV). The latter was used to get a single measure for each sire from RRM to be used for comparison with the CSM; but summarizing over a longer time period is probably not the best selection measure from RRM, because CM changes over time and a summary measure reduces the amount of information used. It has been reported that cases of CM early in lactation are more costly than later cases because they involve a suppressed milk yield throughout lactation (Hagnestam *et al.*, 2007). Therefore, it might be wise, economically speaking, to place more emphasis on the first part of lactation by selecting sires which are good, and thus give daughters with a better mastitis resistance than average, in the beginning of the lactation. Regardless of the method used for GE, there is a risk associated with premature selection of sires based mainly on ongoing records; the risk arises because of the changing sire PBV over time. A drastic illustration of this can be seen, in connection with the sires with the most positive and negative slope (b1) in Paper IV, in Figure 6. There was, however, a strong within-method correlation (>0.9) between the periods 10 days before to either 50 or 150 days after calving for both the summarized PBVs from linear RRM and PBVs from linear CSM.

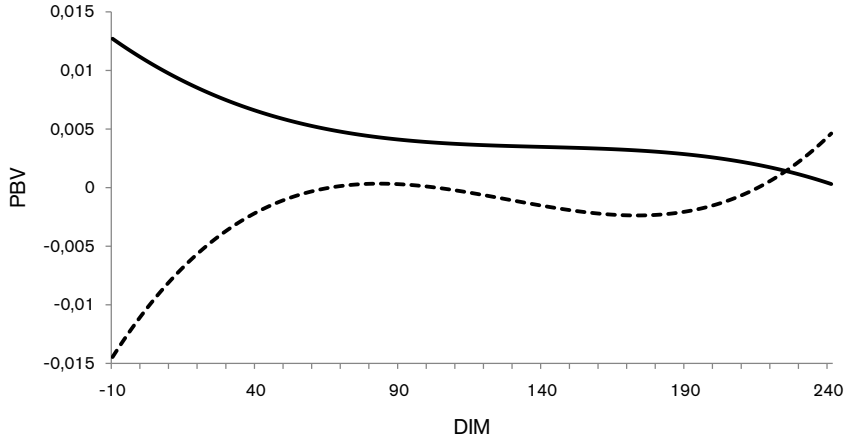

Figure 6. Predicted breeding values (PBV) for the two sires with the most positive (dashed) and negative (solid line) slope up to day 241 in first lactation. Lower PBV implies less CM.

Longitudinal models have theoretical advantages in GE of CM, but the application of these models in a practical setting would be associated with more computational difficulties than, at least, linear CSM. Comparing LMVM and RRM, the latter is somewhat less flexible in capturing different gene expressions in different parts of the lactation, but it also has computational advantages, since in it the number of parameters to estimate is reduced (Jensen, 2001). A threshold RRM for CM solved with Bayesian methodology would be more complicated to apply in routine GE with large data sets and a multiple-trait setting than a linear RRM. A multiple-trait model for GE of mastitis resistance including the correlated traits CM and SCC has been shown to increase the accuracy of selection, especially for the trait with the lowest heritability, i.e. CM (Negussie *et al.*, 2006). RRM for test-day SCC is already in routine use in some countries (Interbull, 2008) and is under development elsewhere. It would be advantageous if in future test-day SCC and longitudinal CM could be analyzed on the same time scale in a bivariate linear RRM. The use of a linear RRM for test-day SCC or milk yield and CM defined as two separate binary traits (before and after day 30 of lactation), for which only an intercept term was fitted, has been reported by Negussie *et al.* (2008). In that study, a preliminary analysis using monthly intervals for longitudinal CM in a bivariate RRM failed, probably as a result of the very low CM frequencies.

Conclusions

The main conclusions to be drawn from the results in this thesis about genetic evaluation of clinical mastitis in dairy cattle are as follows:

- There is an unfavorable genetic correlation between udder health and production, and this emphasizes the need to include mastitis resistance in the breeding goal
- Heritability for CM is low, but considerable genetic variation exists, and this makes genetic selection for mastitis resistance possible
- Accuracy of selection can be increased by combining information on CM and SCC, which are rather strongly genetically correlated
- CM should not be considered as the same trait genetically over time in genetic evaluations, because of the less-than-unity genetic correlations within and between lactations
- In the genetic evaluation for CM, little or no gain in accuracy seems to be achieved by replacing the linear cross-sectional model with a threshold cross-sectional model or survival analysis
- A longer opportunity period for CM will result in increased accuracy of selection, and thus in genetic progress, unless it is counteracted by a longer generation interval
- Longitudinal rather than cross-sectional models are biologically recommendable for genetic evaluation of CM and should be used if it is practically feasible to do so
- The use of a linear random regression model for CM has several appealing features in theory but this method seems sensitive and unstable for parameter estimation
- Choosing the best model for genetic evaluation of CM is far from straightforward, because the suitability of a model is affected by trait definition as well as by theoretical and practical considerations

Future research

In general, further investigations into the most suitable way to model longitudinal CM data, and how to best utilize the results obtained from such models, should be prioritized. The underlying factors causing the pattern of genetic correlations for CM over time also represent an area of major biological interest.

