

Genetic improvement of longevity in dairy cows

Mathijs van Pelt

Thesis committee

Promotor

Prof. Dr R.F. Veerkamp
Special professor Numerical Genetics and Genomics
Wageningen University & Research

Co-promotor

Prof. Dr T.H.E. Meuwissen
Professor of Bioinformatics
Norwegian University of Life Sciences, Ås, Norway

Other members

Prof. Dr H. Hogeveen, Wageningen University & Research
Prof. Dr N. Gengler, University of Liège, Gembloux, Belgium
Prof. M. Coffey, Scotland's Rural College, Edinburgh, UK
Dr S. Rensing, VIT, Verden, Germany

This research was conducted under the auspices of the Graduate School of Wageningen Institute of Animal Sciences (WIAS).

Genetic improvement of longevity in dairy cows

Mathijs van Pelt

Thesis

submitted in fulfillment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 24 March 2017
at 4 p.m. in the Aula.

Van Pelt, M. L.

Genetic improvement of longevity in dairy cows,
188 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2017)
With references, with summary in English

ISBN 978-94-6343-082-1

DOI: 10.18174/404192

Abstract

Van Pelt, M. L. (2017). Genetic improvement of longevity in dairy cows. PhD thesis, Wageningen University, the Netherlands

Improving longevity helps to increase the profit of the farmer, and it is seen as an important measure of improved animal welfare and sustainability of the sector. Breeding values for longevity have been published since in 1999 in the Netherlands. For AI-companies and farmers it is necessary that breeding values are accurately estimated and will remain stable for the rest of life. However, current breeding values for longevity of bulls seem to fluctuate more than expected. The main aim of this thesis was to revisit the genetics of longevity and develop a genetic evaluation model for longevity, where breeding values reflect the true breeding value quicker during early life and therefore breeding values become more stable. Genetic parameters were estimated for survival up to 72 months after first calving with a random regression model (RRM). Survival rates were higher in early life than later in life (99 vs. 95%). Survival was genetically not the same trait across the entire lifespan, because genetic correlations differ from unity between different time intervals, especially when intervals were further apart. Survival in the first year after first calving was investigated more in depth. Survival of heifers has improved considerably in the past 25 years, initially due to the focus on a high milk production. More recently, the importance of a high milk production for survival has been reduced. Therefore functional survival was defined as survival adjusted for within-herd production level. For survival the optimum age at first calving was around 24 months, whereas for functional survival calving before 24 months resulted in a higher survival. Over years, genetic correlations between survival in different 5-yr intervals were below unity, whereas for functional survival genetic correlations did not indicate that survival changed over years. This suggested that a genetic evaluation using historical data should analyze functional survival rather than survival. A new genetic evaluation system for longevity was developed based on a RRM analyzing functional survival. Based on the correlation between the first breeding value of a bull and his later breeding values, the ranking of bulls was shown to be more stable for RRM than the current genetic evaluation. Bias in breeding value was observed, mainly for bulls with a large proportion of living daughters. Adjusting for within-herd production level reduced this bias in the breeding values greatly. Before implementing this new model for genetic evaluation, the cause of this bias needs to be further investigated.

Contents

5	Abstract
9	1 – General introduction
27	2 – Genetic analysis of longevity in Dutch dairy cattle using random regression
59	3 – Changes in the genetic level and the effects of age at first calving and milk production on survival during the first lactation over the last 25 years
83	4 – Genetic changes of survival traits over the past 25 yr in Dutch dairy cattle
109	5 – Genetic evaluation of longevity using a random regression animal model
135	6 – General discussion
159	Summary
167	Curriculum vitae
171	Publications
177	Training and supervision plan
183	Acknowledgments

1

General introduction

1.1 Introduction

Longevity or length of productive life of a cow is a complex trait that reflects total lifespan of a cow. The actual longevity of a cow is determined by the farmer, because the farmer makes the decision to cull the cow or not, unless the cow dies, and the farmer takes into account milk production, health, fertility and workability traits when deciding to cull a cow. Longevity has been of economic importance and Rendel and Robertson (1950) already described how increasing longevity helps to increase profitability of the dairy farmer: 1) reducing annual costs of replacement of cows, 2) increasing the average milk production of the herd through an increase in the proportion of cows in the higher producing age groups, 3) reducing the number of replacements to be reared, and therefore allowing an increase in size of the milking herd for a given acreage and 4) increasing voluntary culling. Which of these factors is more prominent depends on the production circumstances of a farm and the culling decision of the farmer. For example, when a quota system is in place, extra production of the herd has little extra value, but the farmer could reduce the number of replacements, or cull more cows voluntary and herewith increase the average milk production per cow. In both cases, the profit will increase by lowering the costs for the same amount of milk. Van Arendonk (1985) showed that if involuntary culling is decreased, a higher voluntary culling rate can be applied, resulting in a larger profit for the farmer, but not necessarily increasing lifespan as a whole in the herd. Hence, the important economic aim for the farmer should be to reduce involuntary culling, rather than to increase herd life per se. Additional changes in production practices that affect animal health and welfare are needed, particularly given the societal concerns, such as animal welfare and antibiotic resistance, and will likely affect the profitability of the dairy herd (Barkema et al., 2015). Therefore, breeding for longevity is not only important as it will result in more profit, as well as increased animal welfare and social acceptance of the dairy sector.

1.2 Culling reasons

The longevity of a dairy cow is determined by the culling decision of the farmer, and farmers decide to cull dairy cows for various reasons such as mastitis, fertility problems, lameness, or low production (Beaudeau et al., 2000, Zijlstra et al., 2013). Longevity reflects a cow's ability not to be culled. Culling for low production is usually referred to as voluntary culling, culling for disease and low fertility as involuntary culling. Voluntary culling is an economically based decision where it is no longer worth breeding a cow compared with replacing her with a heifer,

1 General introduction

whereas with involuntary culling a cow exists the herd for biological reasons, for example for mastitis or infertility (Van Arendonk, 1985, Fetrow et al., 2006).

In the Netherlands, farmers can voluntarily record culling reasons, but not all farmers record these reasons and data are only available for the past decade. Main culling reasons in the Netherlands are for poor fertility, mastitis and feet and leg disorders (Zijlstra et al., 2013), which was similar to a study in Germany (Heise et al., 2016). Both studies also showed that culling reasons differed within and over lactations. However, recorded culling reasons are subjective scores of the farmer, and culling of a cow is often done for multiple reasons (Fetrow et al., 2006). For example, poor fertility or bad udder health combined with a low daily milk yield will result in a cow not being inseminated again. Therefore it is difficult to evaluate on the basis of voluntary recording of culling reasons whether longevity has changed.

The dairy industry has undergone profound changes in recent decades, that affected productivity, health and welfare of dairy cows, for example, herd size, use of hired labor, housing system, milk price and use of new technology (Barkema et al., 2015). At the same time, milk production per cow has more than doubled in the previous 40 years. Culling reasons are likely to be affected over years by these changes. Further factors that have changed the culling reasons over time are, for example, national regulation and legislation (e.g. quota system). In Europe, a quota system was in place until April 2015 to limit the amount of milk produced annually per country, but also per farmer (Bergevoet et al., 2004). Individual farmers that produced more than the allowed quota, had to pay a penalty for the excess of milk that they produced. This system may have affected culling reasons over the years. Also, the changing viewpoint of dairy farmers might have changed culling reasons. As demonstrated by the fact that worldwide, national selection indexes have changed drastically over the past 2 decades, where the breeding goal moved from selection for production only toward selection for production, longevity, and health traits (Miglior et al., 2005). Moreover, changes in productive life were analyzed over a 15-yr period in the United States, (Tsuruta et al., 2004), and survival of first lactation was analyzed over a 20-yr period in Australia (Haile-Mariam and Pryce, 2015). However, it was difficult to conclude if survival in the United States and Australia changed in those years, as genetic correlations between traits in different years from RRM and multi-trait models differed.

Taking all these factors together, it is very unlikely that longevity has been the same “trait” over the past decades, but there is little insight how culling reasons or longevity have changed over the past decades, and the most important factors that play a role in the culling decisions of dairy farmers.

1.3 Longevity

Without proper objective information on culling reasons, it is preferred to analyze longevity. However longevity can be defined in multiple ways also. Longevity can be measured as a lifetime trait or a stayability trait. Lifetime traits measure the whole lifespan of a cow or length of productive life. These can be measured only after the death of a cow. Stayability traits measure whether or not a cow is alive at a certain point in time (e.g., at a fixed number of months from birth or to a certain parity). These traits can be measured at any point in time, although they do not contain complete information on a cow's longevity as they are binary traits. Use of binary survival implies that only the animals having the opportunity to survive the entire specified period can be used in genetic analysis. Linear traits for longevity are number of lactations (Brotherstone et al., 1997), and length of productive life at 84 months (VanRaden and Klaaskate, 1993).

1.4 Statistical models

Longevity of dairy cows can be analyzed with several models. Survival analysis is a term generally accepted for as a set of methods for analyzing data in human and animal studies (Kalbfleisch and Prentice, 1980, Cox and Oakes, 1984). Survival analysis is characterized by data where the outcome variable is the time until the occurrence of an event of interest. The event can be death, occurrence of a disease, marriage, divorce, etc. The time to event or survival time can be measured in days, weeks, years, etc. For dairy cows the event of interest is culling and the survival time can be the time in days after first calving until a cow is culled. Typical advantages of survival analysis compared to more classical models are: 1) using time to event is more efficient use of information than just whether or not the event has occurred, 2) it may be inconvenient to wait until the event has occurred in all subjects (e.g. for dairy cows this may be 10 years until culling or more for some cows), and 3) possibility to include subjects whose time to event is not known yet (censored). Most common is right-censoring, meaning the event has not occurred at last follow-up. The most popular class of survival models is the class of proportional hazard models (**PHM**) (Cox, 1972, Kalbfleisch and Prentice, 1980). The hazard of an animal (or its risk of being culled) at time t is described as the product of a baseline hazard function and of a positive term which is an exponential function of a vector of covariates w' multiplied by a vector of regression parameters ϑ . The baseline hazard function is either left completely arbitrary (Cox model) or has a parametric form (e.g. exponential, Weibull or gamma).

1 General introduction

In animal breeding PHM is a popular model used in genetic evaluations for longevity, since the software package Survival Kit was introduced in 1994 (Ducrocq and Sölkner, 1994, 1998, Ducrocq et al., 2010). This model attempts to estimate the probability an animal will survive to time t given that it has survived to time $t-1$. The PHM can handle censored data and the non-normal distribution of survival data, but also time-dependent environmental effects can be included in the model. Since the PHM and other survival models have some practical disadvantages, in animal breeding several alternative approaches have been developed. For example to score survival as a binary trait; for example, survived or not up to a specific endpoint (i.e., time, age, lactation) or survived or not in a specific interval of time. This binary trait for survival can then be analyzed with a linear model or a threshold model. If survival is scored as a single binary trait, however, information from culling before the endpoint or information from survival after the endpoint is ignored. In addition, information from living animals that did not have the opportunity to survive a certain specified time interval cannot be used in the analysis. To overcome these major limitations, an expansion of censored records to expected lifespan has been suggested and implemented in the UK (Brotherstone et al., 1997), and US (VanRaden and Klaaskate, 1993), but also multi-trait models including survival to several endpoints have been suggested (Boettcher et al., 1999, Sewalem et al., 2007). The advantage of the multi-trait model is that information for survival after first calving of younger animals still alive can be used to score the traits of the first intervals (they already survived), and the traits for the later intervals can be set to missing; also, variation across the lifetime of a cow is included in the genetic evaluation. Ideally, many traits with different endpoints should be included in the analysis, but when the life of a cow is split in many binary traits, the multi-trait parameterization becomes impossible and random regressions models may be useful to find a more parsimonious parameterization (Schaeffer and Dekkers, 1994). For this reason, Veerkamp et al. (2001) proposed the random regression model (**RRM**) for survival analysis, which was further investigated by Jamrozik et al. (2008). Veerkamp et al. (2001) proved, by comparison of the likelihood, the equivalence between the PHM and the RRM when a large number of survival intervals were defined in the random regression model. Compared with the PHM, the advantage of the RRM is that, in a linear model context, multiple genetic effects can be fitted. Random regression models therefore enable the modeling of longevity as a different but genetically correlated trait across the entire productive life of a cow, whereas the applied PHM in genetic evaluations is restricted to an univariate analysis handling longevity as one genetic effect (Forabosco et al., 2009).

1.5 Genetic evaluation

Genetic evaluation methods evolved over the past 50 years. In 1949 Henderson published the method for best linear unbiased prediction (**BLUP**), but was not actually applied until 1970 because of limitations in computer hardware (Grosu et al., 2013). The application of BLUP has evolved since then from the sire model in 1970, to the sire and maternal grandsire model, the univariate animal model in 1989, to the multivariate animal model, and in 2000 the random regression model for the analysis of longitudinal data was implemented for the genetic evaluation using test day records for milk yield. More recently SNP-BLUP or GBLUP enabled the addition of genomic data in the genetic evaluations (Mrode, 2014). Next to the inclusion of DNA information, over the past 15 years advances in genetic evaluation were made due to automated collection of phenotypes for example for health and fertility traits. All these aspects, computing power, solving algorithms, and data, are connected and have led to the dramatic developments in genetic evaluation procedures, and currently more than a hundred breeding values traits are routinely available in the Netherlands. The focus here is on the genetic evaluation for longevity.

1.5.1 Genetic evaluation for longevity in practice

In animal breeding the application of longevity data is to predict the genetic merit for lifespan of animals and make selection decisions. However, this makes longevity a complex trait, because the true longevity of a cow is available at the end of her life, whereas the selection decisions are made earlier in life. For this purpose it is necessary that breeding values are accurately estimated and will remain stable for the rest of life. Worldwide, different approaches are used for genetic evaluation of longevity (Forabosco et al., 2009). In the Interbull genetic evaluation for longevity 19 countries participate of which 10 countries use survival analysis with a proportional hazards model, 8 countries use a single or a multiple-trait animal model, and one country uses a random regression lactation survival animal model. The current model used in the Netherlands to estimate breeding values is a piecewise Weibull mixed model for longevity (Van der Linde et al., 2004, Van der Linde et al., 2007). Differences in the approaches or models to estimate breeding values for longevity are described in Veerkamp et al. (2001): 1) handling of censored records, 2) use of time-dependent fixed effects, 3) accounting for non-normal distribution of longevity data, 4) combining information from other predictors, for example health and type traits, 5) number of distinct genetic effects assumed during lifetime, and 6) possibility of fitting an animal model to large data sets. No method available is best or worst for all six points.

A crucial aspect for the stability of breeding values appears to be whether the model can extrapolate survival data to predict the full lifespan, and the assumptions that are made in the underlying model. For example, whether longevity is the same trait in terms of (co)variances, or number of distinct genetic effects, during the total lifespan and across lactations. If longevity is not the same trait across time, then a model is needed that has the ability to handle survival data differently during that time span. There are some indications from studies using linear multi-trait models that genetic correlations between parities differ from unity, indicating longevity is not the same trait over the total lifespan (Visscher and Goddard, 1995, Brotherstone et al., 1997, Boettcher et al., 1999, Veerkamp et al., 2001, Sewalem et al., 2007). Similar results were found with a lactation random regression model (Gengler et al., 2005).

At the same time several studies with simulated and/or real data showed correlations of 0.90 or higher between breeding values of different models, indicating the used model may not be a major issue (Boettcher et al., 1999, Lubbers et al., 2000, Veerkamp et al., 2001, Jamrozik et al., 2008). However, Jamrozik et al. (2008) concluded that the random regression model showed a slight superiority over linear multiple-trait and PHM in predicting the proportion of sire's daughters that survived to different endpoints after the first calving. A random regression model can estimate changes over time if they are gradual (Tsuruta et al., 2004). Ducrocq (1999) proposed some suggestions to investigate survival analysis further, one of the suggestions is to use time-dependent sire effects, which is possible with random regression using orthogonal polynomials (Schaeffer, 2004), and in the newest release of the Survival Kit it is also possible to analyse two correlated random effects (Mészáros et al., 2013).

1.5.2 Use of predictor traits

Despite the different models to analyze censored records for survival, predictor traits may also be valuable to predict longevity early in life. These predictor traits, measured early in life, can improve the reliability of the breeding value due to the genetic relationship. Most countries in the Interbull evaluation for longevity combine breeding values of predictors and direct longevity to have a more reliable breeding value when only few cows are culled (Forabosco et al., 2009). Conformation traits, production traits, somatic cell count, calving traits, fertility traits and likability are used as predictors.

Reviews of Dekkers and Jairath (1994), Essl (1998) and Vollema (1998) covered a wide range of studies until the nineties and showed weak to moderate genetic correlations between type traits and longevity. In later studies similar results were

found (Vollema and Groen, 1997, Weigel et al., 1998, Buenger et al., 2001, Larroque and Ducrocq, 2001). Somatic cell count is another predictor with a moderate genetic correlation with longevity as well (Heringstad et al., 2003, Powell and VanRaden, 2003, Holtsmark et al., 2008). Most of these studies estimated genetic correlations between predictors and total lifespan and are indirectly assuming predictors have the same effect over the total lifespan. However, at phenotypic level culling risks or culling decisions for health disorders differ across lactation (Beaudeau et al., 2000, De Vries et al., 2010) and therefore it could be possible that genetic correlations between predictor traits and longevity are different across lifetime as well. If this is the case, the use of predictor traits in a multiple trait breeding value estimation with varying (co)variances may predict longevity better and estimate more stable breeding values. In contrast of using the correlation with predictor traits to estimate the breeding value for longevity, most countries (14 of 19) adjust longevity for cows for production level by including milk production as a covariate in the analysis and herewith provide an approximation of functional longevity (Forabosco et al., 2009). Functional longevity is perceived as the ability to delay involuntary culling, whereas true or productive longevity is the ability to delay culling.

1.5.3 Experience with genetic evaluation in the Netherlands

Breeding values for longevity have been published since the introduction in 1999 in the Netherlands, and have become very important due to the inclusion of the breeding values in the national index, and the introduction of genomic selection. At the same time the breeding value for longevity was included in the national index DPS in the Netherlands, with substantial weight of 33%. Since the introduction the weight has gradually decreased from 33% to 11% in the current national index (NVI), since more health and fertility breeding values became available and were included in the NVI. By weighing these traits directly in the NVI, less weight was required for longevity. Still, the longevity breeding value is perceived as an important breeding value for bulls.

Currently, most breeding companies use genomic information to select and sell bulls. For genomic selection genomic markers are calibrated using genotypes and breeding values of bulls with accurate breeding values (Meuwissen et al., 2001). Utilizing the genomic prediction equation, selection candidates, i.e., young bulls, are now selected at birth with greater accuracy than the parent average information that was available previously, and an accuracy very close to the accuracy of the breeding value based on progeny information (where bulls had to wait five years for). Genomic selection can be very beneficial for traits with a low

1 General introduction

heritability (health traits), traits expressed late in life (longevity), traits expensive to measure (feed efficiency), or traits that cannot be directly measured on the selection candidates (milk production). Genomic selection is very important for longevity, because bulls depend for a long time on the genomic information, as it takes several years before the daughters are culled. Although genomic selection may suggest that classical breeding value estimation became less important, this is not the case. For genomic prediction the breeding values for longevity are important, since these are used as input in the reference population to train the genomic prediction.

Hence, the importance of longevity in the national breeding goal, and the introduction of genomic selection made classical genetic valuation for longevity very important. However, the difficulties in estimating breeding values for longevity as described before, is also observed in practice. For example, current breeding values of bulls seem to fluctuate more than can be expected from theory, i.e., the change in breeding value compared to the increase in reliability (Figure 1.1). This perception was amplified by some famous bulls that were highly favorable for longevity initially, but dropped considerable in their breeding value when more information became available. Also, in most countries, like in the Netherlands, genomic prediction for longevity has poorest validation reliability (Gao et al., 2013; Stoop et al., 2015). Therefore, almost 20 years since the first genetic evaluation for longevity was developed in the Netherlands, an important aim of this thesis was to revisit the genetic aspects influencing the breeding value of longevity.

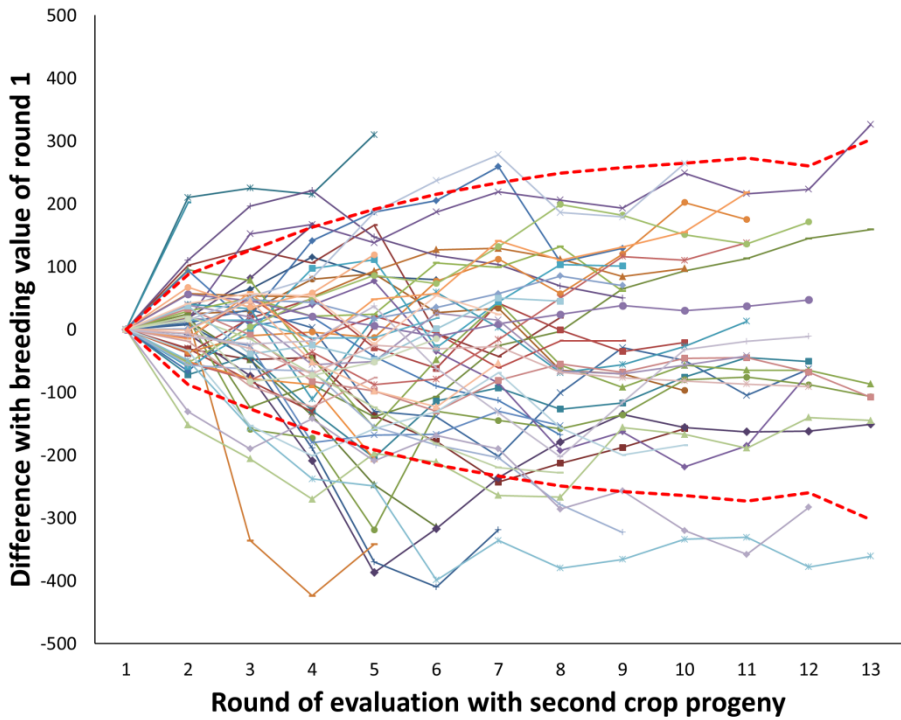


Figure 1.1. Difference in breeding value in days for longevity of individual bulls ($n=54$) with second crop daughters per round of evaluation (interval of 4 months) compared with the last evaluation based on first crop daughters only (round 1); The red dotted lines are the upper and lower bound of the 95%-confidence interval for allowed changes in breeding value based on the increase in reliability between round 1 and the other rounds.