More research into the possibilities and the limitations of LMVM and RRM is needed. It is necessary to discover which one of these approaches is most applicable. This would involve work on deciding the appropriate number and lengths of intervals, as well as on cut points between intervals. Cut points should preferably be based on the pattern of genetic correlation of CM over time, whereas the number of intervals might be more of a practical issue as it affects the number of observations it is appropriate to analyze and, in an LMVM, the number of parameters to estimate. Clearer evidence of the gain achieved by replacing a CSM with a longitudinal model is also desirable, although comparisons are complicated when traits are defined differently. Cross-validation to assess the predictive ability would nicely complement the measures already discussed for field data; alternatively a simulation study could be performed. There is in general a real need for better approaches with which to compare methods of GE, and the development of such approaches should therefore be encouraged.

The use of linear RRM for CM is appealing, but it was far from straightforward for parameter estimation in our studies. To determine whether it was the low CM frequency that gave rise to the problems, it would be interesting to study material with a higher frequency of CM. This could be done, for example, in a simulation where the CM incidence could be varied. Another option would be to further examine threshold RRM, which is less dependent on the frequency than the linear counterpart. Another area to look into further is how to best model the random

regression function and to compare, for example, polynomials and linear splines. It would probably be beneficial, for GE of mastitis resistance, to analyze a longitudinal CM trait together with test-day SCC in a bivariate RRM, and an important research question in this context is whether and if so how, the genetic association between these traits changes over time.

How to best utilize the additional information resulting from longitudinal models for CM needs to be examined. With several time-specific breeding values for each sire we can create new criteria for ranking and selection. Simply summarizing the area under the curve in an RRM in order to obtain a single measure does not exploit the full benefit of this model. The gains and risks associated with a selection criterion which attaches more weight to CM in early lactation would be worth studying, given the higher costs associated with early CM cases.

Our evaluation of SA for CM was somewhat limited, because the models were to be similar to the commonly used CSM and therefore the full power of SA was not exploited. Future research into the possibility of including, for instance, time-dependent random genetic effects of sire in these types of model would therefore be sensible.

Possible biological explanations of the less-than-unity genetic correlation of CM over time present an inspiring area of research. With good quality data on pathogen-specific CM, inferences might presumably be drawn about the presence of different pathogens over time, and about the ways in which these might relate to different genetically determined defense mechanisms. To look more deeply into resource allocation theory for clarifications of the phenomenon would be even more challenging and would probably have to involve data on energy balance. Cows of the Swedish Red breed have better udder health than Swedish Holstein cows, and this appears to be the result, in part, of differences in immune functions. The background to these breed-specific differences in mastitis resistance needs to be further characterized; it is also necessary to ask whether the differences affect the genetic correlations over time. A complementary genetic study of longitudinal CM data in the Swedish Red breed designed to obtain genetic parameters both within and between lactations for this breed would be a suitable starting point.

Avelsvärdering av klinisk mastit hos mjölkkor

Inledning

Traditionellt har avelsarbetet för mjölkkor fokuserat på att öka produktiviteten per ko vilket har bidragit till en dramatisk ökning i mjölkavkastning både i Sverige och i många andra länder där liknande avelsarbete bedrivits. En ensidig avel för ökad produktion medför dock oönskade bieffekter för funktionella egenskaper som fruktsamhet och hälsa. Sjukdomen mastit eller juverinflammation är ett av de största problemen inom mjölkkosektorn då den är mycket vanligt förekommande och orsakar stora ekonomiska förluster för lantbrukaren, lidande för de drabbade korna samt bidrar till en ökad antibiotikaanvändning vilket medför risk för resistensutveckling och är etiskt ohållbart. Kostnader i samband med mastit kommer huvudsakligen från en sänkt mjölkavkastning, fler veterinärbehandlingar, lägre mjölkpris till följd av en försämrad kvalitet och en förhöjd risk för utslagning och nya sjukdomsutbrott. Mastit orsakas vanligen av patogena bakterier som kommer in i juvret via spenkanalen och, om de tillåts föröka sig, ger upphov till inflammation. Det är en komplex sjukdom på grund av att det finns många olika patogener, att mastit uttrycks på många olika sätt och att den har en multifaktoriell bakgrund med många gener och miljöfaktorer inblandade. Mastit kan delas upp i en klinisk och en subklinisk form där båda kännetecknas av en minskad mjölkproduktion och en förändrad mjölksammansättning med en kraftig förhöjning av celltalet (koncentrationen av vita blodkroppar), men där synliga symptom, som ett ömt och svullet juver samt missfärgningar och flockor i mjölken, endast förekommer vid klinisk mastit.

Det ogynnsamma genetiska sambandet mellan mastit och mjölkproduktion innebär att det är viktigt att inkludera mastitresistens i avelsmålet för att förhindra att situationen förvärras. Sverige och de övriga nordiska länderna har en lång tradition av att bedriva ett avelsarbete där både produktion, hälsa och fruktsamhet ingår i avelsmålet. Detta arbete är möjligt tack vare nationella kodatabaser som innehåller uppgifter om djurens identitet, släktskap och registreringar av en mängd viktiga egenskaper. För att bestämma om och hur information om dessa egenskaper ska användas är det viktigt att känna till arvbarheter och genetiska korrelationer (samband). För mastitresistens kan avelsvärdering göras antingen baserat på kliniska mastiter, celltalet eller andra indirekta mått på sjukdomen. I avelsvärderingen får man ett mått på varje djurs nedärvningsförmåga för den aktuella egenskapen som sen kan användas för att rangordna och selektera individer för insättning i avel. Vilket avelsframsteg som kan uppnås för en viss egenskap påverkas av ett antal olika faktorer men en hög säkerhet på de skattade avelsvärdena är en viktig hörnsten för att kunna selektera rätt individer och på så vis bidra till det genetiska framsteget.