1.6 Aim and outline of thesis

The main aim of my thesis is to investigate the genetics of longevity and develop a genetic evaluation model for the prediction of breeding values for longevity, with the objective that breeding values reflect the true breeding value quicker during early life and therefore breeding values become more stable. To develop this model three parts are investigated first:

1. Genetic analysis of longevity in Dutch dairy cattle using random regression. The aim in **Chapter 2** is to investigate if survival across the life of a cow genetically should be treated as one trait and if lifespan needs to be split in time intervals of 1, 3, 6 or 12 mo.
2. Changes in the genetic level and the effects of age at first calving and milk production on survival during the first lactation over the last 25 years. The aim in **Chapter 3** is to investigate cow survival in the Netherlands for the period 1989–2013, and whether the associations of survival with season of calving, genetic level of survival, AFC and within-herd production level have changed over these decades.
3. Genetic changes of survival traits over the past 25 years in Dutch dairy cattle. The aim in **Chapter 4** is to evaluate if survival in first lactation has become a different trait over the past 25 years or not, and if this is affected by adjusting for production level or not.

In **Chapter 5** the prototype for the genetic evaluation model of longevity in the Netherlands and Flanders is developed using a random regression animal. This Chapter uses the findings from Chapter 2 to 4, and to evaluate the impact of these model changes on the breeding values for longevity of bulls, i.e. that breeding values reflect the true breeding value quicker during early life of the bulls and therefore breeding values become more stable, and compare the new random regression animal model with the current genetic evaluation with a PHM. In the general discussion, **Chapter 6**, the proposed model is first discussed within the context of the specifications defined in Chapter 2 to 4 and other models used worldwide. Following this, longevity and its association with other traits are examined and discussed. Finally, practical considerations for the implementation of a routine genetic evaluation of longevity based on the final model are put forward.

References

- Barkema, H. W., M. A. G. von Keyserlingk, J. P. Kastelic, T. J. G. M. Lam, C. Luby, J. P. Roy, S. J. LeBlanc, G. P. Keefe, and D. F. Kelton. 2015. Invited review: Changes in the dairy industry affecting dairy cattle health and welfare. *J. Dairy Sci.* 98:7426-7445.
- Beaudeau, F., H. Seegers, V. Ducrocq, C. Fourichon, and N. Bareille. 2000. Effect of health disorders on culling in dairy cows: a review and a critical discussion. *Ann. Zootech.* 49:293-311.
- Bergevoet, R. H. M., C. J. M. Ondersteijn, H. W. Saatkamp, C. M. J. van Woerkum, and R. B. M. Huirne. 2004. Entrepreneurial behaviour of Dutch dairy farmers under a milk quota system: goals, objectives and attitudes. *Agric. Syst.* 80:1-21.
- Boettcher, P. J., L. K. Jairath, and J. C. W. Dekkers. 1999. Comparison of methods for genetic evaluation of sires for survival of their daughters in the first three lactations. *J. Dairy Sci.* 82:1034-1044.
- Brotherstone, S., R. F. Veerkamp, and W. G. Hill. 1997. Genetic parameters for a simple predictor of the lifespan of Holstein-Friesian dairy cattle and its relationship to production. *Anim. Sci.* 65:31-37.
- Buenger, A., V. Ducrocq, and H. H. Swalve. 2001. Analysis of survival in dairy cows with supplementary data on type scores and housing systems from a region of northwest Germany. *J. Dairy Sci.* 84:1531-1541.
- Cox, D. R. 1972. Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society. Series B (Methodological)*:187-220.
- Cox, D. R. and D. Oakes. 1984. Analysis of survival data. Chapman & Hall.
- De Vries, A., J. D. Olson, and P. J. Pinedo. 2010. Reproductive risk factors for culling and productive life in large dairy herds in the eastern United States between 2001 and 2006. *J. Dairy Sci.* 93:613-623.
- Dekkers, J. C. and L. Jairath. 1994. Requirements and uses of genetic evaluations for conformation and herd life. Pages 61-68 in *Proc. Proceedings of the fifth world congress on genetics applied to livestock production, Guelph, 7-12 August*.
- Ducrocq, V. 1999. Topics that may deserve further attention in survival analysis applied to dairy cattle breeding: some suggestions. *Interbull Bulletin* 21:181.
- Ducrocq, V. and J. Sölkner. 1994. "The Survival Kit", a FORTRAN package for the analysis of survival data. Pages 51-52 in *Proc. Proc. 5th World Congr. Genet. Appl. Livest. Prod., Guelph, Ontario, Canada*.
- Ducrocq, V. and J. Sölkner. 1998. The Survival Kit—V3.0: A Package for Large Analyses of Survival Data. Pages 447–450 in *Proc. Proc. 6th World Congr. Genet. Appl. Livest. Prod., Armidale, Australia*.

1 General introduction

- Ducrocq, V., J. Sölkner, and G. Mészáros. 2010. Survival Kit v6 - a software package for survival analysis. in Proc. Proc. 9th World Congr. Genet. Appl. Livest. Prod., Leipzig, Germany.
- Essl, A. 1998. Longevity in dairy cattle breeding: a review. *Livest. Prod. Sci.* 57:79-89.
- Fetrow, J., K. V. Nordlund, and H. D. Norman. 2006. Invited review: Culling: Nomenclature, definitions, and recommendations. *J. Dairy Sci.* 89:1896-1905.
- Forabosco, F., J. H. Jakobsen, and W. F. Fikse. 2009. International genetic evaluation for direct longevity in dairy bulls. *J. Dairy Sci.* 92:2338-2347.
- Gengler, N., S. Vanderick, P. Mayeres, A. Gillon, and C. Croquet. 2005. Genetic evaluation of cow survival using a lactation random regression model. *Interbull Bulletin* 33:176.
- Grosu, H., J. Carvalheira, C. Draganescu, and S. L. Lungu. 2013. History of genetic evaluation methods in dairy cattle.
- Haile-Mariam, M. and J. E. Pryce. 2015. Variances and correlations of milk production, fertility, longevity, and type traits over time in Australian Holstein cattle. *J. Dairy Sci.* 98:7364-7379.
- Heise, J., Z. T. Liu, K. F. Stock, S. Rensing, F. Reinhardt, and H. Simianer. 2016. The genetic structure of longevity in dairy cows. *J. Dairy Sci.* 99:1253-1265.
- Heringstad, B., R. Rekaya, D. Gianola, G. Klemetsdal, and K. A. Weigel. 2003. Bivariate analysis of liability to clinical mastitis and to culling in first-lactation cows. *J. Dairy Sci.* 86:653-660.
- Holtsmark, M., B. Heringstad, P. Madsen, and J. Ødegård. 2008. Genetic relationship between culling, milk production, fertility, and health traits in Norwegian Red cows. *J. Dairy Sci.* 91:4006-4012.
- Jamrozik, J., J. Fatehi, and L. R. Schaeffer. 2008. Comparison of models for genetic evaluation of survival traits in dairy cattle: a simulation study. *J. Anim. Breed. Genet.* 125:75-83.
- Kalbfleisch, J. D. and R. L. Prentice. 1980. *The Statistical Analysis of Failure Time Data.* John Wiley and Sons, New York.
- Larroque, H. and V. Ducrocq. 2001. Relationships between type and longevity in the Holstein breed. *Genet. Sel. Evol.* 33:39-59.
- Lubbers, R., S. Brotherstone, V. P. Ducrocq, and P. M. Visscher. 2000. A comparison of a linear and proportional hazards approach to analyse discrete longevity data in dairy cows. *Anim. Sci.* 70:197-206.
- Mészáros, G., J. Sölkner, and V. Ducrocq. 2013. The Survival Kit: Software to analyze survival data including possibly correlated random effects. *Computer methods and programs in biomedicine.* doi:<http://dx.doi.org/10.1016/j.cmpb.2013.01.010>.

- Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819-1829.
- Miglior, F., B. L. Muir, and B. J. Van Doormaal. 2005. Selection indices in Holstein cattle of various countries. *J. Dairy Sci.* 88:1255-1263.
- Mrode, R. A. 2014. *Linear models for the prediction of animal breeding values*. 3 ed. Cabi.
- Powell, R. and P. VanRaden. 2003. Correlation of longevity evaluation with other trait evaluations from 14 countries. *Interbull Bulletin* 30:15.
- Rendel, J. M. and A. Robertson. 1950. Some aspects of longevity in dairy cows. *Empire Journal of Experimental Agriculture* 18:49-56.
- Schaeffer, L. R. 2004. Application of random regression models in animal breeding. *Livest. Prod. Sci.* 86:35-45.
- Schaeffer, L. R. and J. C. M. Dekkers. 1994. Random regressions in animal models for test-day production in dairy cattle. Pages 443-446 in *Proc. Proc. 5th World Congr. Genet. Appl. Livest. Prod., Guelph, Ontario, Canada*.
- Sewalem, A., F. Miglior, G. J. Kistemaker, P. Sullivan, G. Huapaya, and B. J. Van Doormaal. 2007. Short communication: Modification of genetic evaluation of herd life from a three-trait to a five-trait model in Canadian dairy cattle. *J. Dairy Sci.* 90:2025-2028.
- Tsuruta, S., I. Misztal, and T. J. Lawlor. 2004. Genetic correlations among production, body size, udder, and productive life traits over time in Holsteins. *J. Dairy Sci.* 87:1457-1468.
- Van Arendonk, J. A. M. 1985. Studies on the replacement policies in dairy cattle. II. Optimum policy and influence of changes in production and prices. *Livest. Prod. Sci.* 13:101-121.
- Van der Linde, C., G. De Jong, and A. Harbers. 2004. Using a piecewise Weibull mixed model in the genetic evaluation for longevity. *Interbull Bulletin*:157.
- Van der Linde, C., A. Harbers, and G. De Jong. 2007. From functional to productive longevity in the Netherlands. *Interbull Bulletin*:203.
- VanRaden, P. M. and E. J. H. Klaaskate. 1993. Genetic evaluation of length of productive life including predicted longevity of live cows. *J. Dairy Sci.* 76:2758-2764.
- Veerkamp, R. F., S. Brotherstone, B. Engel, and T. H. E. Meuwissen. 2001. Analysis of censored survival data using random regression models. *Anim. Sci.* 72:1-10.
- Visscher, P. M. and M. E. Goddard. 1995. Genetic parameters for milk yield, survival, workability, and type traits for Australian dairy cattle. *J. Dairy Sci.* 78:205-220.

1 General introduction

- Vollema, A. R. 1998. Selection for longevity in dairy cattle. Wageningen University, Wageningen.
- Vollema, A. R. and A. F. Groen. 1997. Genetic correlations between longevity and conformation traits in an upgrading dairy cattle population. *J. Dairy Sci.* 80:3006-3014.
- Weigel, K. A., T. J. Lawlor, P. M. Vanraden, and G. R. Wiggans. 1998. Use of linear type and production data to supplement early predicted transmitting abilities for productive life. *J. Dairy Sci.* 81:2040-2044.
- Zijlstra, J., M. Boer, J. Buiting, K. Colombijn-Van der Wende, and E.-A. Andringa. 2013. Rapport 668: Routekaart Levensduur; Eindrapportage van het project "Verlenging levensduur melkvee". Wageningen UR Livestock Research, Wageningen, the Netherlands.

2

Genetic analysis of longevity in Dutch dairy cattle using random regression

M.L. van Pelt ^{1,2}, T.H.E. Meuwissen³, G. de Jong¹, R.F. Veerkamp²

¹ CRV BV, Animal Evaluation Unit, PO Box 454, 6800 AL Arnhem, the Netherlands;

² Animal Breeding and Genomics Centre, Wageningen UR Livestock Research, PO Box 338, 6700 AH Wageningen, the Netherlands;

³ Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, PO Box 5003, N-1432 Ås, Norway

Journal of Dairy Science (2015) 98:4117-4130

DOI: <http://dx.doi.org/10.3168/jds.2014-9090>

Abstract

Longevity, productive life, or lifespan of dairy cattle is an important trait for dairy farmers, and it is defined as the time from first calving to the last test date for milk production. Methods for genetic evaluations need to account for censored data; that is, records from cows that are still alive. The aim of this study was to investigate whether these methods also need to take account of survival being genetically a different trait across the entire lifespan of a cow. The data set comprised 112,000 cows with a total of 3,964,449 observations for survival per month from first calving until 72 mo in productive life. A random regression model with second-order Legendre polynomials was fitted for the additive genetic effect. Alternative parameterizations were (1) different trait definitions for the length of time interval for survival after first calving (1, 3, 6, and 12 mo); (2) linear or threshold model; and (3) differing the order of the Legendre polynomial. The partial derivatives of a profit function were used to transform variance components on the survival scale to those for lifespan. Survival rates were higher in early life than later in life (99 vs. 95%). When survival was defined over 12-mo intervals survival curves were smooth compared with curves when 1-, 3-, or 6-mo intervals were used. Heritabilities in each interval were very low and ranged from 0.002 to 0.031, but the heritability for lifespan over the entire period of 72 mo after first calving ranged from 0.115 to 0.149. Genetic correlations between time intervals ranged from 0.25 to 1.00. Genetic parameters and breeding values for the genetic effect were more sensitive to the trait definition than to whether a linear or threshold model was used or to the order of Legendre polynomial used. Cumulative survival up to the first 6 mo predicted lifespan with an accuracy of only 0.79 to 0.85; that is, reliability of breeding value with many daughters in the first 6 mo can be, at most, 0.62 to 0.72, and changes of breeding values are still expected when daughters are getting older. Therefore, an improved model for genetic evaluation should treat survival as different traits during the lifespan by splitting lifespan in time intervals of 6 mo or less to avoid overestimated reliabilities and changes in breeding values when daughters are getting older.

Key words: dairy cattle, survival, longevity, random regression

2.1 Introduction

Longevity is of economic importance for dairy farmers because increased longevity helps to increase profitability by (1) reducing the annual costs of replacement of cows; (2) increasing the average herd yield through an increase in the proportion of cows in the higher-producing age groups; (3) better use of a given acreage by reducing the number of replacements to be reared, and therefore allowing an increase in size of the milking herd; and (4) increasing voluntary culling (Rendel and Robertson, 1950). Van Arendonk (1985) showed that when involuntary culling decreases, a higher voluntary culling rate can be applied, resulting in a larger profit for the farmer but not necessarily increasing lifespan as a whole. Hence, an important economic aim for the farmer should be to reduce involuntary culling by using genetic selection.

In animal breeding, selection for longevity is complex because the true longevity of a cow is available only at the end of her life, whereas selection and breeding decisions are made earlier in life. Therefore, censored data are used in genetic evaluations, and expected lifespan needs to be extrapolated from survival data during life. To enable this analysis of censored data, different genetic evaluation models are used (Forabosco et al., 2009). Survival analysis with a proportional hazard function is a popular model used in genetic evaluations for longevity, since the software package Survival Kit was introduced in 1994 (Ducrocq and Sölkner, 1994, 1998, Ducrocq et al., 2010). That model attempts to estimate the probability that an animal will survive to time t given that it has survived to time $t - 1$. Although current genetic evaluations are restricted to a single genetic effect during life, in the newest release it is possible to analyze 2 correlated random effects (Mészáros et al., 2013).

An alternative approach to the survival model is to score survival as a binary trait; for example, survived or not up to a specific endpoint (i.e., time, age, lactation) or survived or not in a specific interval of time. This binary trait for survival can then be analyzed with a linear model or a threshold model, although correlations between breeding values based on a linear and a threshold model are very high and no or little reranking is expected (Boettcher et al., 1999, Holtsmark et al., 2009), which has also been shown for other traits (Meijering and Gianola, 1985, Carlén et al., 2006). Therefore, several countries apply a linear model with survival defined as a binary trait in their routine national genetic evaluations (Forabosco et al., 2009).

If survival is scored as a single binary trait, however, information from culling before the endpoint or information from survival after the endpoint is ignored. In

addition, information from living animals that did not have the opportunity to survive a certain specified time interval cannot be used in the analyses. To overcome these major limitations, an expansion of censored records to expected lifespan has been suggested and implemented in the UK (Brotherstone et al., 1997), and data augmentation (Guo et al., 2001) or multi-trait models including survival to several endpoints have been suggested (Boettcher et al., 1999, Sewalem et al., 2007). The advantage of the multi-trait model is that information for survival after first calving of younger animals still alive can be used to score the traits of the first intervals (they already survived), and the traits for the later intervals can be set to missing; also, variation across time is included in the genetic evaluation. Ideally, many traits with different endpoints should be included in the analysis, but when the life of a cow is split in many binary traits, the multi-trait parameterization becomes impossible and random regressions models might be useful to find a more parsimonious parameterization (Schaeffer and Dekkers, 1994). For this reason, Veerkamp et al. (2001) proposed the random regression model (**RRM**) for survival analysis, which was further investigated by Jamrozik et al. (2008). Veerkamp et al. (2001) proved, by comparison of the likelihood, the equivalence between the survival model and the RRM when a large number of survival intervals were defined in the random regression model. Compared with the survival model, the advantage of the RRM is that, in a linear model context, multiple genetic effects can be fitted. Random regression models therefore enable the modeling of longevity as a different but genetically correlated trait across the entire productive life of a cow.

The assumption that longevity is genetically the same trait during the entire productive life of a cow is often violated, because evidence shows that survival is not necessarily the same trait during the total lifespan of a cow, phenotypically as well as genetically. Cows in higher parities have a higher culling risk compared with earlier parities (Ducrocq, 2005, Terawaki and Ducrocq, 2009, De Vries et al., 2010) and dairy producers cull more intensively for low production in first than in later parities (Dürr et al., 1997, Dürr et al., 1999). Studies using linear multi-trait models showed genetic correlations between parities that significantly differ from unity, averaging 0.77, with a range from 0.33 to 0.96, where correlations decrease with increasing distance between parities (Visscher and Goddard, 1995, Brotherstone et al., 1997, Boettcher et al., 1999, Veerkamp et al., 2001, Sewalem et al., 2007, Holtsmark et al., 2009). Similar results were found with a lactation RRM (Gengler et al., 2005). Hence, this supports the question of whether the use of the survival model with 1 (or 2) genetic effects over the entire life of a cow is optimal, and whether a linear RRM with many time intervals across the life of a cow provides an

alternative. A first step to adopting a new model for genetic evaluation is to determine if survival is genetically the same trait across the entire life of a cow.

Therefore, the objective of this study was to estimate genetic parameters for survival across the life of Dutch dairy cows using random regression and multi-trait models of different orders. However, a concern might be that when the lifespan of a cow is split into many short time intervals, the risk of culling becomes too low in each time interval to apply a linear model. Therefore, both a generalized linear model with a logit link function (GLMM) and an ordinary linear model ignoring the binary nature of the trait were used, and 4 different lengths of time interval for survival were evaluated.

2.2 Materials and Methods

2.2.1 Data

Productive life or lifespan was defined as the time from first calving to the last test date for milk production before the animal died or was culled for slaughter (including dry periods). Data were available from the Dutch/Flemish cattle improvement cooperative CRV (CRV, Arnhem, the Netherlands). The data set was constructed from records of pedigree, lactations, and movements of cows in the Netherlands. Records were constructed for each month that a cow had been present in a herd since the first calving. A cow, culled in month j has $j - 1$ records with score 100 (alive) and record j with score 0 (culled). After culling, monthly records were treated as missing. Records for survival per month in productive life were created for cows containing at least 87.5% black and white genes; that is, Holstein Friesian or Dutch Friesian, and having a first calving between 21 and 40 mo of age. Maximum lactation length was 18 mo and only lactations with the last test date after January 1, 1988, were included. Survival records after 72 mo in productive life were discarded. Animals were required to have the first observation in mo 1 in parity 1; that is, left-censored animals were deleted, because those animals had missing milk production data or changed herds. Right-censored animals (e.g., still alive or exported) were included in the final data set. In the period from 1988 to 2012, the total data set comprised 354,014,073 records from 10,004,624 animals in 44,550 herds. For the parameter estimation, the total data set was reduced by criteria on the herd level. Herds with at least 95% black and white genes were selected. Only herds with at least 60 cows present every month in the period from 1993 to 2012 were selected. This was done to exclude herds with nonconventional culling management, because, for example, the farm could have stopped milk production and sold animals all at once or in

batches, or the entire herd could have been culled at once because of a breakout of a disease. Because of the exclusion of left-censored animals, selection of herds was not based on the entire period, but started in 1993. In total, 560 herds fulfilled these criteria and contained 477,105 animals with 15,958,047 records for monthly survival. Additional requirements were that (1) sires had, as sire or as maternal grandsire or combined, at least 12 progeny that could have been productive for at least 36 mo after first calving; (2) sires had progeny in at least 2 herds, and (3) every herd-year-month class had at least 50 observations. These additional requirements for progeny per sire, herds per sire, and observations per herd-year-month were repeated until the final data set met all criteria. The resulting data set consisted of 112,000 cows with 3,964,449 observations for monthly survival originating from 153 herds and 1,329 sires and maternal grandsires. Pedigree information of these sires was traced back 6 generations, resulting in a pedigree file with 1,718 sires.

The data set with monthly survival records was used to create 3 other data sets with different lengths for the time interval for survival of 3, 6, and 12 mo. The records for these 3 data sets were created by combining monthly records for survival to 1 record per 3, 6, or 12 mo for survival. Compared with the initial data set, the resulting data sets contained fewer animals because living animals without information up to the first endpoint were excluded from the data set, because it was not possible to determine whether the cow reached that endpoint or not. An overview of all 4 data sets is given in Table 2.1.

Table 2.1. Number of records and cows for 4 data sets for survival, where each data set has a different length for the time interval for survival

Data set	Time interval (mo)	No. of records	No. of cows
1	1	3,964,449	112,000
2	3	1,338,371	109,458
3	6	682,225	107,595
4	12	354,737	105,716

2.2.2 Statistical Model

To determine whether heritability for survival changes during the total lifespan and to what extent survival is genetically the same trait across the lifespan, genetic parameters were estimated using a random regression sire-maternal grandsire model as base:

2 Genetic analysis of longevity

$$y_{klmno} = \mu + FIXED + \sum_{i=0}^2 hc_{ik} t_l + \left(\sum_{i=0}^p sire_{im} t_l + \frac{1}{2} \times \sum_{i=0}^p mgs_{in} t_l \right) + \varepsilon_{klmno}$$

where y is the binary observation for survival of an animal in herd \times 5-yr period k , with sire m and maternal grandsire n , for time interval 1, 3, 6 or 12 mo; μ is general mean; *FIXED* included fixed effects of parity \times year of calving, a third-order Legendre polynomial on time interval, and a second-order Legendre polynomial on age at first calving; hc_{ik} is the i th-order random regression coefficient of the second-order Legendre polynomial on time interval t_l for the herd curve effect of herd \times 5-yr period k ; $sire_{im}$ and mgs_{in} are the i th-order random regression coefficients of the p th-order Legendre polynomial on time interval t_l for the additive genetic effect of sire m and maternal grandsire n ; and ε_{klmno} is the residual effect. The random sire effects were fitted by overlaying the relationship matrix for sire and one-half the maternal grandsire matrix, resulting in one estimate per effect for a sire. Variance components were estimated with ASReml (Gilmour et al., 2009).