Syftet med avhandlingen var att öka kunskapen om den genetiska bakgrunden till mastitresistens och framförallt att förbättra avelsvärderingen för klinisk mastit genom att utnyttja mer av den tillgängliga informationen och använda bästa möjliga metodik.

Sammanfattning av avhandlingens delarbeten

Nya skattningar av arvbarheter för klinisk mastit och celltal samt genetiska korrelationer mellan dessa juverhälsoegenskaper och produktionsegenskaper i de första tre laktationerna hos Svensk Holstein togs fram i delarbete 1. Metoden som användes var en linjär modell där egenskaperna definierades enligt den då gällande nationella avelsvärderingen. För klinisk mastit innebär det att man skiljer på kor med minst ett mastitfall (1) och kor utan fall (0) inom en viss tidsperiod av laktationen. Ogynnsamma genetiska korrelationer mellan juverhälsa och produktion bekräftade vikten av att inkludera mastitresistens i avelsmålet. Arvbarheten för celltal var något högre (10-14 %) än för klinisk mastit (1-3 %) och den genetiska korrelationen mellan dessa två mått på juverhälsa var hög (+0,7) vilket innebär att celltalet kan bidra till att öka säkerheten i avelsvärderingen för mastitresistens. Säkerheten på skattade avelsvärden är dock högst om information från båda egenskaperna kombineras. Skattningar av genetiska korrelationer mellan laktationer för både klinisk mastit och celltal visade att delvis olika gener har betydelse i olika laktationer och därför bör information från flera laktationer

än den första ingå i avelsvärderingen. Baserat på resultaten från denna studie kan en fleregenskapsmodell med första och senare laktationer som olika egenskaper rekommenderas.

Den traditionella linjära modellen har vissa nackdelar för avelsvärdering av klinisk mastit, som att den befintliga informationen inte utnyttjas fullt ut när upprepade fall och tidsaspekter ignoreras samt att pågående och ofullständiga observationer inte hanteras på ett optimalt sätt. En linjär modell förutsätter dessutom normalfördelade data vilket inte stämmer för klinisk mastit. I delarbetena 2 och 3 undersöktes därför med hjälp av både fältdata och en simuleringsstudie om den teoretiskt bättre metoden överlevandeanalys, där tiden till första kliniska mastitfall inom laktationen analyserades, skulle ge en säkrare avelsvärdering för klinisk mastit än den linjära modellen. I simuleringsstudien ingick även en tröskelmodell i jämförelsen. Resultaten visade dock att det inte var någon skillnad i säkerhet mellan metoderna när de jämfördes med samma längd på riskperioden, antingen de första 150 dagarna eller hela första laktationen. För den längre riskperioden var det emellertid en 8 % högre korrelation mellan tjurarnas simulerade sanna avelsvärden och skattade avelsvärden än för den kortare tidsperioden, men det var inte någon skillnad mellan metoderna.

En longitudinell modell kan hantera de flesta av nackdelarna som finns med en traditionell linjär modell för klinisk mastit och ger dessutom möjlighet att ta hänsyn till att genetiska och miljömässiga effekter kan variera över tiden. Modellen är dock mer komplex och kräver ökad datorkapacitet och kan därför vara svårare att införa i praktisk avelsvärdering. I delarbete 4 undersöktes möjligheten att använda en linjär slumpmässig regressionsmodell för avelsvärdering av klinisk mastit definierat som 12 upprepade binära observationer av samma fenotypiska egenskap inom första laktationen. Den här metoden visade sig dock vara instabil och känslig för parameterskattning under de givna förutsättningarna med detta datamaterial. Trots detta så stämde de skattade kurvorna av arvbarheten och de genetiska korrelationerna över tiden från den slutgiltiga regressionsmodellen med motsvarande punktskattningar från en linjär longitudinell fleregenskapsmodell i vilken klinisk mastit hanterades som 12 olika egenskaper. Detta användes som en informell validering av att den valda regressionsmodellen fungerade tillfredsställande. Båda longitudinella modellerna indikerade att klinisk mastit inte bör betraktas som en och samma genetiska egenskap i avelsvärderingen, baserat på den stora variationen i de genetiska korrelationerna mellan utvalda tidsperioder i första laktationen. En longitudinell modell är därför att föredra framför den traditionella linjära modellen, om den förstnämnda går att tillämpa i

praktiken. I framtiden skulle det vara önskvärt att analysera testdagsdata av celltal och longitudinella data av klinisk mastit samtidigt eftersom detta förväntas öka säkerheten på skattade avelsvärden. Detta talar för användandet av en linjär slumpmässig regressionsmodell men med bakgrund av svårigheterna som uppstod i delarbete 4 så krävs mer studier inom detta område för att klarlägga om det är rätt väg att gå.

Slutsatser

Avelsvärdering för mastitresistens är nödvändig på grund av det ogynnsamma genetiska sambandet med mjölkproduktion. En ökad säkerhet på skattade avelsvärden fås om information från klinisk mastit och celltal kombineras. För att ytterligare förbättra säkerheten vid avelsvärdering för klinisk mastit bör flera laktationer ingå om denna information finns tillgänglig vid tiden för avelsvärderingen. Att ersätta den traditionella linjära modellen med en tröskelmodell eller överlevandeanalys förväntas inte öka säkerheten. En längre riskperiod hade större betydelse än val av metod för avelsvärderingen och kan medföra en ökad säkerhet vilket är gynnsamt för det genetiska framsteget, om förbättringen inte motverkas av ett förlängt generationsintervall. En longitudinell modell är att rekommendera ur ett biologiskt perspektiv, eftersom mer av den tillgängliga informationen utnyttjas och hänsyn tas till att klinisk mastit inte är samma genetiska egenskap över hela laktationen, och bör därför användas om det är praktiskt möjligt. En linjär slumpmässig regressionsmodell visade sig dock vara känslig för parameterskattning under våra förhållanden. Att jämföra metoder på ett rättvisande sätt är komplicerat då det påverkas av hur egenskapen definieras samt teoretiska och praktiska aspekter.

Kortfattat om framtida forskning

Ytterligare studier behövs för att:

- Undersöka hur man bäst analyserar och utnyttjar resultaten av longitudinella data av klinisk mastit genom att studera och jämföra olika modeller och definitioner av klinisk mastit, t.ex. antal intervall och brytpunkter, olika selektionskriteria samt undersöka vilken betydelse frekvensen av klinisk mastit har för resultatet
- Ta reda på den biologiska förklaringen till varför gener har olika stor betydelse för mastitresistensen i olika delar av laktationen genom att studera data med information om patogenspecifik klinisk mastit alternativt söka förklaringar inom resursallokeringsteorin

References

Akers, R.M. 2002. *Lactation and the mammary gland*. Iowa: Iowa State Press.

Andersen-Ranberg, I. M. and Heringstad, B. 2006. Genetic associations between female fertility, mastitis and protein yield in Norwegian Red. In: *Proceedings of the 8th World Congress on Genetics Applied to Livestock Production*. Belo Horizonte, Brazil, August 2006. CD-ROM Commun. no. 1-20.

Barkema, H.W., Schukken, Y.H., Lam, T.J.G.M, Beiboer, M.L., Wilmink, H., Benedictus, G. and Brand, A. 1998. Incidence of clinical mastitis in dairy herds grouped in three categories by bulk milk somatic cell counts. *Journal of Dairy Science* 81, 411-419.

Boettcher, P.J., Dekkers, J.C.M. and Kolstad, B.W. 1998. Development of an udder health index for sire selection based on somatic cell score, udder conformation, and milking speed. *Journal of Dairy Science* 81, 1157-1168.

Boettcher, P.J., Jairath, L.K. and Dekkers, J.C.M. 1999. Comparison of methods for genetic evaluation of sires for survival of their daughters in the first three lactations. *Journal of Dairy Science* 82, 1034-11044.

Chang, Y.M. 2002. *Multivariate and longitudinal models for binary data with applications to clinical mastitis in Norwegian cattle*. Doctoral thesis. University of Wisconsin, US.

Chang, Y.M., Gianola, D., Heringstad, B. and Klemetsdal, G. 2004a. Effects of trait definition on genetic parameter estimates and sire evaluation for clinical mastitis with threshold models. *Animal Science* 79, 355-363.

Chang, Y.M., Gianola, D., Heringstad, B. and Klemetsdal, G. 2004b. Longitudinal analysis of clinical mastitis at different stages of lactation in Norwegian cattle. *Livestock Production Science* 88, 251-261.

Colleau, J.J. and Le Bihan-Duval, E. 1995. A simulation study of selection methods to improve mastitis resistance in dairy cows. *Journal of Dairy Science* 78, 659-671.

de Haas, Y. 1998. *Genetic parameters of mastitis traits and protein production in the first two lactations of Swedish Red and White cows*. Report no.17. Swedish University of Agricultural Sciences. Uppsala, Sweden.

de Haas, Y. 2003. *Somatic cell count patterns – Improvement of udder health by genetics and management*. Doctoral thesis. Wageningen University/Animal Sciences Group Lelystad. The Netherlands.

de Jong, G. and Lansbergen, L. 1996. Udder health index: selection for mastitis resistance. In: Proc. International workshop on genetic improvement of functional traits in cattle. Gembloux, Belgium, January 1996, *Interbull Bulletin* 12, 42-47.

Dohoo, I.R. and Leslie, K.E. 1991. Evaluation of changes in somatic cell counts as indicators of new intramammary infections. *Preventive Veterinary Medicine* 10, 225-237.

Ducrocq, V. 1987. *An analysis of length of productive life in dairy cattle*. Doctoral thesis. Cornell University, Ithaca, NY.