The permanent environment effect of an animal is usually fitted in RRM. However, with survival data, a record of zero indicates that the animal died in the time interval and no subsequent records will follow for this animal. Thus, as only one of the categories can be repeated (i.e., records coded alive), such data structures are not well suited for estimation of permanent environment effects (Ødegård et al., 2006). We tested this effect and the permanent environment variance was close to zero in all models. Therefore, we excluded it from the final analysis.

In our study, we analyzed productive longevity instead of functional longevity. Productive longevity is defined as the ability to delay culling, both voluntary (i.e., for production) as well as involuntary. Functional longevity is the longevity of a cow adjusted for individual milk production relative to the production level of the herd, and it is suggested as a way to exclude voluntary culling from the breeding value (Robertson, 1966). However, this adjustment has been under debate in the Netherlands (e.g., little voluntary culling for yield, health also affects yield), and therefore we did not investigate functional longevity here.

Residual variances were estimated for intervals of 3 mo for the time interval of 1 mo and for each time interval for the time intervals of 3, 6, and 12 mo, and were assumed to be uncorrelated. Heritabilities for each time interval were calculated as $4 \times$ sire variance divided by the sum of $1.25 \times$ sire variance (1.25 because both the sire and the maternal grandsire, i.e., $0.25 \times$ sire variance, were fitted in the model), the herd curve variance, and the residual variance of the corresponding interval.

Estimates of the (co)variance components for the random effects are functions of time and were expanded to (co)variances attributable to time intervals by $\Phi'V\Phi$, where Φ is an $l \times p$ matrix whose elements are the Legendre coefficients for time intervals l , and V is $p \times p$ matrix with the estimated covariance components for the additive genetic effect or the herd curve effect. The resulting $l \times l$ matrix contained the (co)variances for all l /time intervals. Heritabilities for survival per time interval l were derived from these $\Phi'V\Phi$ matrices. To compute approximate standard errors for variances and heritabilities for all time intervals using the method described by Fischer et al. (2004), an automated process (to generate ASReml pin files) was implemented. The statistical significance of the estimated parameters was tested with a chi-squared test with 1 degree of freedom ($P < 0.05$).

2.2.3 Alternative Model Parameterizations

To estimate the genetic parameters, alternative parameterizations of the base statistical model were considered. For each model also, breeding values were predicted and correlations were calculated between the breeding values (**EBV**) for lifespan (LS) based on 555 sires with at least 10 daughters at mo 36 in productive life.

Trait Definition. Traits were defined by varying the length of the time interval for survival, and time intervals of 1, 3, 6, and 12 mo were analyzed using the base model with a second-order Legendre polynomial for the genetic effect. For the length of the time interval of 1 mo, it was possible to correct for the dry period, and fixed classes were used for parity \times month in lactation, with a separate class for dry cows, instead of a third-order Legendre polynomial on the time interval. Because of the short time interval, it was possible to fit fixed effects for year \times season of observation and for parity \times year \times season of calving.

Linear or Threshold Model. To test the suitability of using a linear model for analyzing survival as a binary trait, especially when the risk of culling becomes smaller with shorter time intervals, a generalized linear model with a logit link function (GLMM) was also fitted to all 4 data sets. The same base model was used, with a second-order Legendre polynomial for the genetic effect. Different residual classes were fitted as in the base model, except for the threshold model with the length of the time interval for survival of 12 mo, which showed convergence problems. Therefore, only one residual class was fitted in this case.

Order of Legendre Polynomial. Different orders of the random regression component were tested and, as a result, different numbers of genetic coefficients were estimated in the statistical model. The genetic effect was modeled with a second-order (RR2) or fifth-order (RR5) Legendre polynomial or as the full multi-

trait model (MT) fitting an interaction between genetic effect and time interval. The interaction between genetic effect and the 12 time intervals is equivalent to fitting a 11th-order Legendre polynomial. For this analysis, observations with a time interval of 6 mo were used, because it was feasible to get convergence for RR5 and MT.

2.2.4 Transforming Survival to Lifespan

In animal breeding, we are interested in improving total longevity; that is, LS. In the statistical analysis, survival was modeled as survival in a time interval i conditional on survival up to the previous time interval $i - 1$. An animal that survived time intervals $i - 1$ and i was coded “alive-alive”, and an animal that was culled in time interval i was coded “alive-culled” for time intervals $i - 1$ and i . To transform these conditional survival probabilities into total lifespan, the following transformation from survival into LS was performed.

Let π_{ij} be the probability that an animal j dies in month i , and survives $i - 1$ months. So, the mean lifespan of animal j is $LS_j = \sum_i \pi_{ij}(i - 1)$. The probability that animal j dies in month i follows a binomial distribution and equals the probability that it survives mo 1, 2, ..., $(i - 1)$ times that it dies in month i ; thus, $\pi_{ij} = (1 - p_{ij}) \prod_{k=1}^{(i-1)} p_{kj}$, where p_{ij} is the probability that animal j survives month i ; that is, $p_{ij} = p_i + g_{ij}$, with p_i denoting the average survival rate in month i , and g_{ij} denotes the genetic value of animal j for surviving month i , as estimated by the above analysis. For example, if we consider a maximum LS of 4 mo; that is, every animal that survives up to mo 3 dies in mo 4 ($p_{4j} = 0$), the lifespan of animal j can be written as

$$LS_j = 0(1 - p_{1j}) + 1(p_{1j} - p_{1j}p_{2j}) + 2(p_{1j}p_{2j} - p_{1j}p_{2j}p_{3j}) + 3p_{1j}p_{2j}p_{3j}, \quad [1]$$

$$LS_j = p_{1j} + p_{1j}p_{2j} + p_{1j}p_{2j}p_{3j}, \quad [2]$$

Transforming phenotypic survival to LS is relatively straightforward with equation [2], and extension to more months or different time intervals is also straightforward. However, to obtain breeding values and variance components for LS from the random regressions components, it is important to account for the mean survival rate per time interval in the multiplicative effects. For this transformation, equation [2] can be redefined as a (part-)profit function (other non-LS related parts of the profit function are not relevant here), which expresses lifespan as a linearized equation [2] of survival in each time interval where we have breeding values for (p_{ij}). Thus, estimation of the breeding value for lifespan follows

that of the usual estimation of a breeding value for profit; that is, the aggregate breeding value for LS of animal j is

$$g_{LSj} = v_1 g_{1j} + v_2 g_{2j} + v_3 g_{3j}, \quad [3]$$

where v_1 is the partial derivative of equation [2] to p_{ij} evaluated at the trait means (e.g., $v_1 = 1 + p_2 + p_2 p_3$ for mo 1). Following this definition, calculation of the variance components for LS and the correlation between survival until a certain endpoint and LS becomes straightforward following selection index theory: the variance of the aggregate breeding value for LS is $\mathbf{v}'\Phi'\mathbf{V}\Phi\mathbf{v}$, where \mathbf{v} is a (number of time intervals \times 1) vector of partial derivatives v_i .

To compare the variance components (and EBV) from the threshold model with the linear model, a transformation was needed. The threshold model here used a logit link function; to transform the estimated variances in each time period back to the linear scale, the mean survival rate in each time interval p_i in equation [2] needs to be transformed to the mean survival on the log scale (p_i^*) by using

$$p_i^* = \log\left(\frac{p_i}{1-p_i}\right).$$

Then, the deviance around the mean survival in each time interval was calculated by adding and subtracting 1 genetic standard deviation (calculated with $\Phi'\mathbf{V}\Phi$) to these means on the log scale, $p_{i+1sd}^* = p_i^* + \sigma_{g_i^*}$ and $p_{i-1sd}^* = p_i^* - \sigma_{g_i^*}$. Subsequently, these deviances in each interval were retransformed to the linear scale as

$$p_{i\pm 1SD} = \left(\frac{e^{p_{i\pm 1sd}^*}}{1+e^{p_{i\pm 1sd}^*}}\right).$$

Then, the genetic variance can be calculated by

$$\sigma_{g_i}^2 = \left(\frac{p_{i+1sd}^* - p_{i-1sd}^*}{2}\right)^2,$$

and genetic covariances between time intervals on the linear scale are

$$\sigma_{g_{i,j}} = r_{g_{i,j}}^* \sqrt{\sigma_{g_i}^2 * \sigma_{g_j}^2},$$

where $r_{g_{i,j}}^*$ is the genetic correlation. Having the genetic (co)variances on the linear scale, genetic parameters for LS can be calculated with the partial derivatives of the part-profit function, as described above.

2.3 Results

2.3.1 Mean Survival

The average lifespan was 41.2 mo from first calving to last test date for the culled cows in the data set. After 72 mo in productive life, 17.2% of the cows were still alive and producing milk. Figure 2.1 shows the survival rate per time interval for cows up to 72 mo in productive life for the 4 trait definitions for time interval. The survival rate decreased with an increasing number of months in productive life. Survival rates per time interval were lower with increased length of the time interval, with survival rates between 96 and 100% for survival per 1 mo, between 88 and 98% for survival per 3 mo, between 78 and 95% for survival per 6 mo, and between 61 and 88% for survival per 12 mo. Moreover, survival curves became smoother with increased length of time interval. The survival curves were spikier when we used intervals of 1 and 3 mo, with the peaks approximately 1 yr apart, and smoother when the intervals were 6 and 12 mo. With relatively short time intervals (i.e., 1 and 3 mo), it was possible to capture more detailed information on survival compared with longer time intervals.

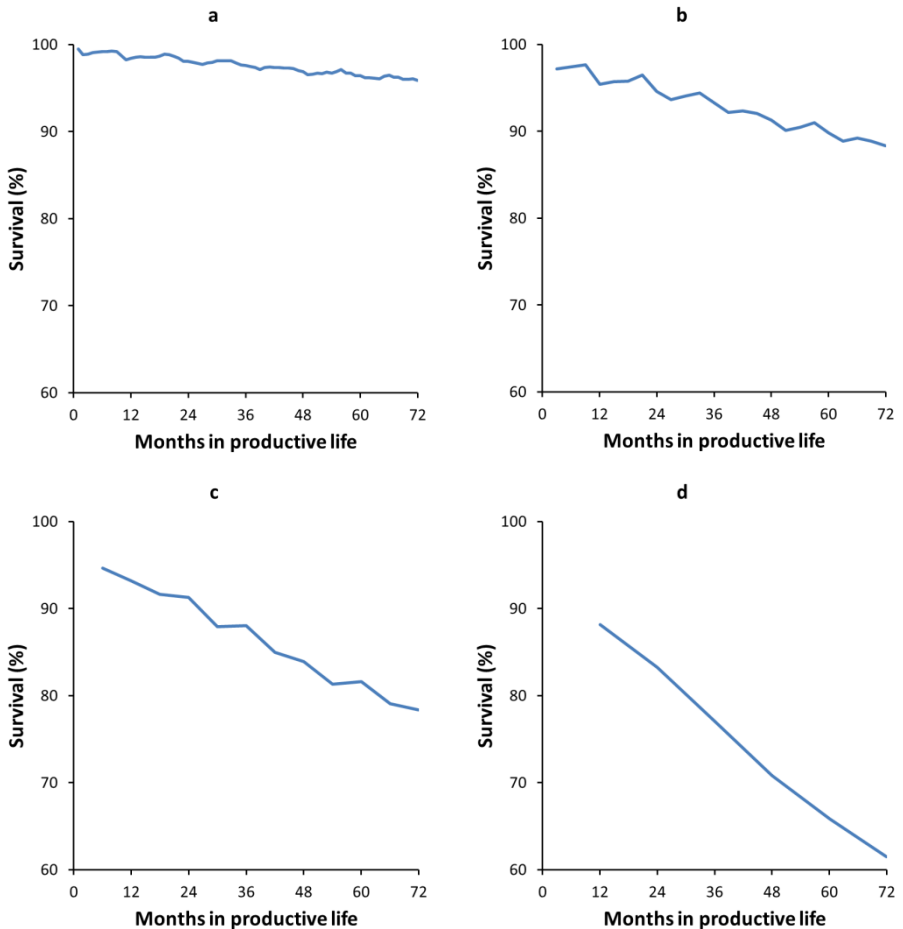


Figure 2.1. Average survival rate for 4 different lengths for the time interval for survival (panels a, b, c, and d show observations per 1, 3, 6, and 12 mo, respectively) for Dutch dairy cows with observations in the period 1988 to 2012.

2.3.2 Trait Definition

Heritabilities for survival in each time interval were (very) low (Table 2.2), especially for the intervals of 1 and 3 mo (0.002 to 0.014). Still, they were significantly different from zero, and the phenotypic variances were similar to the expected variance from the binomial distribution, being $p(1 - p)$. Hence, these low heritabilities reflect the trait definition and the low incidence of culling for time intervals of 1 and 3 mo. The heritabilities for lifespan over the entire period of 72 mo after first calving were considerably higher, ranging from 0.115 to 0.149, with

2 Genetic analysis of longevity

genetic standard deviations ranging from 7.0 to 8.0 mo. Phenotypic standard deviations for lifespan were comparable for the 4 data sets: 20.4 mo for the interval of 1 mo, 20.6 for the interval of 3 mo, 20.7 mo for the interval of 6 mo, and 20.8 mo for the interval of 12 mo.

Table 2.2. Heritabilities for survival (h^2) and genetic standard deviations (σ_g) in months estimated with a random regression model with a second-order Legendre polynomial for the genetic effect¹

Month	Time interval							
	1 mo		3 mo		6 mo		12 mo	
	h^2	σ_g	h^2	σ_g	h^2	σ_g	h^2	σ_g
6	0.003	0.005	0.005	0.03	0.010	0.13		
12	0.002	0.005	0.004	0.04	0.011	0.15	0.022	0.567
18	0.003	0.006	0.007	0.05	0.014	0.19		
24	0.003	0.007	0.008	0.06	0.017	0.22	0.034	0.77
30	0.003	0.008	0.009	0.07	0.018	0.25		
36	0.004	0.009	0.010	0.08	0.021	0.28	0.049	0.991
42	0.004	0.010	0.011	0.08	0.021	0.30		
48	0.005	0.012	0.011	0.09	0.023	0.32	0.060	1.191
54	0.006	0.013	0.012	0.09	0.023	0.35		
60	0.007	0.015	0.012	0.10	0.026	0.37	0.072	1.378
66	0.009	0.017	0.013	0.11	0.027	0.40		
72	0.011	0.020	0.014	0.11	0.031	0.43	0.075	1.581
LS	0.128	7.28	0.115	6.96	0.121	7.22	0.149	8.02

¹Survival was defined for 4 different lengths for the time interval for survival (1, 3, 6, and 12 mo), and heritabilities are shown for every sixth month in productive life up to 72 mo, and heritability for lifespan (LS) is shown.

²Parameters ending in specific months are shown; for the 1-mo interval, parameters of mo 6 are shown; for the 3-mo interval, parameters of mo 4 to 6 are shown; for the 6-mo interval, parameters of mo 1 to 6 are shown; and for the 12-mo interval, there is no estimate at 6 mo but the first parameters are at mo 12. Standard errors ranged from 0.0002 to 0.0015 for the 1-mo time interval, from 0.0005 to 0.0023 for the 3-mo interval, from 0.0012 to 0.0054 for the 6-mo interval, and from 0.0030 to 0.0106 for the 12-mo interval.

For the data set with survival rates defined per 1 mo, genetic correlations between intervals ranged from 0.36 to 1.00 (Table 2.3, below diagonal); for the 3 data sets with other intervals, genetic correlations ranged from 0.66 to 1.00 (Tables 2.4, 2.5, and 2.6, below diagonal). The genetic correlations between time intervals not far apart were close to unity and departed from unity with increasing distance between the intervals, especially for the data set where an interval of 1 mo was used. With the 3 data sets with survival defined over longer time intervals, most

genetic correlations were 0.90 and higher, except for the first time interval, 6 or 12 mo, for which the genetic correlations were 0.80 or lower.

Table 2.3. Genetic correlations between 1-mo intervals for survival, estimated with a linear (below diagonal) and a threshold (above diagonal) random regression model with a second-order Legendre polynomial for the genetic effect¹

Mo	6	12	18	24	30	36	42	48	54	60	66	72
6		0.99	0.94	0.88	0.80	0.72	0.64	0.57	0.49	0.41	0.33	0.25
12	0.97		0.99	0.95	0.89	0.83	0.76	0.68	0.60	0.51	0.40	0.29
18	0.90	0.98		0.99	0.95	0.91	0.85	0.78	0.70	0.60	0.48	0.35
24	0.84	0.95	0.99		0.99	0.96	0.92	0.86	0.78	0.68	0.55	0.40
30	0.78	0.91	0.96	0.99		0.99	0.96	0.92	0.85	0.76	0.62	0.46
36	0.72	0.86	0.93	0.97	0.99		0.99	0.96	0.91	0.82	0.69	0.54
42	0.67	0.80	0.88	0.93	0.97	0.99		0.99	0.95	0.88	0.77	0.62
48	0.60	0.73	0.81	0.88	0.93	0.97	0.99		0.99	0.94	0.85	0.71
54	0.54	0.66	0.74	0.81	0.88	0.93	0.97	0.99		0.98	0.92	0.81
60	0.48	0.58	0.67	0.74	0.82	0.88	0.93	0.97	0.99		0.98	0.91
66	0.42	0.51	0.59	0.67	0.75	0.82	0.89	0.94	0.98	0.99		0.98
72	0.36	0.44	0.52	0.60	0.69	0.77	0.84	0.90	0.95	0.98	1.00	

¹Standard errors for the genetic correlations between 1-mo intervals ranged from 0.0003 to 0.114 for the linear model and from 0.0004 to 0.111 for the threshold model.

Table 2.4. Genetic correlations between 3-mo intervals for survival, estimated with a linear (below diagonal) and a threshold (above diagonal) random regression model with a second-order Legendre polynomial for the genetic effect¹

Mo	6	12	18	24	30	36	42	48	54	60	66	72
6		0.97	0.88	0.77	0.67	0.60	0.56	0.55	0.58	0.64	0.72	0.79
12	0.95		0.97	0.90	0.84	0.78	0.75	0.74	0.76	0.80	0.84	0.84
18	0.87	0.98		0.98	0.94	0.91	0.89	0.88	0.89	0.90	0.90	0.84
24	0.82	0.96	0.99		0.99	0.97	0.96	0.95	0.95	0.94	0.91	0.81
30	0.78	0.94	0.98	1.00		1.00	0.99	0.98	0.98	0.96	0.90	0.77
36	0.76	0.92	0.97	0.99	1.00		1.00	0.99	0.99	0.96	0.90	0.75
42	0.75	0.91	0.96	0.98	0.99	1.00		1.00	0.99	0.97	0.90	0.75
48	0.74	0.90	0.95	0.97	0.98	0.99	1.00		1.00	0.98	0.91	0.77
54	0.74	0.88	0.93	0.95	0.97	0.98	0.99	1.00		0.99	0.94	0.81
60	0.74	0.86	0.91	0.93	0.94	0.96	0.97	0.99	1.00		0.98	0.88
66	0.73	0.84	0.87	0.89	0.91	0.92	0.94	0.96	0.98	1.00		0.96
72	0.72	0.80	0.83	0.84	0.86	0.88	0.91	0.93	0.96	0.98	0.99	

¹Standard errors for the genetic correlations between 3-mo intervals ranged from 0.0002 to 0.106 for the linear model and from 0.0002 to 0.100 for the threshold model.

2 Genetic analysis of longevity

Table 2.5. Genetic correlations between 6-mo intervals for survival, estimated with a linear (below diagonal) and a threshold (above diagonal) random regression model with a second-order Legendre polynomial for the genetic effect¹

Mo	6	12	18	24	30	36	42	48	54	60	66	72
6		0.96	0.85	0.72	0.62	0.54	0.50	0.48	0.50	0.54	0.61	0.70
12	0.94		0.96	0.89	0.82	0.76	0.72	0.71	0.72	0.74	0.78	0.80
18	0.85	0.98		0.98	0.94	0.90	0.88	0.87	0.87	0.88	0.88	0.84
24	0.79	0.95	0.99		0.99	0.97	0.96	0.95	0.94	0.94	0.91	0.84
30	0.75	0.93	0.98	1.00		1.00	0.99	0.98	0.98	0.96	0.92	0.82
36	0.73	0.91	0.97	0.99	1.00		1.00	0.99	0.99	0.97	0.92	0.81
42	0.73	0.90	0.96	0.98	0.99	1.00		1.00	0.99	0.98	0.93	0.81
48	0.73	0.89	0.95	0.97	0.98	0.99	1.00		1.00	0.98	0.94	0.82
54	0.74	0.88	0.92	0.94	0.96	0.97	0.98	1.00		0.99	0.96	0.86
60	0.74	0.86	0.89	0.90	0.92	0.94	0.96	0.98	0.99		0.98	0.91
66	0.74	0.82	0.84	0.85	0.87	0.89	0.92	0.95	0.97	0.99		0.97
72	0.73	0.78	0.78	0.79	0.80	0.83	0.86	0.90	0.94	0.97	0.99	

¹Standard errors for the genetic correlations between 6-mo intervals ranged from 0.0007 to 0.085 for the linear model and from 0.0006 to 0.089 for the threshold model.

Table 2.6. Genetic correlations between 12-mo intervals for survival, estimated with a linear (below diagonal) and a threshold (above diagonal) random regression model with a second-order Legendre polynomial for the genetic effect¹

Mo	12	24	36	48	60	72
12		0.91	0.82	0.77	0.73	0.68
24	0.92		0.98	0.96	0.92	0.85
36	0.83	0.98		0.99	0.97	0.89
48	0.77	0.95	0.99		0.99	0.93
60	0.72	0.90	0.95	0.99		0.97
72	0.66	0.82	0.89	0.94	0.98	

¹Standard errors for the genetic correlations between 12-mo intervals ranged from 0.003 to 0.075 for the linear model and from 0.002 to 0.082 for the threshold model.