Ducrocq, V. and Sölkner, J. 1998. *The Survival Kit – A Fortran package for the analysis of survival data*. Proc. 6th World Congr. Genet. Appl. Livest. Prod., Armidale, Australia, 27, 447-448.

Emanuelson, U. 1997. Clinical mastitis in the population: Epidemiology and genetics. 48th Annual Meeting of the EAAP. Vienna, Austria, Aug 25-28 1997.

Emanuelson, U., Danell, B. and Philipsson, J. 1988. Genetic parameters for clinical mastitis, somatic cell counts, and milk production estimated by multiple-trait restricted maximum likelihood. *Journal of Dairy Science* 71, 467-476.

Emanuelson, U., Oltenacu, P.A. and Gröhn, Y.T. 1993. Nonlinear mixed model analyses of five production disorders of dairy cattle. *Journal of Dairy Science* 76, 2765-2772.

Foulley, J. L. 1992. Prediction of selection response for threshold dichotomous traits. *Genetics* 132, 1187-1194.

Gianola, D. 1982. Theory and analysis of threshold characters. *Journal of Animal Science* 54, 1079-1096.

Gianola, D. and Foulley, J.L. 1983. Sire evaluation for ordered categorical data with a threshold model. *Genetics Selection Evolution* 15, 201-224.

Gilmour, A. R., Gogel B.J., Cullis, B.R., Welham, S.J. and Thompson, R. 2002. *ASReml User Guide Release 1.0*. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Hagnestam, C., Emanuelson, U. and Berglund, B. 2007. Yield losses associated with clinical mastitis occurring in different weeks of lactation. *Journal of Dairy Science* 90, 2260-2270.

Hagnestam-Nielsen, C. and Østergaard, S. 2008. Economic impact of clinical mastitis in a dairy herd assessed by stochastic simulation using different methods to model yield loss. *Animal*. Accepted 2008.

Hansen, M., Lund, M.S., Sørensen, M.K. and Christensen, L.G. 2002. Genetic parameters of dairy character, protein yield, clinical mastitis, and other diseases in the Danish Holstein cattle. *Journal of Dairy Science* 85, 445-452.

Harmon, R.J. 1994. Physiology of mastitis and factors affecting somatic cell counts. *Journal of Dairy Science* 77, 2103-2112.

Heringstad, B., Klemetsdal, G. and Ruane, J. 2000. Selection for mastitis resistance in dairy cattle: a review with focus on the situation in the Nordic countries. *Livestock Production Science* 64, 95-106.

Heringstad, B., Klemetsdal, G. and Ruane, J. 2001. Variance components of clinical mastitis in dairy cattle – effects of trait definition and culling. *Livestock Production Science* 67, 265-272.

Heringstad, B., Rekaya, R., Gianola, D., Klemetsdal, G. and Weigel, K. A. 2003a. Genetic change for clinical mastitis in Norwegian cattle: A threshold model analysis. *Journal of Dairy Science* 86, 369-375.

Heringstad, B., Chang, Y.M., Gianola, D. and Klemetsdal, G. 2003b. Genetic analysis of longitudinal trajectory of clinical mastitis in first-lactation Norwegian cattle. *Journal of Dairy Science* 86, 2676-2683.

Heringstad, B., Chang, Y.M., Gianola, D. and Klemetsdal, G. 2004. Multivariate threshold model analysis of clinical mastitis in multiparous Norwegian dairy cattle. *Journal of Dairy Science* 87, 3038-3046.

Heringstad, B., Chang, Y.M., Gianola, D. and Klemetsdal, G. 2005. Genetic analysis of clinical mastitis, milk fever, ketosis, and retained placenta in three lactations of Norwegian Red cows. *Journal of Dairy Science* 88, 3273-3281.

Heringstad, B., Chang, Y.M., Andersen-Ranberg, I.M. and Gianola, D. 2006. Genetic analysis of number of mastitis cases and number of services to conception using a censored threshold model. *Journal of Dairy Science* 89, 4042-4048.

Hinrichs, D., Stamer, E., Junge, W. and Kalm, E. 2005. Genetic analyses of mastitis data using animal threshold models and genetic correlation with production traits. *Journal of Dairy Science* 88, 2260-2268.

Hogan, J. S., Smith, K. L., Hoblet, K. H., Schoenberger, P. S., Todhunter, D. S., Hueston, W. D., Pritchard, D. E., Bowman, G. L., Heider, L. E., Brockett, B. L. and Conrad, H. R. 1989. Field survey of clinical mastitis in low somatic cell count herds. *Journal of Dairy Science* 72, 1547-1556.

Holmberg, M. 2007. *Genetic dissection of functional traits in dairy cattle*. Doctoral thesis. Swedish University of Agricultural Sciences. Uppsala, Sweden. Electronic version available at http://epsilon.slu.se/eng

Interbull. 2008. Description of national genetic evaluation systems for dairy cattle traits as applied in different Interbull member countries. Retrieved September 4, 2008 from http://www-interbull.slu.se/national_ges_info2/framesida-ges.htm

Jensen, J. 2001. Genetic evaluation of dairy cattle using test-day models. *Journal of Dairy Science* 84, 2803-2812.