In addition, genetic correlations between lifespan and cumulative time intervals up to mo 6 in productive life ranged from 0.79 to 0.85 for intervals of 1, 3, and 6 mo, and genetic correlations increased to unity with longer cumulative intervals for survival (Figure 2.2). Genetic correlations with lifespan were >0.95 when the cumulative interval was longer than 30 mo for the interval of 1 mo or 24 mo for the other time intervals.

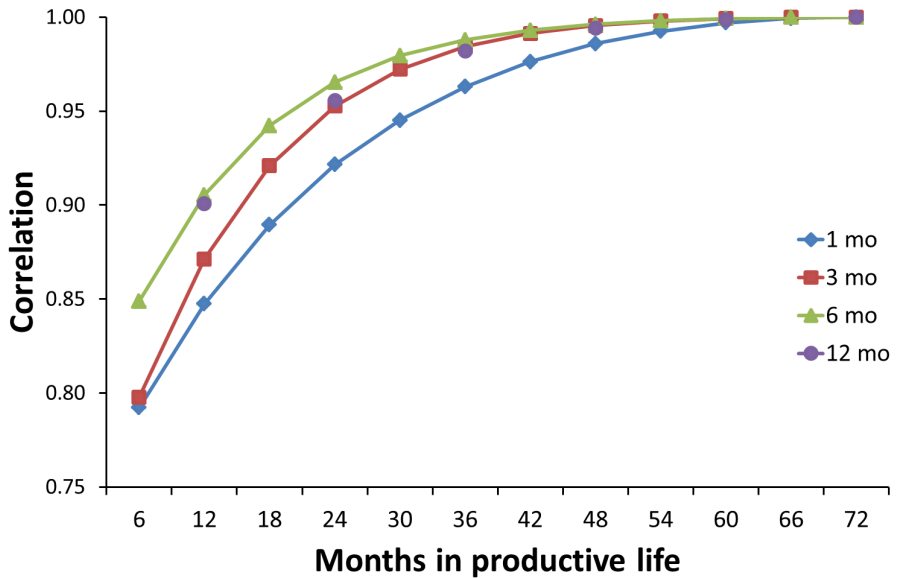


Figure 2.2. Genetic correlations between cumulative intervals for survival for 4 different lengths for the time interval for survival (1, 3, 6, and 12 mo) and lifespan, where the genetic effect is estimated with a linear random regression model with a second-order Legendre polynomial.

2.3.3 Linear or Threshold Model

The genetic correlations between time intervals of the threshold model were comparable with those of the linear model, where the greatest differences were between the intervals at the beginning or end of the 72-mo period with the other intervals (Tables 2.3, 2.4, 2.5, and 2.6, above diagonal). The greatest difference in genetic correlation was 0.25 between the linear and threshold models. For genetic correlations between cumulative intervals and lifespan (Figure 2.3), differences between the linear and threshold models were smaller (<0.10) and occurred only for correlations of cumulative survival in early life with lifespan.

2 Genetic analysis of longevity

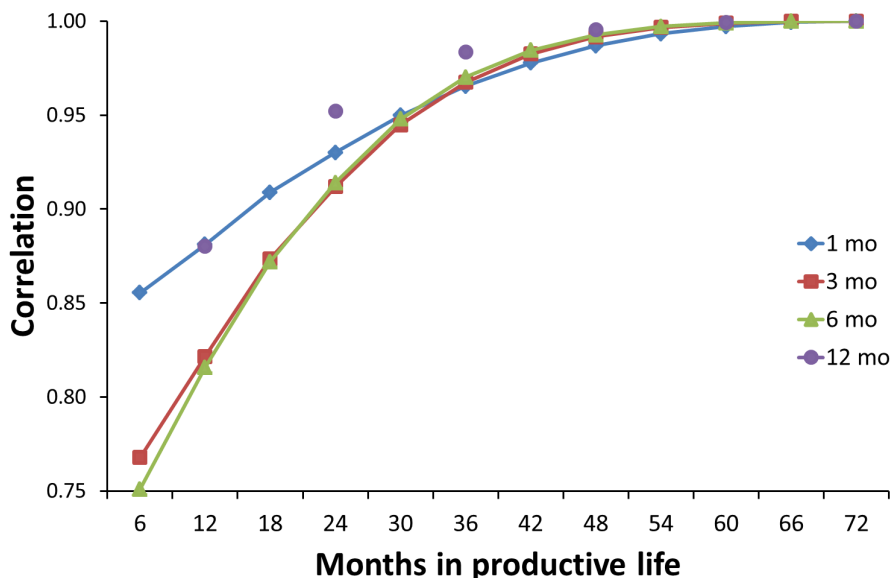


Figure 2.3. Genetic correlations between cumulative intervals for survival for 4 different lengths for the time interval for survival (1, 3, 6, and 12 mo) and lifespan, where the genetic effect is estimated with a threshold random regression model with a second-order Legendre polynomial.

2.3.4 Order of Legendre Polynomial

Heritabilities for the 3 models (RR2, RR5, and MT) fitted to the data set with survival per 6 mo are shown in Figure 2.4. The heritability for RR2 ranged from 0.010 to 0.031, for RR5 from 0.012 to 0.032, and for MT from 0.013 to 0.036. Heritability for MT showed a more spiked pattern, as expected, because the RR fits a more parsimonious model. For MT, the first 2 intervals only include survival of first-parity cows and show that survival in early lactation is less heritable than survival in late lactation. Two other peaks are clearly visible at mo 24 and 36. These peaks are mainly from cows in second and third lactations, and indicate higher heritabilities in late lactation compared with early lactation. The heritabilities and genetic standard deviations for lifespan over the entire period of 72 mo for the 3 parameterizations were similar: 0.121 and 7.2 mo for RR2, 0.124 and 7.3 mo for RR5, and 0.130 and 7.5 mo for MT. Heritabilities for lifespan were similar, although heritabilities for intervals of 36 to 60 mo were higher for RR5 and MT compared with RR2. Lifespan parameters were calculated with $\mathbf{v}'\Phi'\mathbf{V}\Phi\mathbf{v}$, where \mathbf{v} is the vector with partial derivatives. The intervals in early life have more weight in \mathbf{v} than the intervals later in life; therefore, intervals later in life influenced lifespan

heritability less than did intervals early in life. Besides the partial derivatives v , the genetic correlations between life stages also affected the heritability of lifespan, and genetic correlations for RR5 and MT were on average lower than those for RR2 (-0.04 and -0.15).

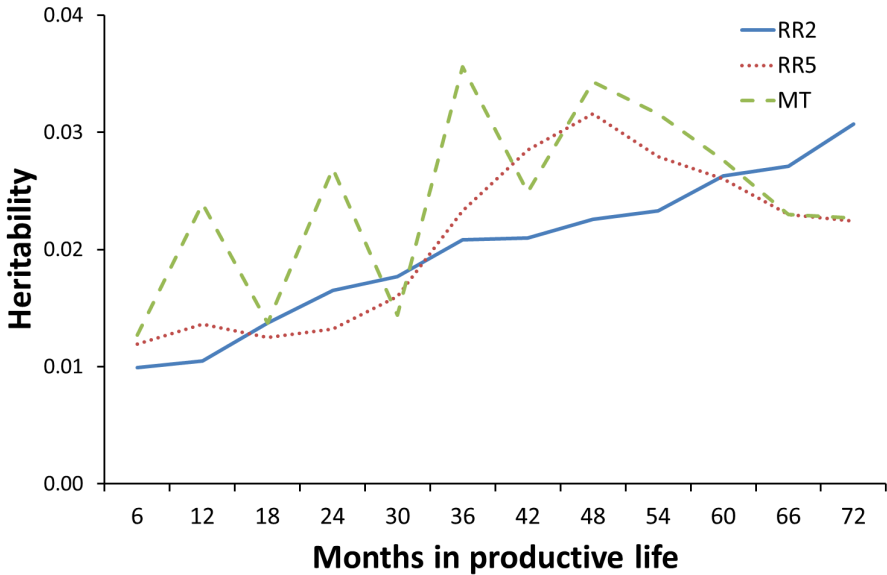


Figure 2.4. Heritability for survival with a 6-mo interval, where the genetic effect is estimated with a random regression model with a second-order (RR2) or a fifth-order (RR5) Legendre polynomial or as interaction between genetic effect and interval of 6 mo (MT).

For comparison with the genetic correlations for RR2 (Table 2.4), genetic correlations for RR5 and MT are shown in Table 2.7. With more genetic effects fitted in the model, genetic correlations between adjacent intervals deviated more from unity. For example, the correlation between mo 6 and 12 was 0.94 for RR2, 0.73 for RR5, and 0.51 for MT. Genetic correlations between cumulative survival and LS were comparable between the parsimonious models RR2 and RR5 as well as MT (Figure 2.5). In very early life, up to mo 6, the genetic correlations with LS became lower when a higher order of fit was used, with the lowest correlation for MT (0.76) and highest correlation for RR2 (0.85).

2 Genetic analysis of longevity

Table 2.7. Genetic correlations between 6-mo intervals for survival, where the genetic effect is modeled with a random regression model with a fifth-order Legendre polynomial (RR5, below diagonal) and as interaction between genetic effect and interval of 6 mo (MT, above diagonal)¹

Mo	6	12	18	24	30	36	42	48	54	60	66	72
6		0.49	0.62	0.51	0.75	0.59	0.67	0.59	0.71	0.69	0.65	0.71
12	0.73		0.31	0.91	0.58	0.88	0.75	0.81	0.63	0.64	0.52	0.48
18	0.72	0.96		0.34	0.77	0.47	0.62	0.48	0.71	0.66	0.69	0.71
24	0.75	0.86	0.96		0.69	0.92	0.83	0.89	0.69	0.69	0.61	0.55
30	0.76	0.85	0.94	0.99		0.73	0.85	0.79	0.91	0.84	0.80	0.79
36	0.74	0.88	0.94	0.96	0.99		0.94	0.94	0.80	0.83	0.73	0.70
42	0.71	0.89	0.92	0.92	0.96	0.99		0.95	0.91	0.91	0.83	0.80
48	0.70	0.88	0.91	0.91	0.94	0.98	1.00		0.86	0.88	0.80	0.74
54	0.72	0.82	0.87	0.91	0.94	0.96	0.97	0.98		0.95	0.90	0.88
60	0.72	0.67	0.77	0.86	0.89	0.88	0.87	0.89	0.96		0.94	0.93
66	0.69	0.54	0.67	0.78	0.81	0.78	0.75	0.78	0.88	0.98		0.97
72	0.73	0.78	0.82	0.83	0.85	0.88	0.89	0.91	0.95	0.94	0.90	

¹Standard errors for the genetic correlations between 6-mo intervals ranged from 0.002 to 0.143 for RR5 and from 0.047 to 0.167 for MT.

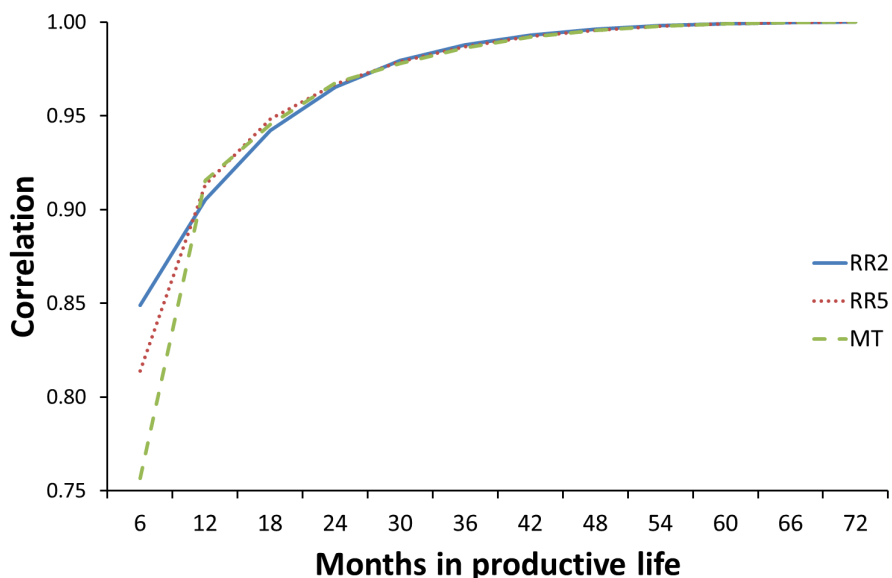


Figure 2.5. Genetic correlations between cumulative intervals for survival with a 6-mo interval and lifespan, where the genetic effect is estimated with a random regression model with a second-order (RR2) or a fifth-order (RR5) Legendre polynomial or as interaction between genetic effect and interval of 6 mo (MT).

2.3.5 Breeding Values for Lifespan

Correlations among EBV for lifespan of sires from all linear and threshold random regression models with a second-order Legendre polynomial for the genetic effect are presented in Table 2.8. Only bulls with at least 10 daughters at mo 36 in productive life were included. In general, correlations between EBV of different time intervals were (very) high between linear and threshold models (≥ 0.97). However, EBV were more sensitive to trait definition, and correlations between EBV for LS became much lower when estimated from survival data split across different intervals. Especially comparing LS breeding values from 12-mo intervals with those coming from 1- and 3-mo intervals gave correlations around 0.60 and 0.80, respectively. When computing EBV based on a time interval of 6 mo, very little reranking was expected with the different parameterizations because all correlations were ≥ 0.99 (Table 2.9).

Table 2.8. Correlations between EBV of lifespan for sires based on linear and threshold random regression models with a second-order Legendre polynomial for the genetic effect with 4 different lengths for the time interval for survival (1, 3, 6, and 12 mo)¹

Time interval	1 mo	3 mo	6 mo	12 mo
1 mo	0.99	0.93	0.85	0.59
3 mo	0.96	0.98	0.98	0.82
6 mo	0.91	0.98	0.97	0.91
12 mo	0.61	0.78	0.87	0.98

¹Correlations between linear and threshold model for same time interval in bold on diagonal, correlations between different time intervals with linear model below diagonal, and correlations between different time intervals with threshold model above diagonal.

Table 2.9. Correlations between EBV of lifespan for sires based on linear and threshold models with a 6-mo interval for survival, where the genetic effect is modeled as second- or fifth-order Legendre polynomial (RR2 and RR5) or as interaction between genetic effect and interval of 6 mo (MT)¹

Model	RR2	RR5	MT
RR2	0.97	1.00	1.00
RR5	1.00	0.98	1.00
MT	0.99	1.00	0.98

¹Correlations between linear and threshold model for same model in bold on diagonal, correlations between different models with linear model below diagonal, and correlations between different models with threshold model above diagonal.

2.4 Discussion

The major objective of this paper was to investigate whether survival is the same trait across the life of a cow. For this purpose, analyses were performed until 72 mo in productive life. Phenotypically, survival rates were not constant during the productive life of a cow in this study, because monthly survival percentages were higher in early life than in later life, and survival curves became smoother when intervals of 12 mo were taken to define survival compared with shorter time intervals. The cyclic pattern observed for survival was also shown by Pritchard et al. (2013) for the UK population, with clear peaks of numbers of culled animals at the end of lactation or the start of a new lactation. In first and second lactation, the peaks for numbers of culled animals were more prominent and in later lactation, the peaks faded out over multiple months, which is in accordance with our results (Figure 2.1), because cows could be in different lactations at the same month in productive life because of different calving intervals. Roxström et al. (2003) proposed a piecewise baseline hazard function to model survival for each lactation number \times lactation stage group to have a better fit of the data in studies using the Survival Kit software. Differences found in our study in survival rates within and between lactations were, although expressed on a different scale, in accordance with the larger relative culling risks with increasing lactation number and lactation stage (Roxström and Strandberg, 2002, Ducrocq, 2005, Terawaki and Ducrocq, 2009, Sasaki et al., 2012).

All applied alternative parameterizations of survival—different trait definitions, linear versus threshold model, and different orders of the Legendre polynomial—resulted in nonunity genetic correlations. From all of these parameterizations, the interval over which survival was defined was clearly the most critical factor. Correlations between EBV for LS of linear and threshold models were very high and resulted in little reranking of bulls. Also, the order of the Legendre polynomial (RR2, RR5, or MT) for the genetic effect had some effect on the correlation structure for survival at different time intervals but had little effect on the correlations with LS or the correlation between the breeding values.

Genetic correlations seem to be lower for survival defined over shorter time intervals than when survival was defined for a whole year, as was shown at the level of the mean in Figure 2.1. In addition, the correlation with LS became much weaker. The reason for this is that there were different reasons for culling in the first year at different stages in lactation, and the shorter time intervals allow for these differences, whereas the 12-mo interval included a mixture of culling reasons. Hence, this resulted in clearly different breeding values for LS for bulls.

2.4.1 Random Regression Parameterization

A linear RRM was used to estimate variance components for survival by fitting one random regression curve across the entire lifespan. Only a few applications of random regression for survival in cattle have been reported so far (Veerkamp et al., 2001, Schaeffer, 2004, Gengler et al., 2005, Jamrozik et al., 2013). One difference between our study and that of Schaeffer (2004) was the trait definition of the observation. We modeled the observation for survival as a record for each time interval a cow was “alive” and one record “dead” in the time interval the cow was culled, followed by missing records up to the end of the observed period (following Veerkamp et al. (2001)). Schaeffer (2004), however, proposed having the same number of records up to a certain endpoint for each uncensored cow. After culling, a cow has records coded “dead” up to a certain endpoint instead of missing records. Because of Schaeffer’s definition of the observation, (co)variances for survival can be summed using standard matrix algebra (Jamrozik et al., 1997). The latter definition is close to the survival function in the proportional hazard model, whereas our definition was close to the hazard function in the proportional hazard model. By modeling the hazard, we were able to estimate covariances between independent time intervals, which was the main aim of our study. The parameterization used by Schaeffer (2004) yields a more continuous parameterization, resulting in part-whole correlations between time intervals.

A second difference compared with the earlier studies using random regression was the length of the time intervals for survival. In the earlier studies, the trait was defined as survival per lactation, whereas we analyzed 4 different lengths for the time interval for survival. Our study showed that it is critical to define survival in smaller time intervals (e.g., 6 mo or less) to account for within-lactation variation. Random regression and multiple trait models are 2 approaches to estimate (co)variances for longitudinal survival data, and we used an RRM with a second-order Legendre polynomial. However, compared with a multiple trait model, our model does not require an a priori definition of the trait based on lactation number or year of productive life but is more flexible because a curve is fitted through the data. Therefore, strictly speaking, we did not need to partition the observations per month, and a parsimonious description with a second-order Legendre polynomial for the genetic effect for survival was shown to be sufficient because correlations between EBV of RR2, RR5, and MT were 0.99 or higher.

One change that could be made in the model for time interval of 1 mo would be to estimate residual variances for each month instead of per 3 mo. We think this could be an over-parameterization of the model but we did not test it. Another point is that we compared the order of the Legendre polynomial to test if genetic

2 Genetic analysis of longevity

parameters were comparable between different parameterizations. Therefore, we chose 2 extreme parameterizations (RR2 and MT) and 1 parameterization in the middle (RR5). Other parameterizations (e.g., RR3 and RR4) should also be investigated before choosing a model for the genetic evaluation.

We also compared computing time and number of iterations between the linear and threshold model on all trait definitions with an RRM with a second-order Legendre polynomial (Table 2.10), in addition to comparing genetic parameters. For both the linear and threshold models, computing time per iteration decreased when the time interval length increased. However, the number of iterations decreased for the linear model and increased for the threshold model. Computing time between linear and threshold models did not differ much for time intervals of 1 and 3 mo, whereas the linear model was 5 times faster than the threshold model for the time interval of 6 mo and 10 times faster for the time interval of 12 mo.

Table 2.10. Computing time in minutes and number of iterations for linear and threshold random regression models with a second-order Legendre polynomial for the genetic effect with 4 different lengths for the time interval for survival (1, 3, 6, and 12 mo)

Time interval (mo)	Linear model		Threshold model	
	Computing time	No. of iterations	Computing time	No. of iterations
1	426	23	539	14
3	514	59	344	30
6	72	10	332	37
12	64	11	666	93

2.4.2 Transformation of EBV for Monthly Survival to EBV for Lifespan

In this study, we approximated the variance of lifespan by the nonlinear profit function for estimated parameters. Brotherstone et al. (1997) and Visscher et al. (1994) derived expected lifespan and the variance of lifespan from a geometric distribution with $E(LS) = 1/(1 - p)$ and $var(LS) = p/(1 - p)^2$, where p is the survival probability from lactation n to lactation $n + 1$. Heritability for lifespan is based on a scaling of the heritability for survival per lactation with $1/(1 - p)$ and results in 3 to 5 times higher heritability for lifespan compared with survival per lactation. However, the approach suggested by Brotherstone et al. (1997) and Visscher et al. (1994) assumed all lactations or periods to be genetically the same trait with equal variances. In our study, estimated (co)variances between all months over the entire period of 72 mo were different and, using the

approximation for the variance of lifespan, we approximated the variance for lifespan across the entire period of 72 mo.

For similar model parameterizations for RR2, we estimated genetic standard deviations for LS ranging from 7.0 to 8.0 mo for the linear models and from 7.3 to 9.3 mo for the threshold models. The higher values were found with survival defined across 12-mo intervals. Using the estimated variances from our study, we can also mimic the current model, assuming genetic correlations of unity across all intervals for the 4 trait definitions for the length of the time interval, and we found genetic standard deviations in the range of 7.3 to 8.6 mo. Hence, the trend is that the more we assume that survival is the same trait across the entire life, the higher the estimated genetic variance will be. This leads to a higher suggested heritability in the range of 0.125 to 0.174 for unity genetic correlations versus 0.115 to 0.149 with nonunity genetic correlations with the 4 trait definitions for the length of the time interval. Also, applying RR5 or MT instead of RR2 to the data set with time intervals of 6 mo, with genetic correlations deviating more from unity, resulted in higher genetic standard deviations, 7.8 and 8.7 versus 7.3 mo, and higher heritabilities, 0.141 and 0.175 versus 0.125. Therefore, ignoring the nonunity genetic correlations in the calculation of LS leads to overestimation of the genetic standard deviation and reliability of breeding values. Thus, the method presented here is, in that sense, a refinement of the method described by Brotherstone et al. (1997) and Visscher et al. (1994), because correlations between time intervals not equal to unity are taken into account.

2.4.3 Genetics of Survival

In this study, the heritability of lifespan was in the range of 0.115 to 0.149 for the linear RR2 models and within the range of 0.016 to 0.181 used in different national genetic evaluations (Forabosco et al., 2009) and is within the range of 0.003 to 0.197 listed in the review of Sasaki (2013). The linear increase in total heritability for lifespan up to 36 mo in productive life is similar to the heritabilities reported by Ducrocq (2005).

Heritabilities for short time intervals for survival were very low compared with that for lifespan. This is caused by the definition of the trait, where each interval of 1, 3, 6, or 12 mo is an observation for survival. Regardless of the prediction equation, genetic or phenotypic, it is difficult to predict the probability that a cow is culled in a certain time interval. The noise-to-signal ratio is very low with short time intervals, but the ratio is higher with more time intervals (e.g., lifespan).

Genetic correlations between different time intervals suggest that survival is genetically not the same trait during the productive life of a cow, because most

2 Genetic analysis of longevity

correlations were <0.90 . In previous studies, genetic correlations of survival between different (part)lactations were in the range of 0.33 to 0.96, where lowest correlations were estimated between more distant periods (Visscher and Goddard, 1995, Brotherstone et al., 1997, Boettcher et al., 1999, Veerkamp et al., 2001, Sewalem et al., 2007, Holtsmark et al., 2009). The range of genetic correlations in our study was slightly wider because of a greater number of analyzed time points with shorter time intervals and because the analyzed period was longer. The genetic correlations indicate that survival early in productive life is a different trait than survival later in life; for example, young cows are culled more for fertility problems and low production, whereas older cows are culled more for poor udder health and poor claw health (Zijlstra et al., 2013).

2.4.4 Genetic Evaluation

The current genetic evaluation for longevity in the Netherlands uses a piecewise Weibull model implemented in the Survival Kit software (Van der Linde et al., 2007). The applied heritability is 0.12 and the genetic standard deviation is 270 d or 9 mo; in the current study, the heritability for lifespan was comparable for most parameterizations but genetic standard deviations were smaller. The Survival Kit used in the current genetic evaluation is based on the assumption that longevity is genetically the same trait over the total lifespan; in the current study, estimated genetic correlations were not unity between individual time intervals in productive life and not unity between individual time intervals and lifespan. Based on data from the first 6 mo in productive life, breeding values estimated with the Survival Kit can have reliabilities up to 99% when bulls have very large progeny groups. With results from this study, where the genetic correlation between survival up to 6 mo and the entire period of 72 mo was between 0.79 and 0.85 for the linear RR2 models, the maximum reliability would be between 62 and 72%. Therefore, reliabilities are overestimated in the current Dutch genetic evaluation, and changes in the breeding value for the entire period of 72 mo are still expected. In the current genetic evaluation, a waiting period of 9 mo to include data is applied and this is a short-term fix to reduce bias until a suitable model is developed that can treat longevity as different traits during lifespan. Moreover, it should be tested whether the estimated genetic parameters are robust enough to be applied in the national genetic evaluation. Carlén et al. (2009) concluded that a linear RRM was unstable and sensitive for the estimation of parameters for longitudinal binary observations for clinical mastitis, but it is expected to work for breeding value estimation when the genetic parameters are known. Before implementing an RRM for survival in the genetic evaluation, a formal comparison should be made with the

Weibull model using the national data set, comparing stability of breeding values over time instead of only comparing the genetic parameters. Cross-validation procedures could be used to test the stability of genetic parameters to be used for genetic evaluation.

2.5 Conclusions

Our results indicate that survival is genetically not the same trait across the entire lifespan after first calving because genetic correlations differ from unity between different time intervals in productive life, especially when intervals are further apart. Heritabilities for survival after first calving within a time interval (1, 3, 6, and 12 mo) were low. However, heritability for lifespan across the entire period of 72 mo after first calving was higher and ranged from 0.115 to 0.149. To estimate breeding values for lifespan, the trait definition of the length of the time interval for survival is important. The breeding values were more sensitive to the trait definition than to whether a linear or threshold model was used or to the order of Legendre polynomial in the model. Based on the results of this study, genetic evaluations would be improved if survival were treated as different traits during lifespan by splitting lifespan in time intervals of 6 mo or less to avoid overestimated reliabilities and changes in breeding values as daughters of bulls get older.

2.6 Acknowledgments

The authors acknowledge funding from the Dutch Dairy Board (PZ; Zoetermeer, the Netherlands), Genetic Evaluation for Sires (GES) and cattle improvement cooperative CRV (Arnhem, the Netherlands) and the Breed4Food project (program “Kennisbasis Dier,” code KB-12-006.-03-008-ASG-LR). We also thank Vincent Ducrocq (Institut National de la Recherche Agronomique, Jouy-en-Josas, France) and Mario Calus (Animal Breeding and Genomics Centre, Wageningen UR Livestock Research, Wageningen, the Netherlands) for their suggestions and comments on the manuscript.

2.7 References

- Boettcher, P. J., L. K. Jairath, and J. C. W. Dekkers. 1999. Comparison of methods for genetic evaluation of sires for survival of their daughters in the first three lactations. *J. Dairy Sci.* 82:1034-1044.
- Brotherstone, S., R. F. Veerkamp, and W. G. Hill. 1997. Genetic parameters for a simple predictor of the lifespan of Holstein-Friesian dairy cattle and its relationship to production. *Anim. Sci.* 65:31-37.
- Carlén, E., U. Emanuelson, and E. Strandberg. 2006. Genetic evaluation of mastitis in dairy cattle using linear models, threshold models, and survival analysis: A simulation study. *J. Dairy Sci.* 89:4049-4057.
- Carlén, E., K. Grandinson, U. Emanuelson, and E. Strandberg. 2009. Random regression models for genetic evaluation of clinical mastitis in dairy cattle. *Animal* 3:1100-1108.
- De Vries, A., J. D. Olson, and P. J. Pinedo. 2010. Reproductive risk factors for culling and productive life in large dairy herds in the eastern United States between 2001 and 2006. *J. Dairy Sci.* 93:613-623.
- Ducrocq, V. 2005. An improved model for the French genetic evaluation of dairy bulls on length of productive life of their daughters. *Anim. Sci.* 80:249-256.
- Ducrocq, V. and J. Sölkner. 1994. "The Survival Kit", a FORTRAN package for the analysis of survival data. Pages 51-52 in Proc. Proc. 5th World Congr. Genet. Appl. Livest. Prod., Guelph, Ontario, Canada.
- Ducrocq, V. and J. Sölkner. 1998. The Survival Kit—V3.0: A Package for Large Analyses of Survival Data. Pages 447–450 in Proc. Proc. 6th World Congr. Genet. Appl. Livest. Prod., Armidale, Australia.
- Ducrocq, V., J. Sölkner, and G. Mészáros. 2010. Survival Kit v6 - a software package for survival analysis. in Proc. Proc. 9th World Congr. Genet. Appl. Livest. Prod., Leipzig, Germany.
- Dürr, J. W., H. G. Monardes, and R. I. Cue. 1999. Genetic analysis of herd life in Quebec Holsteins using Weibull models. *J. Dairy Sci.* 82:2503-2513.
- Dürr, J. W., H. G. Monardes, R. I. Cue, and J. C. Philpot. 1997. Culling in Quebec Holstein herds. 1. Study of phenotypic trends in herd life. *Can. J. Anim. Sci.* 77:593-600.
- Fischer, T. M., A. R. Gilmour, and J. H. J. van der Werf. 2004. Computing approximate standard errors for genetic parameters derived from random regression models fitted by average information REML. *Genet. Sel. Evol.* 36:363-369.

- Forabosco, F., J. H. Jakobsen, and W. F. Fikse. 2009. International genetic evaluation for direct longevity in dairy bulls. *J. Dairy Sci.* 92:2338-2347.
- Gengler, N., S. Vanderick, P. Mayeres, A. Gillon, and C. Croquet. 2005. Genetic evaluation of cow survival using a lactation random regression model. *Interbull Bulletin* 33:176.
- Gilmour, A. R., B. Gogel, B. Cullis, R. Thompson, D. Butler, M. Cherry, D. Collins, G. Dutkowski, S. Harding, and K. Haskard. 2009. ASReml user guide release 3.0. VSN International Ltd, Hemel Hempstead, UK.
- Guo, S. F., D. Gianola, R. Rekaya, and T. Short. 2001. Bayesian analysis of lifetime performance and prolificacy in Landrace sows using a linear mixed model with censoring. *Livest. Prod. Sci.* 72:243-252.
- Holtmark, M., B. Heringstad, and J. Ødegård. 2009. Predictive abilities of different statistical models for analysis of survival data in dairy cattle. *J. Dairy Sci.* 92:5730-5738.
- Jamrozik, J., J. Fatehi, and L. R. Schaeffer. 2008. Comparison of models for genetic evaluation of survival traits in dairy cattle: a simulation study. *J. Anim. Breed. Genet.* 125:75-83.
- Jamrozik, J., S. McGrath, R. A. Kemp, and S. P. Miller. 2013. Estimates of genetic parameters for stayability to consecutive calvings of Canadian Simmentals by random regression models. *J. Anim. Sci.* 91:3634-3643.
- Jamrozik, J., L. R. Schaeffer, and J. C. M. Dekkers. 1997. Genetic evaluation of dairy cattle using test day yields and random regression model. *J. Dairy Sci.* 80:1217-1226.
- Meijering, A. and D. Gianola. 1985. Linear versus nonlinear methods of sire evaluation for categorical traits - a simulation study. *Genet. Sel. Evol.* 17:115-131.
- Mészáros, G., J. Sölkner, and V. Ducrocq. 2013. The Survival Kit: Software to analyze survival data including possibly correlated random effects. *Computer methods and programs in biomedicine* 110:503-510.
- Ødegård, J., I. Olesen, B. Gjerde, and G. Klemetsdal. 2006. Evaluation of statistical models for genetic analysis of challenge test data on furunculosis resistance in Atlantic salmon (*Salmo salar*): Prediction of field survival. *Aquaculture* 259:116-123.
- Pritchard, T., M. Coffey, R. Mrode, and E. Wall. 2013. Understanding the genetics of survival in dairy cows. *J. Dairy Sci.* 96:3296-3309.
- Rendel, J. M. and A. Robertson. 1950. Some aspects of longevity in dairy cows. *Empire Journal of Experimental Agriculture* 18:49-56.
- Robertson, A. 1966. A mathematical model of culling process in dairy cattle. *Animal Production* 8:95-108.

2 Genetic analysis of longevity

- Roxström, A., V. Ducrocq, and E. Strandberg. 2003. Survival analysis of longevity in dairy cattle on a lactation basis. *Genet. Sel. Evol.* 35:305-318.
- Roxström, A. and E. Strandberg. 2002. Genetic analysis of functional, fertility-, mastitis-, and production-determined length of productive life in Swedish dairy cattle. *Livest. Prod. Sci.* 74:125-135.
- Sasaki, O. 2013. Estimation of genetic parameters for longevity traits in dairy cattle: A review with focus on the characteristics of analytical models. *Anim. Sci. J.* 84:449-460.
- Sasaki, O., M. Aihara, K. Hagiya, A. Nishiura, K. Ishii, and M. Satoh. 2012. Genetic evaluation of the longevity of the Holstein population in Japan using a Weibull proportional hazard model. *Anim. Sci. J.* 83:95-102.
- Schaeffer, L. R. 2004. Application of random regression models in animal breeding. *Livest. Prod. Sci.* 86:35-45.
- Schaeffer, L. R. and J. C. M. Dekkers. 1994. Random regressions in animal models for test-day production in dairy cattle. Pages 443-446 in *Proc. 5th World Congr. Genet. Appl. Livest. Prod., Guelph, Ontario, Canada.*
- Sewalem, A., F. Miglior, G. J. Kistemaker, P. Sullivan, G. Huapaya, and B. J. Van Doormaal. 2007. Short communication: Modification of genetic evaluation of herd life from a three-trait to a five-trait model in Canadian dairy cattle. *J. Dairy Sci.* 90:2025-2028.
- Terawaki, Y. and V. Ducrocq. 2009. Nongenetic effects and genetic parameters for length of productive life of Holstein cows in Hokkaido, Japan. *J. Dairy Sci.* 92:2144-2150.
- Van Arendonk, J. A. M. 1985. Studies on the replacement policies in dairy cattle. II. Optimum policy and influence of changes in production and prices. *Livest. Prod. Sci.* 13:101-121.
- Van der Linde, C., A. Harbers, and G. de Jong. 2007. From functional to productive longevity in the Netherlands. *Interbull Bulletin* 37:203.
- Veerkamp, R. F., S. Brotherstone, B. Engel, and T. H. E. Meuwissen. 2001. Analysis of censored survival data using random regression models. *Anim. Sci.* 72:1-10.
- Visscher, P. M., P. J. Bowman, and M. E. Goddard. 1994. Breeding objectives for pasture based dairy production systems. *Livest. Prod. Sci.* 40:123-137.
- Visscher, P. M. and M. E. Goddard. 1995. Genetic parameters for milk yield, survival, workability, and type traits for Australian dairy cattle. *J. Dairy Sci.* 78:205-220.
- Zijlstra, J., M. Boer, J. Buiting, K. Colombijn-van der Wende, and E.-A. Andringa. 2013. Rapport 668: Routekaart Levensduur; Eindrapportage van het project

"Verlenging levensduur melkvee". Wageningen UR Livestock Research, Wageningen, the Netherlands.

3

Changes in the genetic level and the effects of age at first calving and milk production on survival during the first lactation over the last 25 years

M.L. van Pelt^{1,2}, G. de Jong¹, R.F. Veerkamp²

¹ CRV BV, Animal Evaluation Unit, PO Box 454, 6800 AL Arnhem, the Netherlands;

² Animal Breeding and Genomics Centre, Wageningen UR Livestock Research, PO Box 338, 6700 AH Wageningen, the Netherlands

The Animal Consortium (2016) 10:2043-2050
DOI: <https://doi.org/10.1017/S1751731116001282>

Abstract

Survival during the first year after first calving was investigated over the last 25 years, 1989–2013, as well as how the association of survival with season of calving, age at first calving (AFC) and within-herd production level has changed over that period. The data set contained 1 108 745 Dutch black-and-white cows in 2185 herds. Linear models were used to estimate (1) effect of year and season and their interaction and (2) effect of AFC, within-herd production level, and 5-year intervals and their two-way interactions, and the genetic trend. All models contained AFC and percentage of Holstein Friesian as a fixed effect, and herd-year-season, sire and maternal grandsire as random effects. Survival and functional survival were analysed. Functional survival was defined as survival adjusted for within-herd production level. Survival rate increased by 8% up to 92% in the last 25 years. When accounting for pedigree, survival showed no improvement up to 1999, but improved since then. Genetically, survival increased 3% to 4% but functional survival did not increase over the 25 years. We found an interesting difference between the genetic trends for survival and functional survival for bulls born between 1985 and 1999, where the trend for survival was still increasing, but was negative for functional survival. Since 1999, genetic trend picked up again for both survival and functional survival. AFC, season of calving and within-herd production level affected survival. Survival rate decreased 0.6%/month for survival and 1.5% for functional survival between AFC of 24 and 32 months. Calving in summer resulted in 2.0% higher survival than calving in winter. Within herd, low-producing cows had a lower survival rate than high-producing cows. However, these effects became less important during the recent years. Based on survival optimum AFC is around 24 months, but based on functional survival it is better to have an AFC < 24 months. Overall, survival rate of heifers has improved considerably in the past 25 years, initially due to the focus on a high milk production. More recently, the importance of a high milk production has been reduced towards attention for functional survival.

Key words: dairy cattle, longevity, survival, age at first calving, within-herd production level

3.1 Implications

Survival rate in the first year after calving increased from 8% to 92% between 1989 and 2013. Genetic selection made a positive contribution of 4% for survival, whereas functional survival – adjusted for within-herd production level – declined until 1999 and is since then increasing again. Culling risk increased with older age at first calving (**AFC**) and decreased with higher production within herd. However, over the years, the effect of AFC and production level on survival in first lactation has reduced significantly.

3.2 Introduction

The dairy industry has undergone profound changes in recent decades, that potentially affect the productivity, health and welfare of dairy cows, for example, herd size, use of hired labour, housing system, milk price and use of new technology (Barkema et al., 2015). At the same time, milk production per cow has more than doubled in the previous 40 years, and till the end of the last century single trait selection dominated breeding programmes. All these changes over time have stimulated the discussion in Western Europe about the effects of these changes on health and welfare and the underlying lifespan of cows. For example, as discussed in more detail by Veerkamp et al. (2008), there is clear evidence that genetic selection solely for milk yield has negative consequences for health and fertility, but it is not the absolute milk yield that apparently created the problems (Weigel, 2006, Windig et al., 2006). Also, studies on the trend in longevity over the past decades vary from the opinion that the effect of larger farms and (selection for) higher milk production have decreased the survival rate of dairy cows (Oltenacu and Broom, 2010, Froidmont et al., 2013), to the opinion that improved management and multi-trait genetic selection have had a positive impact on the survival of dairy cows (Dematawewa and Berger, 1998, Dechow and Goodling, 2008, Miglior et al., 2012). But, apart from the conflicting literature, there is little insight into how longevity has changed over the past decades, and the most important factors that play a role in the culling decisions of dairy farmers.

Culling decisions are likely to be affected by changes in national regulation (e.g. milk quota system), legislation, feed costs, milk price and revenues for culled animals, and therefore culling reasons are likely to have changed over the years or seasons as well. In Europe, a milk quota system was in place from April 1984 until March 2015 in order to limit the amount of milk produced annually per country, but also per farmer (Bergevoet et al., 2004). Individual farmers that produced more than the allowed quota, had to pay a penalty for the excess of milk they produced. This system might have affected culling reasons towards the end of the quota year

(31 March), because farmers might have decided to cull more cows in order to avoid the penalty. How such a national regulation affected the culling, on a seasonal basis for example, has not yet been investigated.

In the literature, it has been suggested that animals calving in the optimal range of 22 to 26 months for AFC achieved highest lifetime milk yield and longest productive life, and resulted in highest profit per cow (Nilforooshan and Edriss, 2004, Froidmont et al., 2013, Wathes et al., 2014). But within-herd production level also affects survival. Cows are culled on both voluntary and involuntary bases; the voluntary case is, for example, when the farmer decides to remove a healthy cow for low production, and the involuntary case is, for example, when the farmer is forced to remove the cow because of poor health, injury or poor fertility. With a high level of involuntary culling, a farmer has less opportunity to cull low-producing cows voluntarily. In the past, voluntary culling was assumed to be mainly for milk yield, and there is evidence that low-producing cows are more likely to be culled than high-producing cows (Vollema et al., 2000, Sewalem et al., 2005, Terawaki and Ducrocq, 2009). However, in expanding herds in Wisconsin, for low-producing cows the relative risk for culling decreased during the period 1981–2000, whereas the relative risk increased for high-producing cows (Weigel et al., 2003). Hence, in Western Europe it is unclear what the association is between production level, AFC and survival, and whether these associations have changed over decades. For example, breeding goals have changed over the past 15 years, from selection mainly on yield towards selection related to more extensive breeding goals including health, fertility and longevity (Miglior et al., 2005, Miglior et al., 2012). Therefore, the objective of this study was to investigate cow survival in the Netherlands for the period 1989–2013, and whether the associations of survival with season of calving, genetic level of survival, AFC and within-herd production level have changed over these decades.

3.3 Material and methods

3.3.1 Data

For this study, survival after first calving was defined: (1) survival until month 12 (**surv_12mo**), that is, a cow was considered to have survived until month 12 if she did not die and was not culled for slaughter and (2) survival to parity 2 (**surv_1st_lac**), that is, a cow was considered to have survived her first lactation if she had initiated her second lactation by having at least 1 test day belonging to the second lactation.

3 Changes of survival over the last 25 years

We have chosen to analyse survival in early life instead of total lifespan, because `surv_12mo` had a high genetic correlation (~ 0.85) with survival up to 72 months (Van Pelt et al., 2015), information for `surv_12mo` and `surv_1st_lac` is rather more rapidly available for each animal compared with lifespan, and more importantly modelling of the data is less complex. With only first lactation animals without censored information, all animals within a calving year can be compared, instead of comparing a mixture of different age groups at the same time within a herd.

Survival was coded as 1 for animals that survived `surv_12mo/surv_1st_lac`, and as 0 for animals that died or were culled for slaughter. Data were available from the Dutch/Flemish cattle improvement cooperative (CRV) from 1989 to 2013. Records for survival were constructed from the national movement database considering herdbook-registered cows. Compared to using milk recording information, the use of the movement database allowed accurate differentiating between animals that died, that were slaughtered, that were exported or that moved to another herd. All animals were a combination of at least 87.5% Holstein Friesian and Dutch Friesian, and AFC was between 21 and 40 months. Herds with at least 95% Holstein Friesian and Dutch Friesian genes were selected. Animals were required to have the first observation in month 1 in parity 1; that is, left-censored animals were deleted, because those animals had missing milk production data or changed herds. Data were created for all cows at herd level, containing all milking cows in all age classes. Only herds with at least 30 cows present every month, in the period 1994–2013, were selected in order to exclude herds with a non-conventional culling management, because, for example, the farm stopped operating or the entire herd was culled at once because of an outbreak of a veterinary disease. Due to the exclusion of left-censored animals, selection of herds was not based on the entire period, but started in 1994. Additional requirements were that (1) sires had (i.e. as sire or as maternal grandsire or combined) at least 15 progeny that could have been productive for at least 12 months after first calving, (2) sires had progeny in at least two herds and (3) every herd-year-month class had at least 25 observations. These three additional requirements for progeny per sire, herds per sire and observations per herd-year-month had to be repeated 11 times until the final data set met all three criteria. The creation of the final data set with `surv_12mo` and `surv_1st_lac` was as follows: for `surv_12mo` animals were selected that could have been in the herd for at least 12 months after first calving, that is, animals calving after 31 December 2013 were excluded. The data set with `surv_12mo` contained 1 108 745 animals. For `surv_1st_lac` animals were selected that could have been in the herd for at least 18 months after first calving, that is, animals calving after 30 June 2013 were excluded. The data set with `surv_1st_lac` contained 1 062 276

animals. In both data sets 2185 herds were included. Pedigree information of the sires and maternal grandsires was traced back six generations, resulting in a pedigree file with 11 268 sires.