Johansson, K., Eriksson, S., Pösö, J., Toivonen, M., Sander Nielsen, U., Eriksson, J-Å. and Pedersen Aamand, G. 2006. Genetic evaluation of udder health traits for Denmark, Finland, and Sweden. In: Proc. of the 2006 Interbull meeting. Kuopio, Finland, June 2006, *Interbull Bulletin* 35, 92-96.

Kadarmideen, H.N., Thompson, R. and Simm, G. 2000. Linear and threshold model genetic parameters for disease, fertility and milk production in dairy cattle. *Animal Science* 71, 411-419.

Kehrli, M.E. and Shuster, D.E. 1994. Factors affecting milk somatic cells and their role in health of the bovine mammary gland. *Journal of Dairy Science* 77, 619-627.

Koenen, E.P.C., Berglund, B., Philipsson, J. and Groen, A.F. 1994. Genetic parameters of fertility disorders and mastitis in the Swedish Friesian breed. *Acta Agriculturae Scandinavica, Section A - Animal Sciences* 44, 202-207.

Lund M. S., Jensen, J. and Petersen, P. H. 1999. Estimation of genetic and phenotypic parameters for clinical mastitis, somatic cell production deviance, and protein yield in dairy cattle using Gibbs sampling. *Journal of Dairy Science* 82, 1045-1051.

Luttinen, A. and Juga, J. 1997. Genetic correlations between milk yield, somatic cell count, mastitis, milkability and leakage in Finnish dairy cattle population. In: Proc. International

workshop on genetic improvement of functional traits in cattle; health. Uppsala, Sweden, June 1997, *Interbull Bulletin* 15, 78-83.

Madsen, P. and Jensen, J. 2008. *An user's guide to DMU*. A package for analysing multivariate mixed models. Version 6, release 4.7. University of Aarhus, Tjele, Denmark.

Mark, T. 2004. Applied genetic evaluations for production and functional traits in dairy cattle. *Journal of Dairy Science* 87, 2641-2652.

Meijering, A. and Gianola, D. 1985. Linear versus nonlinear methods of sire evaluation for categorical traits: a simulation study. *Genetics Selection Evolution* 17, 115-132.

Mrode, R.A. and Swanson, G.J.T. 1996. Genetic and statistical properties of somatic cell count and its suitability as an indirect means of reducing the incidence of mastitis in dairy cattle. *Animal Breeding Abstracts* 64, 847-857.

Nash, D.L., Rogers, G.W., Cooper, J.B., Hargrove, G.L., Keown, J.F. and Hansen, L.B. 2000. Heritability of clinical mastitis incidence and relationships with sire transmitting abilities for somatic cell score, udder type traits, productive life, and protein yield. *Journal of Dairy Science* 83, 2350-2360.

Negussie, E., Koivula, M. and Mäntysaari, E. A. 2006. Genetic parameters and single versus multi-trait evaluation of udder health traits. *Acta Agriculturae Scandinavica, Section A - Animal Sciences* 56, 73-82.

Negussie, E., Strandén, I. and Mäntysaari E. A. 2008. Genetic associations of clinical mastitis with test-day somatic cell count and milk yield during first lactation of Finnish Ayrshire. *Journal of Dairy Science* 91, 1189-1197.

Norberg, E. 2004. *Electrical conductivity of milk as a phenotypic and genetic indicator of bovine mastitis*. Doctoral thesis. The Royal Veterinary and Agricultural University, Fredriksberg/Danish Institute of Agricultural Sciences, Tjele. Denmark.

Nyman, A-K. 2007. *Epidemiological studies of risk factors for bovine mastitis*. Doctoral thesis. Swedish University of Agricultural Sciences. Uppsala, Sweden. Electronic version available at http://epsilon.slu.se/eng

Philipsson, J. and Lindhé, B. 2003. Experiences of including reproduction and health traits in Scandinavian dairy cattle programmes. *Livestock Production Science* 83, 99-112.

Philipsson, J., Ral, G. and Berglund, B. 1995. Somatic cell count as a selection criteria for mastitis resistance in dairy cattle. *Livestock Production Science* 41, 195-200.

Rauw, W.M., Kanis, E., Noordhuizen-Stassen, E.N. and Grommers, F.J. 1998. Undesirable side effects of selection for high production efficiency in farm animals: a review. *Livestock Production Science* 56, 15-33.

Rekaya, R., Gianola, D., Weigel, K. and Shook, G. 2003. Longitudinal random effects models for genetic analysis of binary data with application to mastitis in dairy cattle. *Genetics Selection Evolution* 35, 457-468.

Rodrigues-Motta, M., Gianola, D., Heringstad, B., Rosa, G.J.M. and Chang, Y.M. 2007. A zero-inflated Poisson model for genetic analysis of the number of mastitis cases in Norwegian Red cows. *Journal of Dairy Science* 90, 5306-5315.

Roxström, A. and Strandberg, E. 2002. Genetic analysis of functional, fertility-, mastitis-, and production-determined length of productive life in Swedish dairy cattle. *Livestock Production Science* 74, 125-135.

Rupp, R. and Boichard, D. 1999. Genetic parameters for clinical mastitis, somatic cell score, production, udder type traits, and milking ease in first lactation Holsteins. *Journal of Dairy Science* 82, 2198-2204.