3.3.2 Statistical Model

Two analyses were performed to test the effect on survival for (1) year and season, (2) AFC, within-herd production level, 5-year interval and to estimate the genetic trend. The following base model was used:

$$y = \mu + \text{FIXED} + hys + \left(\text{sire} + \frac{1}{2}mgs\right) + e$$

where y is the observation for `surv_12mo` or `surv_1st_lac`, μ the overall mean, hys the random effect of herd-year-season of first calving, $sire$ the random effect for sire, mgs the random effect for maternal grandsire and e the residual. The random sire effects were fitted by overlaying the relationship matrix for sire and 1/2 times the maternal grandsire matrix, resulting in one estimate per effect for a sire. Both analyses contained fixed effects for AFC (15 classes: ≤ 21 , 22, ..., 34, ≥ 35 months) and percentage of Holstein Friesian genes (five classes: $\leq 50.0\%$, 62.5%, 75.0%, 87.5% and 100%). The other fixed effects differed per analysis.

The first analysis contained fixed effects for year, season and the interaction between year and season; year was year of first calving (1989–2013) and season was season of first calving (winter, 1 January to 31 March; spring, 1 April to 30 June; summer, 1 July to 30 September; autumn, 1 October to 31 December).

The second analysis contained fixed effects for 5-year interval (5-year intervals of first calving were 1989–93, till 2009–13), and the interaction between 5-year interval and AFC. This model was used to analyse survival. The same model was used to analyse functional survival, where survival is adjusted for individual milk production relative to the production level of the herd, and it is suggested as a way to exclude voluntary culling from the breeding value (Robertson, 1966). Therefore, for functional survival the fixed effect for within-herd production level was included, together with the interactions between 5-year interval and within-herd production level, and between AFC and within-herd production level. Within-herd production level was a ranking of animals within a herd by 5-year interval for predicted or realized 305-day yield of combined kg fat and protein, and animals were ranked into seven classes from worst to best, with 1: 1% to 5%, 2: 6% to 20%, 3: 21% to 40%, 4: 41% to 60%, 5: 61% to 80%, 6: 81% to 95% and 7: 96% to 100%. The genetic trend for bulls was obtained by averaging estimated breeding values

3 Changes of survival over the last 25 years

(EBV) by birth year for bulls that were at least 87.5% Holstein Friesian. The genetic trend for cows was approximated by averaging for each animal 1/2 EBV of sire plus 1/4 EBV of maternal grandsire, as a sire model was used no EBV for the maternal grandam were available.

Effects were estimated with ASReml (Gilmour et al., 2009). Subsequently, in order to compare levels within an effect, corrected for all other effects in the model, y -values were predicted using least squares means of the effects with the PREDICT statement in ASReml (Gilmour et al., 2009). Predicted y -values were tested for significance based on the t -statistic. Fixed effects in the models were tested for significance based on the F -test ($P < 0.01$).

3.4 Results

In a 25-year period, the analysed farms showed changes in mean survival rate, but also in herd size and milk production figures (Table 3.1). When we compare 1989–93 with 2009–13, herd size increased by 48%, percentage Holstein Friesian increased from 80% to nearly 100%, AFC remained stable around 26 months, lactation length increased by 13%, and 305-day production increased by 18% for kg milk, 11% for kg fat and 18% for kg protein. With the increase in herd size and production, $surv_{12mo}$ increased by 7.3% and $surv_{1st_lac}$ increased by 7.1%. Overall, the results were comparable for $surv_{12mo}$ and $surv_{1st_lac}$, therefore only results of $surv_{12mo}$ are presented here.

Table 3.1. Number of cows, number of herds, means for percentage of Holstein Friesian (HF), survival rate at 12 months after first calving ($surv_{12mo}$), survival rate of first lactation ($surv_{1st_lac}$), age at first calving (AFC), lactation length, 305-day yield for milk, fat and protein, divided into five 5-year intervals.

5-year interval	1988-93	1994-98	1999–2003	2004-08	2009-13
Cows (n)	175 822	209 941	226 278	236 362	260 342
Herds (n)	2185	2185	2185	2185	2185
HF (%)	80.2	91.7	98.0	99.4	99.4
$Surv_{12mo}$ (%)	83.4	84.9	88.0	90.3	90.7
$Surv_{1st_lac}$ (%)	79.5	80.1	83.0	85.8	86.6
AFC (mo)	25.8	26.0	25.9	25.8	25.7
Lactation length	309	322	339	350	348
305-day milk yield (kg)	6664	7266	7541	7796	7883
305-day fat yield (kg)	299	315	326	331	332
305-day protein yield (kg)	231	250	259	268	272

Raw and predicted means for surv_12mo are presented per calving year in Figure 3.1. Mean survival rate increased phenotypically between 1989 and 2013 by about 8%. Survival rates fluctuated between years, and was lowest in 1991 (82.0%) and highest in 2007 (91.8%). When survival was predicted as the least squares means in the model with effects for year, season and year-season, then there was less variation between years and there was no improvement of survival up to 1999, in contrast to the trend observed in the raw means. The effect in the model that caused this shift in curve was by accounting for the pedigree in the model. We tested this by excluding the pedigree from the model. Not accounting for genetic covariance between records led to a too optimistic trend for survival up to 1999, and improvement between 1991 and 2007 was reduced to 6.5% when accounting for pedigree. The direct effect of Holsteinization was relatively small. Compared with a model without accounting for percentage Holstein Friesian, a model including percentage Holstein Friesian gave predicted means for survival that were lower, the maximum difference was 0.5% between 1989 and 2013, the years where the mean percentage Holstein Friesian differed the most.

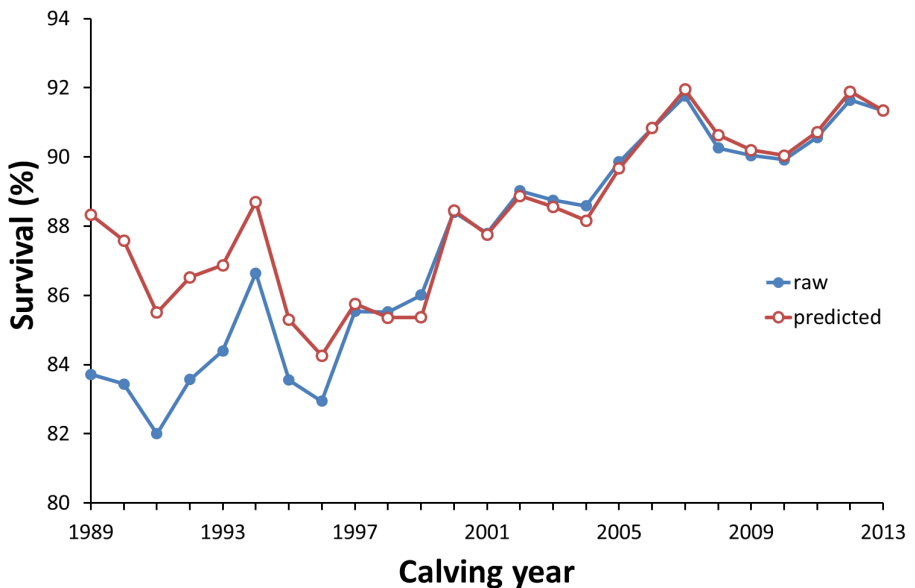


Figure 3.1. Raw and predicted means for survival at 12 months for calving year (reference: 2006).

3 Changes of survival over the last 25 years

Cows that calved in winter had the lowest survival with 86.8%. Compared with calving in winter, survival was 0.5% higher in spring, 2.0% higher in summer and 0.9% higher in autumn. Within a year, the difference in predicted means for survival between cows calving in summer and winter was 2% or more in 1989–2001, and the largest difference (–5.6%) in this period was between summer 1994 and winter 1995 (Figure 3.2). In later decades, this difference between summer and winter became much smaller and the difference in predicted means for survival between summer and winter was 1.2% or less in 2009–13.

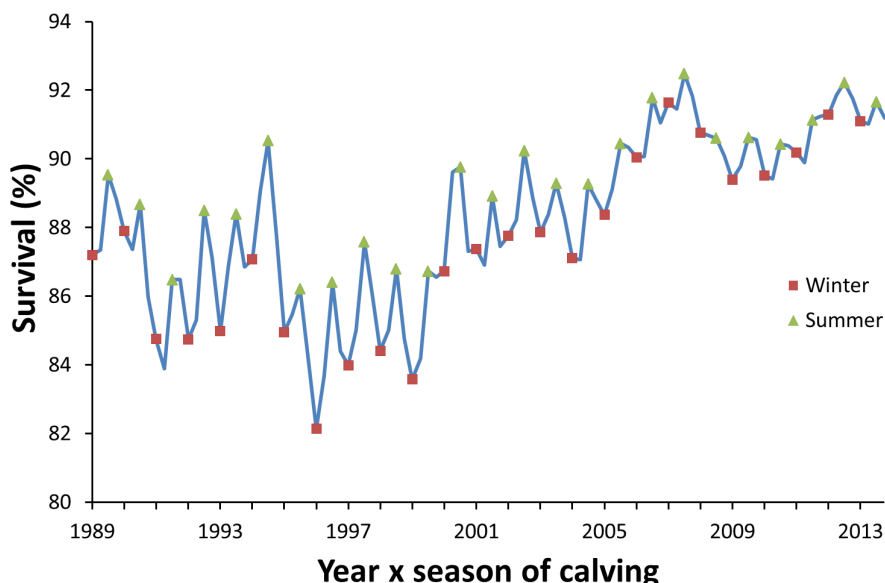


Figure 3.2. Predicted means for survival at 12 months for year x season of calving (reference: winter 2006, winters are marked by squares and summers are marked by triangles).

In the raw data, survival showed a clear optimum of 89% at AFC of 23 to 24 months and cows calving at a younger or older age had a lower survival (Figure 3.3). For predicted means for survival, that is, not adjusted for within-herd production level, the same pattern was observed as for the raw means. However, for functional survival, that is, survival adjusted for within-herd production level, no optimum was observed for survival at 24 months and AFC < 24 months resulted in higher survival rates. This suggests that calving at a younger age resulted in a higher survival rate as long as production level was not decreased. Calving at a higher AFC resulted always in a decline in survival rate, especially when production level did not improve. The average decrease in survival rate per extra month AFC was 0.6% for

survival and 1.5% for functional survival between AFC of 24 and 32 months. Also over the last decades the effect of AFC on functional survival changed, not only at the level of the intercept, as expected because mean survival increased over the years (Figure 3.1), but the effect of AFC on survival also reduced (Figure 3.4). The difference in survival rate between AFC of 24 and 32 months declined from 14.8% in 1989–93 to 10.0% in 2009–13.

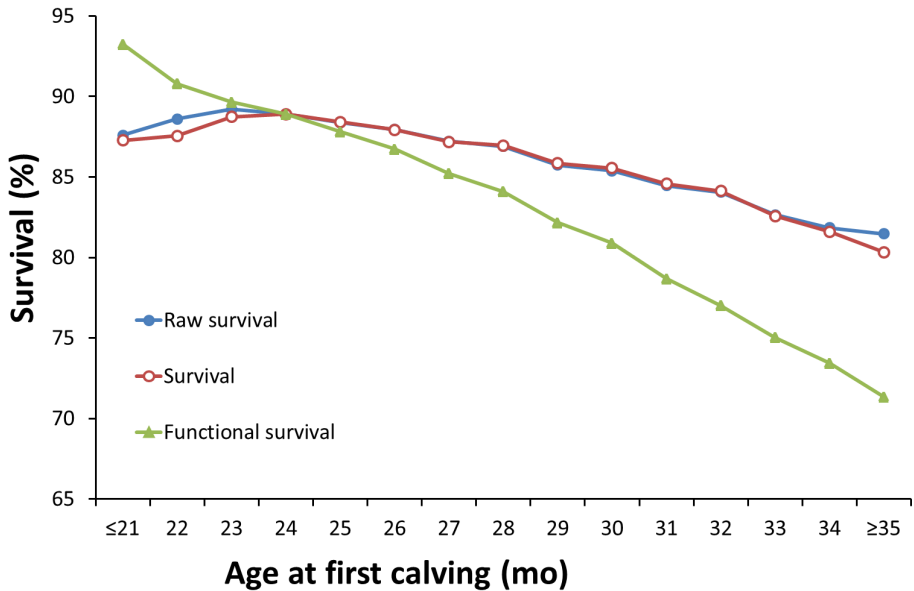


Figure 3.3. Raw means for survival and predicted means for (functional) survival at 12 months for age at first calving (reference: 24 months).

3 Changes of survival over the last 25 years

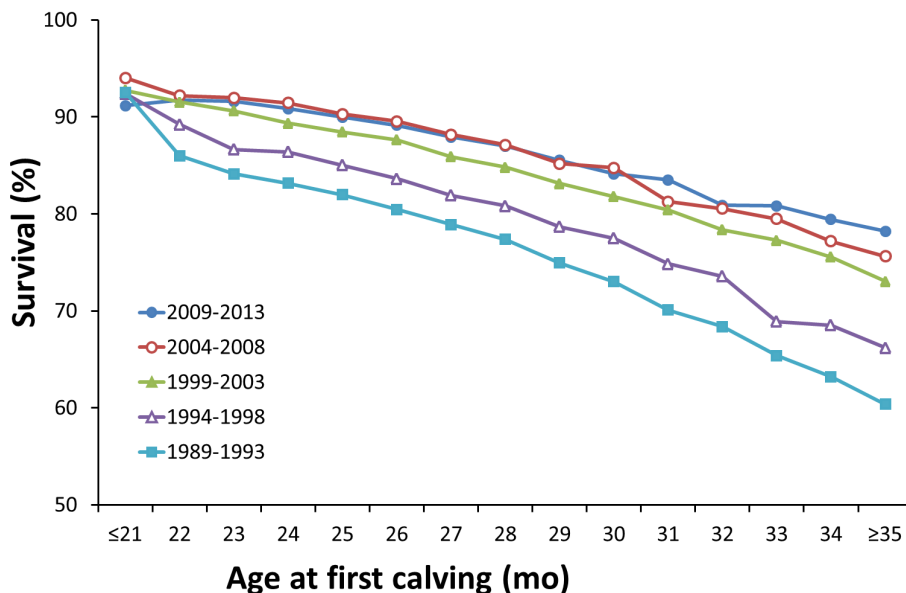


Figure 3.4. Predicted means for survival at 12 months for age at first calving (reference: 24 months, 2004–08).

Similar to the effect of AFC on survival, the effect of within-herd production level on survival changed over the past decades (Figure 3.5). In 1989–93, the difference in survival rate between the lowest (1% to 5%) and average (41% to 60%) within-herd production level was 43.6%, and this difference decreased to 28.6% in 2009–13. In addition, the difference in survival rate between the highest (96% to 100%) and average production level decreased from 12.4% in 1989–93 to 7.4% in 2009–13. Only 41.8% of the lowest-producing cows were surviving the first year in 1989–93, and in 2009–13 this survival rate increased to 64.3%. Hence, these results suggest that the within-herd production level became less important for culling over the past decades. The effect of the interaction between AFC and production level on functional survival (Figure 3.6) showed that for the high-producing cows, AFC did not influence the survival rate as much as for the low-producing cows. Cows producing below herd-average already had a lowered survival rate, but when they also calved at an older age, 24 v. 32 months, the survival rate decreased faster (–27.1%) compared with high-producing herd mates (–3.1%).

3 Changes of survival over the last 25 years

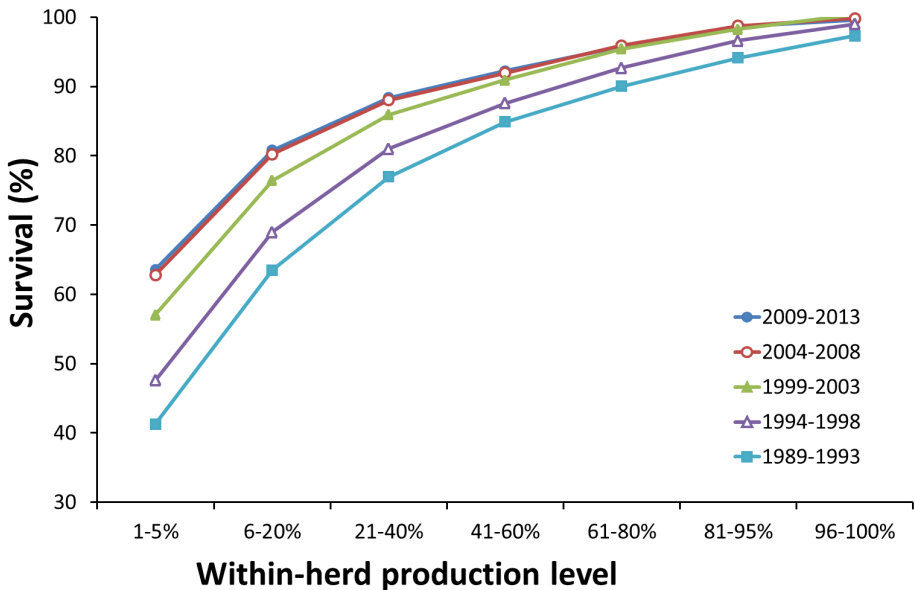


Figure 3.5. Predicted means for survival at 12 months for within-herd production level per 5-year interval (reference level: 41% to 60%, 2004–08). Class 96% to 100% is the highest within-herd production level.

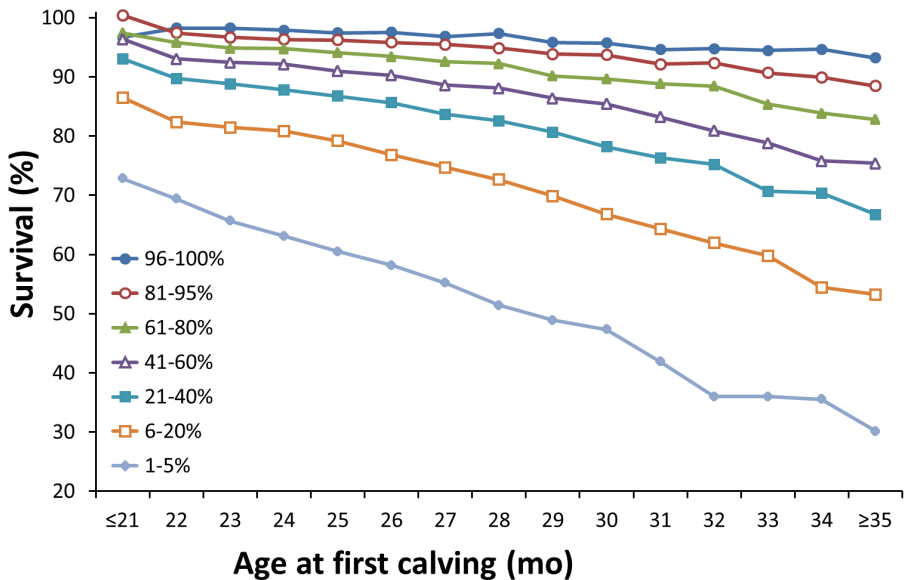


Figure 3.6. Predicted means for survival at 12 months for age at first calving per within-herd production level (reference: 24 months, level 41% to 60%). Class 96% to 100% is the highest within-herd production level.

3 Changes of survival over the last 25 years

Genetic levels for survival increased over the 25-year period with 3.7% for bulls and 2.9% for cows, and functional survival did not increase for bulls and declined 0.4% for cows (Figure 3.7). However, there was an interesting difference between survival and functional survival of bulls born between 1985 and 1999. The genetic trend for survival increased (0.16%/year), but genetic progress for functional survival declined over this period (-0.10%/year). Initially, the genetic improvement of the bulls came due to the higher production of their daughters within herd, rather than a better functional survival per se. For survival and functional survival the genetic trend picked up again for the bulls after 1999, and both increased 0.15%/year. Also for cows we see a positive genetic trend since that time albeit lower than the trend for bulls.

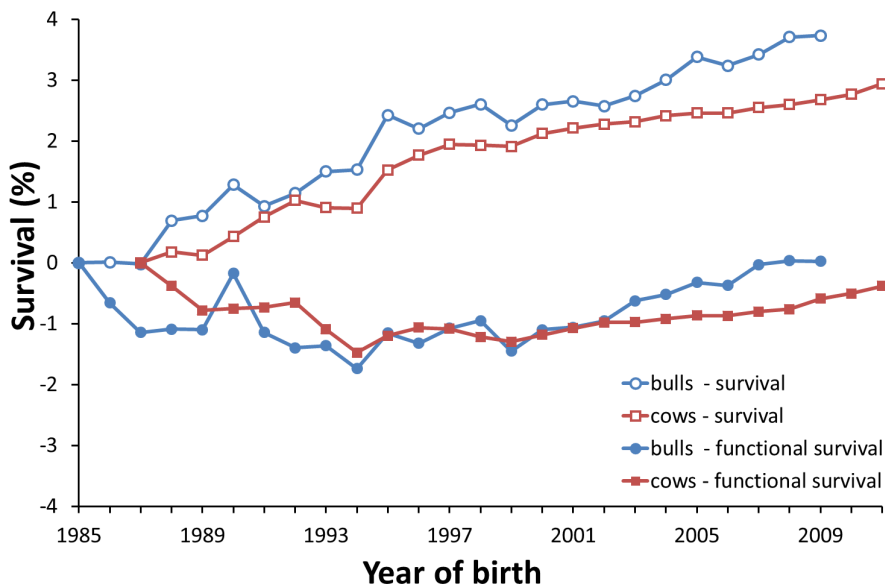


Figure 3.7. Estimates of genetic trend for (functional) survival for Holstein bulls and cows (reference: bulls 1985, cows 1987).

3.5 Discussion

The objective of this study was to investigate whether cow survival changed between 1989 and 2013, and whether the associations of survival with year and season of calving, AFC, within-herd production level and genetics have changed over these decades. Analysis of a unique data set collected over the past 25 years, which includes movement records to decide if an animal was really culled or just sold to another farm, demonstrated clear phenotypic and genetic trends over time,

but also how trends are affected by taking into account the pedigree or not, or by adjusting for milk yield or not. Furthermore, it was demonstrated how effects of AFC and within-herd milk production interact and changed over time. Although results are based on Dutch dairy cows, results might be applicable to other countries in Western Europe.