Rupp, R. and Boichard, D. 2003. Genetics of resistance to mastitis in dairy cattle. *Veterinary Research* 34, 671-688.

Saebø, S. and Frigessi, A. 2004. A genetic and spatial Bayesian analysis of mastitis resistance. *Genetics Selection Evolution* 36, 527-542.

Saebø, S., Almøy, T., Heringstad, B., Klemetsdal, G. and Aastveit A.H. 2005. Genetic evaluation of mastitis resistance using a first-passage time model for Wiener processes for analysis of time to first treatment. *Journal of Dairy Science* 88, 834-841.

Sander Nielsen, U., Aamand Pedersen, G., Pedersen, J. and Jensen, J. 1997. Genetic correlations among health traits in different lactations. In: Proc. International workshop on genetic improvement of functional traits in cattle; health. Uppsala, Sweden, June 1997, *Interbull Bulletin* 15, 68-77.

SAS. 2002. SAS Release 9.1, 2002-2003. SAS Inst. Inc., Cary, NC, USA.

Schaeffer, L.R. 2004. Application of random regression models in animal breeding. *Livestock Production Science* 86, 35-45.

Schepers, A.J., Lam, T.J.G.M., Schukken, Y.H., Wilmink, J.B.M. and Hanekamp, W.J.A. 1997. Estimation of variance components for somatic cell counts to determine thresholds for uninfected quarters. *Journal of Dairy Science* 80, 1833-1840.

Schneider, M. del P. 2006. *Fertility, mastitis and longevity in dairy cattle analyzed using survival models*. Doctoral thesis. Swedish University of Agricultural Sciences. Uppsala, Sweden. Electronic version available at http://epsilon.slu.se/eng

Schneider, M. del P., Strandberg, E., Emanuelson, U., Grandinson, K. And Roth, A. 2007. The effect of veterinary-treated clinical mastitis and pregnancy status on culling in Swedish dairy cows. *Preventive veterinary Medicine* 80, 179-192.

Schukken, Y.H., Grommers, F.J, Van de Geer, D., Erb, H.N. and Brand, A. 1990. Risk factors for clinical mastitis in herds with a low bulk milk somatic cell count. 1. Data and risk factors for all cases. *Journal of Dairy Science* 73, 3463-3471.

Schukken, Y.H., Lam, T. J. G. M. and Barkema, H. W. 1997. Biological basis for selection on udder health traits. In: Proc. International workshop on genetic improvement of functional traits in cattle; health. Uppsala, Sweden, June 1997, *Interbull Bulletin* 15, 27-33.

Seegers, H., Fourichon, C. and Beaudeau, F. 2003. Production effects related to mastitis and mastitis economics in dairy cattle herds. *Veterinary Research* 34, 475-491.

Shook, G.E. 1989. Selection for disease resistance. *Journal of Dairy Science* 72, 1349-1362.

Strandberg, E. and Shook, G.E. 1989. Genetic and economic responses to breeding programs that consider mastitis. *Journal of Dairy Science* 72, 2136-2142.

Swedish Dairy Association. 2007a. *Djurhälsovård* (Animal health care) 2006/2007. Svensk Mjölk, Sweden.

Swedish Dairy Association. 2007b. *Husdjursstatistik* (Cattle statistics) 2007. Svensk Mjölk, Sweden.

Vaarst, M. and Enevoldsen, C. 1997. Patterns of clinical mastitis manifestations in Danish organic dairy herds. *Journal of Dairy Research* 64, 23-37.

Veerkamp, R.F. and de Haas, Y. 2005. Genetic improvement in mastitis control programmes. In: Hogeveen, H. (Ed.). *Mastitis in dairy production*. 115-122. Wagening Academic Publishers.

Weller, J.I. and Ron, M. 1992. Genetic analysis of fertility traits in Israeli Holsteins by linear and threshold models. *Journal of Dairy Science* 75, 2541-2548.

Weller, J.I., Saran, A. and Zeliger, Y. 1992. Genetic and environmental relationships among somatic cell count, bacterial infection, and clinical mastitis. *Journal of Dairy Science* 75, 2532-2540.

Yazdi, M.H., Visscher, P.M., Ducrocq, V. and Thompson, R. 2002. Heritability, reliability of genetic evaluations and response to selection in proportional hazard models. *Journal of Dairy Science* 85, 1563-1577.

Zwald, N.R., Weigel, K.A., Chang, Y.M., Welper, R.D. and Clay, J.S. 2004. Genetic selection for health traits using producer-recorded data. I. Incidence rates, heritability estimates, and sire breeding values. *Journal of Dairy Science* 87, 4287-4294.

Zwald N.R., Weigel K.A., Chang Y.M., Welper R.D. and Clay J.S. 2006. Genetic analysis of clinical mastitis data from on-farm management software using threshold models. *Journal of Dairy Science* 89, 330-336.

Acknowledgements

To produce a PhD thesis might seem like a ”one-man-show” to some people, but those of you that have been there know that behind this little book lies efforts from a lot of people. With this book I will finish a chapter of my life that has been filled with challenges to overcome, problems to solve, successes and failures. I could not have done this without you! Don't feel bad if this part of the thesis is the first, and maybe only, part you will read. As a friend of mine stated, this is probably because we humans are very interested in the social bounds of other people and this curiosity might in itself be a great character in our daily life, not at least within research!