To enable proper modelling, and to facilitate the investigation of the effect of milk yield on survival, this study was limited to survival during first lactation only. Obviously, survival in the first year is a prerequisite for a longer lifespan of an animal. However, it is more common to analyse lifespan, and culling during first lactation differs from culling patterns in later lactations (Boettcher et al., 1999, Sewalem et al., 2007, Van Pelt et al., 2015). Notwithstanding, we think we can compare these results with full lifespan because *surv_12mo* had a high genetic correlation (~ 0.85) with survival up to 72 months, and a similar change as we observed for survival in the first year in genetic trend was observed in the national evaluation for total lifespan, both for bulls and cows (CRV, 2015a and 2015b). Numerically, it is possible to extrapolate survival rate till 12 months to lifespan at 72 months, with some simple assumptions on survival rate during later parities. Van Pelt et al. (2015) found survival rates (mean across all years in a subset of this data set) for each 12-month period up to 72 months of 0.88, 0.83, 0.77, 0.70, 0.66 and 0.61, and when they were extrapolated they gave an average lifespan of 3.00 years. If we assume that only survival in the first 12 months improved over the past decades, then we expect that a change in survival rate from 84% to 92% at 12 months, which is about the change between intervals 1989–93 and 2009–13, would have increased lifespan by 0.27 years from 2.87 to 3.14 years. However, it could be hypothesized when survival rate in the first year has improved, that survival in later years has improved as well, and the impact on lifespan would be greater. Assuming this improved survival rate across whole lifespan, then lifespan is expected to be improved by 0.79 years from 2.64 to 3.43 years over the last decades. Thus, probably the increase in lifespan over the past 25 years that is predicted from the results in this study is in the range of 0.27 to 0.79 years. Comparing this prediction with the Dutch cow population, lifespan improved by 0.60 years for cows culled during the period 2000–08 (CRV, 2015a).

Only farms that were in business during the entire period (1989–2013) were selected. The advantage was that the effects of year, AFC and within-herd production level could be evaluated across the same farms. In the analysed years, the number of dairy farms in a milk recording scheme decreased in the Netherlands from 30 000 to 15 000 and the number of cows per farm almost doubled from 46 to 87 (CRV, 2015a). The number of cows also increased in the farms that we

3 Changes of survival over the last 25 years

selected, albeit that this increase was only 48% (Table 3.1). Therefore, our selection of farms might not be a precise reflection for the whole dairy industry, but the selected farms stayed in business and probably reflect the best sample of farms to evaluate changes in survival rate over these decades.

The peak in survival rate in calving year 2007 coincided with a high milk price in 2008 in the Netherlands (Jongeneel and Van Berkum, 2015). Hence, milk price and economic circumstances were important factors affecting the survival rate of cows and might also explain that heifers calving in summer had on average a 2.0% higher survival than heifers calving in winter. In the Netherlands, dairy factories gave a premium on the milk price for milk produced in fall and winter, and a penalty in spring and summer. Therefore, farmers were likely to be more tolerant in culling their summer calving heifers. However, the opposite results for the seasonal effect were found across countries. In the United States, survival rate for Holstein Friesian was higher for heifers calving in summer and fall (Hadley et al., 2006). In Wallonia, Belgium, negligible effects of season of calving were found on herd life (Froidmont et al., 2013), whereas in Spain survival rate was higher for heifers calving in winter and spring (Bach, 2011). These opposite results for season could be due to climatic differences, where cows in hot climates had impaired reproduction and also had lower survival rates in the hotter summer months (Vitali et al., 2009). Furthermore, national regulations, like the milk quota system, might have affected seasonal differences. Autumn and winter calving heifers that were not pregnant at the end of the quota year, that is, before 1 April, had a higher risk of being culled compared with summer calving heifers, because the farmer had to decide which cows to cull and especially when exceeding the milk quota. Over the years though, seasonal differences in survival rate became smaller (Figure 3.2), and it could be argued that farmers were better able to handle the quota system. In the beginning of the quota system, farmers probably culled animals more drastically to avoid exceeding their quota, and over time they became better in planning the introduction of their young stock to the milking herd and consequently better in planning the culling of animals.

The phenotypic increase for survival rate observed in the current study was not consistently found in all countries. In the United States, the phenotypic trend for productive herd life was negative for the past decades (Nieuwhof et al., 1989, Hare et al., 2006), apparently because of more intense culling primarily due to management decisions rather than genetics (Hare et al., 2006). Oltenacu and Algers (2005) reported that in the United States the proportion of cows still alive at 48 months of age decreased from 80% to 60% for the period 1957–2002. However, survival to second parity stabilized, and similar to our results showed an increase

for all analysed breeds since 1996 (Hare et al., 2006). In Austria, the average herd life decreased by 0.5 years to 3.5 years for the period 1990–2005 (Fürst and Fürst-Waltl, 2006). The differences between these countries might be partly explained by the introduction of genetic evaluations for longevity since the mid-1990s, for example, in the United Kingdom (Veerkamp et al., 1995) and in the Netherlands (Vollema et al., 2000). Normally, more awareness among farmers arises, AI (artificial insemination) companies excluded inferior bulls and breeding goals change when a genetic evaluation for a new trait is introduced, leading to a change in genetic trend.

In 1999, the genetic evaluation for functional longevity was introduced, and was changed to longevity in 2008 (Van der Linde et al., 2007). Before the introduction of the genetic evaluation there was a strong genetic trend for survival, that was reduced after introduction of the genetic evaluation. However, for functional survival we saw a decline in genetic trend before introduction of the genetic evaluation. After introduction of the genetic evaluation for functional survival a positive trend was observed, which is in line with expectation. The initially strong genetic progress for survival was probably due to a narrow breeding goal with strong emphasis on milk production, and the strong emphasis on increasing yield through heifer selection. That there was no accompanying genetic trend for functional survival during those years is probably due to the well-known association between selection for milk yield only and negative correlated responses for health and fertility (Pryce et al., 1997). Thus, genetic progress was due to the fact that sires that inherited a high milk production were used more often and their daughters were favoured during first lactation. This imbalance in the data is adjusted for by taking into account the pedigree structure with both sire and maternal grandsire in the model, which had also a clear impact on the phenotypic trend over those years (Figure 3.1). Furthermore, 305-day milk production increased from 6664 to 7266 kg between 1989–93 and 1994–98 and increased from 7796 to 7883 kg between 2004–08 and 2009–13. This also demonstrates that increasing milk production per se became of less importance for farmers. Also, in the early 1990s there was a stronger association between within-herd production level and the opportunity to survive (Figure 3.5), but genetic trend for functional survival declined until birth year 1999 (Figure 3.7). Another observation was that the genetic trend was sensitive to the inclusion of fixed effects for AFC and within-herd production level in the model. Here we presented the genetic trend from an analysis that included the effects of the interaction of AFC, within-herd production level and 5-year intervals. These effects were clearly important (Figures 3.4 and 3.5) in the data. Initially, when AFC was fitted as one fixed effect across 25 years,

3 Changes of survival over the last 25 years

we saw that genetic trend for functional survival was affected to such an extent that it even declined after 1999. Modelling fixed effects as if these are the same across production levels and years is clearly too simplistic. The effect of AFC on survival had an interaction with 5-year interval (Figure 3.4) and within-herd production level (Figure 3.6). Also, within-herd production level is affected directly by AFC, as later calving leads to a higher milk yield (Nor et al., 2013). These results show that it is important to adjust survival correctly for milk production and AFC over the years in genetic evaluation and when evaluating genetic trends for longevity of animals properly.

With survival there was an optimum AFC of 23 to 24 months for survival. Also, in most other countries an AFC around 24 months was found to be related with highest survival rate, but also with highest lifetime production: the Netherlands (Nor et al., 2013), Wallonia, Belgium (Froidmont et al., 2013), France (Ducrocq, 2005), Ireland (Evans et al., 2006, Berry and Cromie, 2009), United Kingdom (Wathes et al., 2014), Italy (Pirlo et al., 2000), Israel (Weller and Ezra, 2015), Iran (Nilforooshan and Edriss, 2004), Australia (Haworth et al., 2008) and Canada (Sewalem et al., 2005). However, with functional survival, that adjusts all cows to an average production level, no optimum AFC was found for survival. A lower AFC resulted always in a higher survival (Figure 3.3). Although cows with AFC<24 months tend to have lower within-herd production levels (results not shown), this lower production was why they were culled and not AFC alone. If rearing management ensures sufficient development of heifers before starting to breed them, generally influenced by nutrition and pre-pubertal growth rate during the rearing period (Wathes et al., 2008), then cows that calve at AFC of 24 months or younger are more likely to survive first lactation than when calving at a higher AFC. Furthermore, AFC offers a good option to improve survival of first lactation animals.

3.6 Conclusions

In the Netherlands, survival during first year after first calving increased by 8% up to 92% in the last 25 years, but when accounting for the pedigree, survival showed no improvement up to 1999. Genetically, survival increased 3% to 4% and functional survival did not increase over this period. We found an interesting difference between the genetic trends for survival and functional survival for bulls born between 1985 and 1999, where the trend for survival was increasing, but was negative for functional survival. Since 1999, genetic trend picked up again for both survival and functional survival. AFC, season of calving and within-herd production

level affected survival. However, these effects became less important in the most recent years. Based on survival the optimum AFC is around 24 months, but based on functional survival it is better to have an AFC<24 months. Overall, the survival rate of heifers has increased considerably in the past 25 years, initially due to the focus on a high milk production. More recently, the importance of a high milk production has been reduced towards attention on functional survival.

3.7 Acknowledgments

The authors acknowledge funding from the Dutch Dairy Board (PZ; Zoetermeer, the Netherlands), Genetic Evaluation for Sires and CRV (Arnhem, the Netherlands), and the Breed4Food project (programme 'Kennisbasis Dier', code: KB-12-006.-03-008-ASG-LR).

3.8 References

- Bach, A. 2011. Associations between several aspects of heifer development and dairy cow survivability to second lactation. *J. Dairy Sci.* 94:1052-1057.
- Barkema, H. W., M. A. G. von Keyserlingk, J. P. Kastelic, T. J. G. M. Lam, C. Luby, J. P. Roy, S. J. LeBlanc, G. P. Keefe, and D. F. Kelton. 2015. Invited review: Changes in the dairy industry affecting dairy cattle health and welfare. *J. Dairy Sci.* 98:7426-7445.
- Bergevoet, R. H. M., C. J. M. Ondersteijn, H. W. Saatkamp, C. M. J. van Woerkum, and R. B. M. Huirne. 2004. Entrepreneurial behaviour of Dutch dairy farmers under a milk quota system: goals, objectives and attitudes. *Agric. Syst.* 80:1-21.
- Berry, D. P. and A. R. Cromie. 2009. Associations between age at first calving and subsequent performance in Irish spring calving Holstein-Friesian dairy cows. *Livest. Sci.* 123:44-54.
- Boettcher, P. J., L. K. Jairath, and J. C. W. Dekkers. 1999. Comparison of methods for genetic evaluation of sires for survival of their daughters in the first three lactations. *J. Dairy Sci.* 82:1034-1044.
- CRV. 2015a. CRV Jaarstatistieken Nederland 2015. Arnhem, the Netherlands.
- CRV. 2015b. Genetisch trend van koeien in Nederland. Accessed: April 6, 2016. https://www.crv4all.nl/wp-content/uploads/2016/03/gen_trend_koe_nl_20151203.pdf.
- Dechow, C. D. and R. C. Goodling. 2008. Mortality, culling by sixty days in milk, and production profiles in high- and low-survival Pennsylvania herds. *J. Dairy Sci.* 91:4630-4639.
- Dematawewa, C. M. B. and P. J. Berger. 1998. Genetic and phenotypic parameters for 305-day yield, fertility, and survival in Holsteins. *J. Dairy Sci.* 81:2700-2709.
- Ducrocq, V. 2005. An improved model for the French genetic evaluation of dairy bulls on length of productive life of their daughters. *Anim. Sci.* 80:249-256.
- Evans, R. D., M. Wallace, D. J. Garrick, P. Dillon, D. P. Berry, and V. Olori. 2006. Effects of calving age, breed fraction and month of calving on calving interval and survival across parities in Irish spring-calving dairy cows. *Livest. Sci.* 100:216-230.
- Froidmont, E., P. Mayeres, P. Picron, A. Turlot, V. Planchon, and D. Stilmant. 2013. Association between age at first calving, year and season of first calving and milk production in Holstein cows. *Animal* 7:665-672.
- Fürst, C. and B. Fürst-Waltl. 2006. Breeding aspects of the calving process, rate of stillbirths and longevity in dairy cattle. *Zuchtungskunde* 78:365-383.

- Gilmour, A. R., B. Gogel, B. Cullis, R. Thompson, D. Butler, M. Cherry, D. Collins, G. Dutkowsky, S. Harding, and K. Haskard. 2009. ASReml user guide release 3.0. VSN International Ltd, Hemel Hempstead, UK.
- Hadley, G. L., C. A. Wolf, and S. B. Harsh. 2006. Dairy cattle culling patterns, explanations, and implications. *J. Dairy Sci.* 89:2286-2296.
- Hare, E., H. D. Norman, and J. R. Wright. 2006. Survival rates and productive herd life of dairy cattle in the United States. *J. Dairy Sci.* 89:3713-3720.
- Haworth, G. M., W. P. Tranter, J. N. Chuck, Z. Cheng, and D. C. Wathes. 2008. Relationships between age at first calving and first lactation milk yield, and lifetime productivity and longevity in dairy cows. *Vet. Rec.* 162:643-647.
- Jongeneel, R. and S. Van Berkum. 2015. What will happen after the EU milk quota system expires in 2015? An assessment of the Dutch dairy sector. LEI Wageningen UR, Wageningen, the Netherlands.
- Miglior, F., J. Chesnais, and B. J. Van Doormaal. 2012. Genetic improvement: A major component of increased dairy farm profitability. in Invited presentation at 38th ICAR Biennial Session held in Cork, Ireland. ICAR Technical Series. International Committee for Animal Recording, Rome, Italy.
- Miglior, F., B. L. Muir, and B. J. Van Doormaal. 2005. Selection indices in Holstein cattle of various countries. *J. Dairy Sci.* 88:1255-1263.
- Nieuwhof, G. J., H. D. Norman, and F. N. Dickinson. 1989. Phenotypic trends in herd life of dairy cows in the United States. *J. Dairy Sci.* 72:726-736.
- Nilforooshan, M. A. and M. A. Edriss. 2004. Effect of age at first calving on some productive and longevity traits in Iranian Holsteins of the Isfahan Province. *J. Dairy Sci.* 87:2130-2135.
- Nor, N. M., W. Steeneveld, T. van Werven, M. C. M. Mourits, and H. Hogeveen. 2013. First-calving age and first-lactation milk production on Dutch dairy farms. *J. Dairy Sci.* 96:981-992.
- Oltenucu, P. A. and B. Algers. 2005. Selection for increased production and the welfare of dairy cows: are new breeding goals needed? *AMBIO: A Journal of the Human Environment* 34:311-315.
- Oltenucu, P. A. and D. M. Broom. 2010. The impact of genetic selection for increased milk yield on the welfare of dairy cows. *Anim. Welf.* 19:39-49.
- Pirlo, G., F. Miglior, and M. Speroni. 2000. Effect of age at first calving on production traits and on difference between milk yield returns and rearing costs in Italian Holsteins. *J. Dairy Sci.* 83:603-608.
- Pryce, J. E., R. F. Veerkamp, R. Thompson, W. G. Hill, and G. Simm. 1997. Genetic aspects of common health disorders and measures of fertility in Holstein Friesian dairy cattle. *Anim. Sci.* 65:353-360.

3 Changes of survival over the last 25 years

- Robertson, A. 1966. A mathematical model of culling process in dairy cattle. *Animal Production* 8:95-108.
- Sewalem, A., G. J. Kistemaker, V. Ducrocq, and B. J. Van Doormaal. 2005. Genetic analysis of herd life in Canadian dairy cattle on a lactation basis using a Weibull proportional hazards model. *J. Dairy Sci.* 88:368-375.
- Sewalem, A., F. Miglior, G. J. Kistemaker, P. Sullivan, G. Huapaya, and B. J. Van Doormaal. 2007. Short communication: Modification of genetic evaluation of herd life from a three-trait to a five-trait model in Canadian dairy cattle. *J. Dairy Sci.* 90:2025-2028.
- Terawaki, Y. and V. Ducrocq. 2009. Nongenetic effects and genetic parameters for length of productive life of Holstein cows in Hokkaido, Japan. *J. Dairy Sci.* 92:2144-2150.
- Van der Linde, C., A. Harbers, and G. de Jong. 2007. From functional to productive longevity in the Netherlands. *Interbull Bulletin* 37:203.
- Van Pelt, M. L., T. H. E. Meuwissen, G. de Jong, and R. F. Veerkamp. 2015. Genetic analysis of longevity in Dutch dairy cattle using random regression. *J. Dairy Sci.* 98:4117-4130.
- Veerkamp, R., J. Windig, M. Calus, W. Ouweltjes, Y. de Haas, and B. Beerda. 2008. Selection for high production in dairy cattle. Pages 243-260 in *Resource allocation theory applied to farm animal production*. W. Rauw, ed. CAB International, Wallingford, UK.
- Veerkamp, R. F., W. G. Hill, A. W. Stott, S. Brotherstone, and G. Simm. 1995. Selection for longevity and yield in dairy-cows using transmitting abilities for type and yield. *Anim. Sci.* 61:189-197.
- Vitali, A., M. Segnalini, L. Bertocchi, U. Bernabucci, A. Nardone, and N. Lacetera. 2009. Seasonal pattern of mortality and relationships between mortality and temperature-humidity index in dairy cows. *J. Dairy Sci.* 92:3781-3790.
- Vollema, A. R., S. van der Beek, A. G. F. Harbers, and G. de Jong. 2000. Genetic evaluation for longevity of Dutch dairy bulls. *J. Dairy Sci.* 83:2629-2639.
- Wathes, D. C., J. S. Brickell, N. E. Bourne, A. Swali, and Z. Cheng. 2008. Factors influencing heifer survival and fertility on commercial dairy farms. *Animal* 2:1135-1143.
- Wathes, D. C., G. E. Pollott, K. F. Johnson, H. Richardson, and J. S. Cooke. 2014. Heifer fertility and carry over consequences for life time production in dairy and beef cattle. *Animal* 8:91-104.
- Weigel, K. A. 2006. Prospects for improving reproductive performance through genetic selection. *Anim. Reprod. Sci.* 96:323-330.

- Weigel, K. A., R. W. Palmer, and D. Z. Caraviello. 2003. Investigation of factors affecting voluntary and involuntary culling in expanding dairy herds in Wisconsin using survival analysis. *J. Dairy Sci.* 86:1482-1486.
- Weller, J. I. and E. Ezra. 2015. Environmental and genetic factors affecting cow survival of Israeli Holsteins. *J. Dairy Sci.* 98:676-684.
- Windig, J. J., M. P. L. Calus, B. Beerda, and R. F. Veerkamp. 2006. Genetic correlations between milk production and health and fertility depending on herd environment. *J. Dairy Sci.* 89:1765-1775.

4

Genetic changes of survival traits over the past 25 yr in Dutch dairy cattle

M.L. van Pelt^{1,2}, V. Ducrocq³, G. de Jong¹, M.P.L. Calus², R.F. Veerkamp²

¹ CRV BV, Animal Evaluation Unit, PO Box 454, 6800 AL Arnhem, the Netherlands;

² Animal Breeding and Genomics Centre, Wageningen UR Livestock Research, PO Box 338, 6700 AH Wageningen, the Netherlands;

³ GABI INRA, AgroParisTech, Université Paris-Saclay 78350 Jouy-en-Josas, France

Journal of Dairy Science (2016) 99:9810-9819

DOI: <http://dx.doi.org/10.3168/jds.2016-11249>

Abstract

Genetic correlations and heritabilities for survival were investigated over a period of 25 yr to evaluate if survival in first lactation has become a different trait and if this is affected by adjusting for production level. Survival after first calving until 12 mo after calving (surv_12mo) and survival of first lactation (surv_1st_lac) were analyzed in Dutch black-and-white cows. The data set contained 1,108,745 animals for surv_12mo and 1,062,276 animals for surv_1st_lac, with first calving between 1989 and 2013. The trait survival as recorded over 25 yr was split in five 5-yr intervals to enable a multitrait analysis. Bivariate models using subsets of the full data set and multitrait and autoregressive models using the full data set were used. Survival and functional survival were analyzed. Functional survival was defined as survival adjusted for within-herd production level for 305-d yield of combined kilograms of fat and protein. Mean survival increased over time, whereas genetic variances and heritability decreased. Bivariate models yielded large standard errors on genetic correlations due to poor connectedness between the extreme 5-yr intervals. The more parsimonious models using the full data set gave nonunity genetic correlations. Genetic correlations for survival were below 0.90 between intervals separated by 1 or more 5-yr intervals. Genetic correlations for functional survival did not indicate that definition of survival changed (≥ 0.90). The difference in genetic correlations between survival and functional survival is likely explained by lower emphasis of dairy farmers on culling in first lactation for low yield in more recent years. This suggests that genetic evaluation for longevity using historical data should analyze functional survival rather than survival.

Key words: longevity, survival, genetic correlation

4.1 Introduction

Longevity, also known as survival, is an important trait for dairy farmers because it is related to economic and animal welfare reasons. In the Netherlands, pressure exists from society to increase the longevity of dairy cows (LTO, 2011). The longevity of a dairy cow is determined by the culling decision of the farmer, and farmers decide to cull dairy cows for various reasons such as mastitis, fertility problems, lameness, or low production (Beaudeau et al., 2000, Zijlstra et al., 2013). These culling reasons are likely to be affected over years by changes; for example, in national regulation and legislation (e.g., quota system), feed costs, milk price, and revenues for culled animals. In Europe, a quota system was in place until April 2015 to limit the amount of milk produced annually per country, but also per farmer (Bergevoet et al., 2004). Individual farmers that produced more than the allowed quota had to pay a penalty for the excess of milk they produced. This system might have affected culling reasons over the years. From 1990 to 2014 the average herd size increased, milk production per cow increased, and fertility declined in the Netherlands (CRV, 2015). Also, worldwide, national selection indexes have changed drastically over the past 2 decades, where the breeding goal moved from selection for production only toward selection for production, longevity, and health traits (Miglior et al., 2005). This might have changed the perception of the right culling reasons for dairy farmers. Currently, in the genetic evaluation in the Netherlands, 30 yr of records are included, but it is still assumed to be 1 trait.