The work behind this thesis was performed at the Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences. Financial support was given by the department and travel grants from “Knut och Alice Wallenberg stiftelse” and “SLU-medel för internationalisering av forskarutbildningen” enabled me to participate in an international PhD course in Belgium, 2005, and an international conference in Lithuania, 2008. All financial support is greatly acknowledged.

To my excellent supervisors Erling Strandberg and Ulf Emanuelson: Thank you for your never-ending support and enthusiasm! These characteristics of a supervisor are of great importance for the well-being of a PhD student! Thanks also for always believing in me and my work, even at times when I failed to do so myself. The results did not always turn out as expected (do they ever?) and your way of turning this into something interesting and challenging is admirable…I am grateful for the good co-operation we had during the last couple of years and I am looking forward to continuous team work!

To my main supervisor Erling: Thanks for catching my interest for animal breeding and genetics in the first place during the undergraduate course and my MSc project, and for giving me great conditions and possibilities to develop my academic skills during my PhD project by encouraging me to participate in courses, conferences and teaching. Apart from being my statistical expert and a very smart guy, you have been a very present supervisor that always took your time to answer questions, discuss results and give me many new ideas to try out….

To my co-supervisor Uffe: Thanks for your ability to combine professional and humane skills! Your feed-back on my work was always fast and very valuable, not at least during the last part of my project. Our discussions about the up's and down's in research, the sometimes difficult task in combining job and family, and future career possibilities made me always feel better. An encouraging ☺ in an e-mail can be of great importance for the receiver!

To my different co-authors and "resource persons":

Katja Grandinson, you have been a great mentor and friend for me during my whole PhD studies and I am glad that you wanted to act as a resource person for the last part of my project.

Pilar Schneider, I would like to thank you for your excellent help and guidance with my survival analyses, for introducing me to great people at different conferences and for giving my life some Latin American touch!

Anki Roth, you played an important role for me in the first and last chapter of my PhD studies and I want thank you for being a funny and inspiring MSc supervisor/co-author/carrier coach!

Freddy Fikse, thanks for your help and engagement in my PhD project and for the great ideas you gave me for my last paper, often during a coffee break or an informal corridor meeting!

Kjell Johansson, thanks for your interest in my project and for great help with data and ideas especially for my first paper!

To all my colleagues at the department: Even at times when work itself felt less fun (there were times…), it was always nice to "hang around" with you and I would like to acknowledge you all for this! You are a great mix of people that are all important for the success of the department and for individual persons. I would like to give a special thanks to the administrative staff, especially Jörgen, Bijan, Monica, Siw and Dan, who are always there to assist with practical things. Ni är klippor! I would also like to express my gratitude to all the past and present foreign guests. You enrich our working environment and make the department full of life!

To all the past and present PhD students at the department: It has been great being part of this tight group! We have shared so many nice work-related and social things together and I will miss being one of you! A special thanks to Mia, Jennie, Karl-Johan, Sofia and Christel for shared

moments and personal talks about life as a PhD student, human and parent and the combination of these roles! Thanks also to Emma TH and Linda for being supportive dissertation coaches!

Mia and Jennie, it is nice to have two such great friends as colleagues! I hope our families will stay in touch! Mia, I trust you completely, enjoy your company and would like to thank you for always being so cool. Jennie, you are a great listener (and speaker!) and your laughter is contagious!

To Lava, my dear dog, thanks for keeping my company at my office during all these years and for all nice lunch walks that kept me fit and filled my brain cells with new energy! Thanks also to all other dog owners and dogs at the department for your company and nice talks during our walks!

To all my dear friends, who make life worth living! You and your families, not at least all the new members, give an extra dimension to my life. I wish there was more time to spend together! Cissi, Kajsa, Nina and Emma, original members of "Kulturklubben", thanks for past and future social activities together! Christina, my newest friend, thanks for all the things we share☺! Ylva and Johanna, some of my nicest nature experiences in the mountains I shared with you and a pair of trekking skies or muddy hiking boots! My present and ex-neighbors next-door and "above", thanks for making life in a small cottage in Krusenberg even more pleasant. A special thanks to Sara & Björn for the great house-partys!

To my wonderful family:

Mum and dad, your support and believe in me is really appreciated! It feels so great to know that you are always there to listen when I need to talk (which is pretty often on dog walks!) and to help out with practical things such as caring for Tim and Lava!

My brothers, Fredrik and Erik, I am grateful for our good contact and friendship and for giving me the best "sisters" I could ever wish, Mia and Hanna, and a little lovely niece, Lina. Let us always be friends! A special thanks to Hanna and Erik for a cover illustration that makes me happy!

My Belgian family, Annie & Paul and Steven & Yamina, thanks for always opening your hearts and homes to us, and for the chocolate! It has been great getting to know you and your culture during our stays in Belgium!

Kristof, my life companion, thanks for your love, support and "coping ability", especially during the most intensive period of finishing my PhD thesis. I am grateful for you spontaneity (most of the time…), your endless believe in my abilities and for giving me the greatest gift of all!

Tim, min älskade son, tack för att du är en underbar liten människa som skänker mig glädje och ger mig perspektiv på tillvaron♥