In the Netherlands, farmers can voluntarily record culling reasons, but not all farmers record these reasons and data are only available for the most recent years. Recorded culling reasons are subjective scores of the farmer, and culling of a cow is often done for multiple reasons (Fetrow et al., 2006). For example, poor fertility or bad udder health combined with a low daily milk yield will result in a cow not being inseminated again. For these reasons, it is difficult to evaluate if culling reasons have changed over the years. Another way to evaluate changes in the trait survival is to analyze phenotypic records for longevity and compare the mean and standard deviation across years; however, this provides little information if the trait definition (i.e., culling reasons) of survival changed. The trait definition can be also evaluated by analyzing if daughters of the same bulls survive longest in each year or if a reranking of bulls occurred. Genetic links between animals culled in different years allow estimation of genetic correlations between years, and a unity correlation is expected when culling reasons are not different between years. In the United States, changes in productive life were analyzed over a 15-yr period

(Tsuruta et al., 2004), and in Australia survival of first lactation was analyzed over a 20-yr period (Haile-Mariam and Pryce, 2015). In both studies multitrait and random regression models were used to detect changes in their survival trait, and in both studies it appeared that with random regression models genetic correlations between survival in different years were high. However, with multitrait models, estimated genetic correlations were as low as 0.21, with large standard errors (Haile-Mariam and Pryce, 2015), and 0.22 (Tsuruta et al., 2004); therefore, it was difficult to conclude if survival in the United States and Australia changed in those years.

Two different definitions for survival were proposed by (Ducrocq, 1987): survival and functional survival. Survival is defined as the ability to delay culling, and functional survival is defined as the ability to delay involuntary culling. Adjustment of survival for individual deviation from within-herd production levels provides an approximation of functional survival. By adjusting for production level, the assumption is made that functional survival is independent of voluntary culling, assuming that voluntary culling is primarily for low production. Voluntary culling is an economic decision of the farmer because more profit from a new heifer is expected. If involuntary culling is decreased, a higher voluntary culling rate can be applied, resulting in a larger profit for the farmer, but not necessarily increasing longevity as a whole (Van Arendonk, 1985). It is of interest to investigate whether survival and functional survival have evolved differently over time.

The objective of this study was to evaluate if survival in first lactation has become a different trait over the past 25 yr or not, and if this is affected by adjusting for production level or not. This was achieved by computing the genetic correlation between survival of first parities starting in different years. For the Dutch situation, it was the first time that such a long time span could be analyzed, as data on survival has been available in the Netherlands since 1988.

4.2 Material and Methods

4.2.1 Data

In this study, 2 definitions for survival after first calving were used: (1) survival until mo 12 (**surv_12mo**; i.e., the animal did not die or was not culled for slaughter), and (2) survival of first lactation (**surv_1st_lac**; i.e., a cow was considered to have survived first lactation if she had initiated her second lactation). Besides **surv_12mo**, we analyzed **surv_1st_lac** because over the past decades calving interval has increased and culling for infertility could have shifted beyond mo 12. Survival in early life was analyzed instead of total lifespan, because information for

4 Genetic changes of survival traits

surv_12mo and surv_1st_lac is more rapidly available for each animal compared with lifespan, and, more importantly, because the modeling of the data is less complex. With only first-lactation animals, all animals within a calving year can be directly compared instead of comparing a mixture of different age groups at the same time within a herd.

The surv_12mo and surv_1st_lac were coded as 1 for animals that survived and as 0 for animals that died or were culled for slaughter. Data were available from the Dutch and Flemish cattle improvement cooperative CRV, and cows with first calving between 1989 and 2013 were used. Records for survival were constructed from records of pedigree, lactations, and movements of cows in the Netherlands. All animals were a combination of at least 87.5% Holstein Friesian and Dutch Friesian, and the required age at first calving was between 21 and 40 mo. Herds with at least 95% Holstein Friesian and Dutch Friesian genes were selected. Only herds with at least 30 cows present every month, in the period from 1994 to 2013, were selected to exclude herds with a nonconventional culling management; for example, because the farm has stopped or the entire herd was culled at once because of the outbreak of a disease. Additional requirements were that sires had at least 15 progeny or grand-progeny that could have been productive for at least 12 mo after first calving, sires had progeny in at least 2 herds, and every herd-year-month class had at least 25 observations. The additional requirements for progeny per sire, herds per sire, and observations per herd-year-month had to be repeated 11 times until the final data set met all criteria. The creation of the final data set with surv_12mo and surv_1st_lac was as follows. For surv_12mo, animals were selected that could have been in the herd for at least 12 mo after first calving (i.e., animals calving after 31 December 2013 were excluded); the data set with surv_12mo contained 1,108,745 animals. For surv_1st_lac, animals were selected that could have been in the herd for at least 18 mo after first calving (i.e., animals calving after 30 June 2013 were excluded); the data set with surv_1st_lac contained 1,062,276 animals. In both data sets 2,185 herds were included. Pedigree information of the sires and maternal grandsires was traced back 6 generations, resulting in a pedigree file with 11,268 sires.

4.2.2 Statistical Model

To determine whether genetic parameters for survival changed across years, the period of 25 yr was split into 5 intervals of 5 yr (1989–1993, 1994–1998, 1999–2003, 2004–2008, and 2009–2013) and analyzed as 5 traits. The basic mixed model equation for all models to analyze survival was:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + (\mathbf{Z}_s + \mathbf{Z}_{mgs}/2)\mathbf{u}_s + \mathbf{e}$$

where \mathbf{y} is a vector of observations (0/1) for survival (surv_12mo or surv_1st_lac); \mathbf{X} is an incidence matrix linking the observations to the fixed effects; \mathbf{b} is a vector of fixed effects of herd-year of first calving, year-month of first calving, fixed regression on age at first calving (15 classes: 21, 22,..., 34, ≥ 35 mo) with a second-order Legendre polynomial, and percentage of Holstein Friesian genes, divided into 5 classes (≤ 50.0 , 62.5, 75.0, 87.5, and 100%); \mathbf{Z}_s and \mathbf{Z}_{mgs} are the incidence matrices for sire and maternal grandsire, which are overlaid to estimate random sire effects; \mathbf{u}_s is the vector of random sire effects; and \mathbf{e} is a vector of random residual effects. The variances were defined as

$$Var \begin{bmatrix} \mathbf{u} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{G} \otimes \mathbf{A} & 0 \\ 0 & \mathbf{R} \otimes \mathbf{I} \end{bmatrix}$$

where \mathbf{G} is the 5×5 sire (co)variance matrix for five 5-yr intervals; \mathbf{A} is the additive genetic relationship matrix between sires; \mathbf{R} is a 5×5 diagonal heterogeneous residual (co)variance matrix for five 5-yr intervals; \mathbf{I} is identity matrix; and \otimes is the Kronecker product.

4.2.3 Bivariate Analysis

We first used a bivariate model to estimate genetic correlations between all pairwise combinations of 5-yr intervals. To validate the consistency of the genetic correlations for survival, 2 alternative models were considered. The first alternative model was surv_12mo analyzed as functional survival, where survival was adjusted for individual production level relative to the production level of the herd; this has been suggested as a way to correct for voluntary culling due to low production (Robertson, 1966). Therefore, for functional survival, a fixed effect for within-herd production level was included, where within-herd production level was the cow ranking within a herd by 5-yr interval for predicted or realized 305-d yield of combined kilograms of fat and protein. Animals were ranked into 7 classes from worst to best for combined kilograms of fat and protein, as (1) 1 to 5%, (2) 6 to 20%, (3) 21 to 40%, (4) 41 to 60%, (5) 61 to 80%, (6) 81 to 95%, and (7) 96 to 100%. The second alternative model was surv_12mo analyzed with a threshold model, using the logit link function, to test for a difference in estimated parameters of linear and threshold models with high survival rates.

Our base model was the linear model for surv_12mo because our interest was if surv_12mo changed genetically over 25 yr. In subsequent analyses we changed the

base model to explore sensitivity of the assumptions. In analysis 2 we changed the trait from surv_12mo to surv_1st_lac to analyze the effect of culling in late lactation on the genetic correlations between intervals; in analysis 3 we added the fixed effect for within-herd production level to analyze functional surv_12mo; and in analyses 4 and 5 we used a threshold model instead of a linear model to validate that ranking of bulls was similar to a linear model for surv_12mo and functional surv_12mo, respectively.

Variance components were estimated with ASReml (Gilmour et al., 2015). Heritability for each 5-yr interval was calculated as $4 \times$ sire variance divided by the sum of $1.25 \times$ sire variance (1 for the sire and 0.25 for the maternal grandsire) plus the residual variance of the corresponding interval.

4.2.4 Alternative Parameterizations

For some combinations of 5-yr intervals the number of common sires and maternal grandsires was very low (Table 4.1), indicating poor genetic links. Because of the poor connectedness between the extreme 5-yr intervals, we also examined more parsimonious models to estimate the genetic (co)variance matrix \mathbf{G} on the full data set. The following models were used [examples of genetic (co)variance matrix \mathbf{G} are displayed for the multivariate and autoregressive models in Figure 4.1].

- a) Univariate model with one genetic effect and homogeneous residual variances for 5-yr intervals (\mathbf{UN}_{hom}).
- b) Univariate model with one genetic effect and heterogeneous residual variances for 5-yr intervals (\mathbf{UN}_{het}).
- c) Uncorrelated multivariate model with heterogeneous variances (\mathbf{MT}_{uc}), which is equivalent to a separate univariate analysis for each 5-yr interval. All covariances in \mathbf{G} are zero.
- d) Multivariate model with a constrained correlation of 0.95 between all 5-yr intervals and with heterogeneous variances for 5-yr intervals ($\mathbf{MT}_{0.95}$).
- e) Multivariate model with a constrained correlation of 0.99 between all 5-yr intervals and with heterogeneous variances for 5-yr intervals ($\mathbf{MT}_{0.99}$).
- f) Full multivariate model with heterogeneous variances for 5-yr intervals, where the additive genetic (co)variance matrix \mathbf{G} is an unstructured matrix (\mathbf{MT}_{us}).
- g) First-order autoregressive (\mathbf{AR}) model with heterogeneous variances for 5-yr intervals ($\mathbf{AR1}$). Matrix \mathbf{G} has an autoregressive correlation structure with the correlation function $C_{ii} = 1$, $C_{i+1,i} = \phi_1$, and $C_{ij} = \phi_1 C_{i-1,j}$. Only one extra parameter, ϕ_1 , is needed, besides the variances, to estimate the genetic correlation structure.

- h) First-order autoregressive model with heterogeneous variances for 5-yr intervals and constrained parameter ϕ_1 of 0.90 (**AR1_{0.90}**).
- i) Second-order autoregressive model with heterogeneous variances for 5-yr intervals (**AR2**). Matrix **G** has an autoregressive correlation structure with the correlation function where correlation $C_{ii} = 1$, $C_{i+1,i} = \phi_1/(1 - \phi_2)$, and $C_{ij} = \phi_1 C_{i-1,j} + \phi_2 C_{i-2,j}$. Two extra parameters, ϕ_1 and ϕ_2 , are needed, besides the variances, to estimate the genetic correlation structure.

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

MT_{UC}

$$\begin{bmatrix} 1 & a & b & d & g \\ a & 1 & c & e & h \\ b & c & 1 & f & i \\ d & e & f & 1 & j \\ g & h & i & j & 1 \end{bmatrix}$$

MT_{US}

$$\begin{bmatrix} 1 & a & a & a & a \\ a & 1 & a & a & a \\ a & a & 1 & a & a \\ a & a & a & 1 & a \\ a & a & a & a & 1 \end{bmatrix}$$

MT_{0.95} and MT_{0.99}

$$\begin{bmatrix} 1 & a & b & c & d \\ a & 1 & a & b & c \\ b & a & 1 & a & b \\ c & b & a & 1 & a \\ d & c & b & a & 1 \end{bmatrix}$$

AR1, AR1_{0.90} and AR2

Figure 4.1. Genetic covariance structures fitted to full data set for survival for the multitrait (MT) and autoregressive (AR) models. Different letters (a , b , etc.) represent different values of correlation. MT_{UC} has an uncorrelated covariance matrix. MT_{0.95} and MT_{0.99} have constrained correlations, where a is 0.95 and 0.99. MT_{US} has an unstructured (US) covariance matrix. First order autoregressive (AR1) correlation $C_{ii} = 1$, $C_{i+1,i} = \phi_1$, and $C_{ij} = \phi_1 C_{i-1,j}$. In AR1_{0.90} parameter ϕ_1 is constrained to 0.90. Second order autoregressive (AR2) correlation $C_{ii} = 1$, $C_{i+1,i} = \phi_1/(1 - \phi_2)$, and $C_{ij} = \phi_1 C_{i-1,j} + \phi_2 C_{i-2,j}$. For AR1, AR1_{0.90} and AR2 a is $C_{i+1,i}$, and b , c , d are C_{ij} .

4 Genetic changes of survival traits

Table 4.1. Number of common sires and maternal grandsires (in parentheses) in same 5-yr interval (diagonal) and in two 5-yr intervals (below diagonal)

Item	1989–1993	1994–1998	1999–2003	2004–2008	2009–2013
1989–1993	2,195 (1,988)				
1994–1998	697 (1,855)	2,557 (3,679)			
1999–2003	179 (737)	759 (2,463)	3,186 (4,649)		
2004–2008	90 (168)	259 (894)	1,029 (2,992)	3,189 (5,239)	
2009–2013	43 (56)	130 (242)	349 (1,228)	1,125 (3,336)	2,569 (5,281)

4.2.5 Model Comparison

For selection of the best-fit model among the 9 alternative models, Akaike's information criteria (AIC), and Bayesian information criteria (BIC) were used. The AIC and BIC were defined as:

$$AIC = -2 \log(L_i) + 2t_i$$

$$BIC = -2 \log(L_i) + t_i \log v$$

where L_i is the log-likelihood of model i , t_i is the number of (co)variance parameters to estimate in model i and $v = n - p$ is the residual degrees of freedom, where n was the number of observations and p was the number of estimable fixed effects in the model. The AIC (Akaike, 1973) and BIC (Schwarz, 1978) were calculated for each model and the model with the smallest value was considered the preferred model. In addition, when AIC and BIC gave a different preferred model, the ranking based on AIC and BIC were combined and the model with the best overall ranking was considered the preferred model.

4.3 Results

Mean survival rate was 87.8% for surv_12mo and 83.2% for surv_1st_lac in the total data set. Between 1989 to 1993 and 2009 to 2013 survival increased by 7.3% for surv_12mo and 7.1% for surv_1st_lac (Table 4.2). The largest increase in mean survival rate was between 1994 to 1998 and 2004 to 2008 for both surv_12mo (5.4%) and surv_1st_lac (5.7%). Mean survival rates were more similar for 5-yr intervals 1989 to 1993 and 1994 to 1998, as well as for intervals 2004 to 2008 and 2009 to 2013. The difference between surv_12mo and surv_1st_lac can be

considered as culling at the end of lactation, and their difference varied between 4 and 5% across all 5-yr intervals. However, the relative reduction of percentage culled animals from 1989 to 1993 and 2009 to 2013 was 44% for surv_12mo and 35% for surv_1st_lac. Therefore, in more recent years, animals are more likely culled after 12 mo.

Table 4.2. Number of cows and mean survival rate (%) for survival after first calving and survival of first lactation per 5-yr interval

5-yr interval	Survival at 12 mo		Survival of first lactation	
	No. of cows	Mean	No. of cows	Mean
1989–1993	175,822	83.4	174,216	79.5
1994–1998	209,941	84.9	207,547	80.1
1999–2003	226,278	88.0	222,360	83.0
2004–2008	236,362	90.3	231,588	85.8
2009–2013	260,342	90.7	226,565	86.6

4.3.1 Variances and Heritabilities

Parameter estimates for survival across 5-yr intervals based on bivariate analyses are shown in Tables 4.3, 4.4, 4.5, 4.6, and 4.7. Across 5-yr intervals genetic and residual standard deviations reduced for all 4 analyses. The decrease of genetic standard deviations was relatively larger for surv_12mo (Table 4.3) and surv_1st_lac (Table 4.4) than for functional surv_12mo (Table 4.5). Residual standard deviations also decreased across 5-yr intervals, except for surv_12mo and functional surv_12mo analyzed with the threshold model, which has by definition equal residual variances for both traits (Table 4.6 and 4.7). The changes in genetic and residual standard deviations were not proportional; the relative decrease of residual standard deviations was smaller compared with the relative decrease of the genetic variance. This resulted in decreasing heritabilities across 5-yr intervals. When surv_12mo was adjusted for within-herd production level, this is functional survival, heritability for the first three 5-yr intervals was constant, followed by lower heritabilities in the later 5-yr intervals (Table 4.5). Heritabilities of survival across 5-yr intervals estimated with the threshold model were decreasing as well, similar to the linear models (Table 4.6), where heritabilities for functional survival were more constant again (Table 4.7). When heritabilities estimated with a linear model for surv_12mo were transformed to the underlying scale with the formula of Dempster and Lerner (1950), they were similar to the heritabilities from the threshold model (0.04–0.14). Heritabilities on the underlying scale for functional surv_12mo ranged from 0.07 to 0.10 and for surv_1st_lac from 0.06 to 0.12.

4 Genetic changes of survival traits

Table 4.3. Genetic standard deviation (σ_A), residual standard deviation (σ_E), heritability (in bold on the diagonal), and genetic correlations (offdiagonal) for survival at 12 mo after first calving divided in five 5-yr intervals (SE in parentheses)

5-year interval	σ_A	σ_E	heritability and genetic correlations				
			1989 – 1993	1994 – 1998	1999 – 2003	2004 – 2008	2009 – 2013
1989 – 1993	0.093	0.365	0.063 (0.007)				
1994 – 1998	0.084	0.353	0.88 (0.04)	0.056 (0.006)			
1999 – 2003	0.059	0.320	0.95 (0.06)	0.98 (0.02)	0.034 (0.006)		
2004 – 2008	0.044	0.293	0.74 (0.19)	0.83 (0.08)	0.96 (0.03)	0.022 (0.006)	
2009 – 2013	0.033	0.288	0.97 (0.22)	0.99 (0.09)	0.86 (0.09)	0.90 (0.05)	0.013 (0.006)

Table 4.4. Genetic standard deviation (σ_A), residual standard deviation (σ_E), heritability (in bold on the diagonal), and genetic correlations (offdiagonal) for survival of first lactation divided in five 5-yr intervals (SE in parentheses)

5-year interval	σ_A	σ_E	heritability and genetic correlations				
			1989 – 1993	1994 – 1998	1999 – 2003	2004 – 2008	2009 – 2013
1989 – 1993	0.097	0.397	0.059 (0.006)				
1994 – 1998	0.091	0.393	0.84 (0.05)	0.053 (0.006)			
1999 – 2003	0.078	0.370	0.94 (0.06)	0.92 (0.03)	0.044 (0.006)		
2004 – 2008	0.066	0.344	0.76 (0.17)	0.80 (0.08)	0.94 (0.03)	0.036 (0.006)	
2009 – 2013	0.051	0.337	0.80 (0.31)	0.91 (0.12)	0.87 (0.07)	0.95 (0.03)	0.023 (0.005)

4 Genetic changes of survival traits

Table 4.5. Genetic standard deviation (σ_A), residual standard deviation (σ_E), heritability (in bold on the diagonal), and genetic correlations (offdiagonal) for functional survival at 12 mo after first calving divided in five 5-yr intervals (SE in parentheses)

5-year interval	σ_A	σ_E	heritability and genetic correlations				
			1989 – 1993	1994 – 1998	1999 – 2003	2004 – 2008	2009 – 2013
1989 – 1993	0.070	0.346	0.041 (0.005)				
1994 – 1998	0.068	0.336	0.92 (0.04)	0.040 (0.005)			
1999 – 2003	0.062	0.308	0.93 (0.07)	1.00 (0.01)	0.039 (0.005)		
2004 – 2008	0.048	0.284	0.96 (0.12)	0.91 (0.06)	1.00 (0.01)	0.028 (0.005)	
2009 – 2013	0.043	0.279	0.90 (0.18)	0.86 (0.13)	0.90 (0.06)	0.94 (0.03)	0.023 (0.005)

Table 4.6. Genetic standard deviation (σ_A), residual standard deviation (σ_E), heritability (in bold on the diagonal), and genetic correlations (offdiagonal) for survival at 12 mo after first calving divided in five 5-yr intervals (SE in parentheses) estimated with a threshold model

5-year interval	σ_A	σ_E	heritability and genetic correlations				
			1989 – 1993	1994 – 1998	1999 – 2003	2004 – 2008	2009 – 2013
1989 – 1993	0.721	1.814	0.150 (0.015)				
1994 – 1998	0.656	1.814	0.85 (0.05)	0.125 (0.014)			
1999 – 2003	0.566	1.814	0.93 (0.07)	0.97 (0.02)	0.095 (0.013)		
2004 – 2008	0.500	1.814	0.69 (0.19)	0.80 (0.09)	0.94 (0.04)	0.075 (0.013)	
2009 – 2013	0.412	1.814	0.83 (0.27)	0.98 (0.11)	0.84 (0.09)	0.89 (0.05)	0.052 (0.012)

4 Genetic changes of survival traits

Table 4.7. Genetic standard deviation (σ_A), residual standard deviation (σ_E), heritability (in bold on the diagonal), and genetic correlations (offdiagonal) for functional survival at 12 mo after first calving divided in five 5-yr intervals (SE in parentheses) estimated with a threshold model

5-year interval	σ_A	σ_E	heritability and genetic correlations					
			1989 – 1993	1994 – 1998	1999 – 2003	2004 – 2008	2009 – 2013	
1989 – 1993	0.648	1.814	0.120 (0.014)					
1994 – 1998	0.600	1.814	0.89 (0.05)	0.106 (0.013)				
1999 – 2003	0.648	1.814	0.89 (0.08)	1.00 (0.01)	0.123 (0.013)			
2004 – 2008	0.600	1.814	0.91 (0.13)	0.89 (0.07)	1.00 (0.02)	0.106 (0.012)		
2009 – 2013	0.539	1.814	0.90 (0.19)	0.83 (0.14)	0.89 (0.07)	0.90 (0.04)	0.086 (0.012)	

4.3.2 Genetic Correlations

Most genetic correlations obtained with the bivariate analyses were higher than 0.85 and did not differ significantly from unity. The few genetic correlations that did differ significantly from unity were high (>0.80). Standard errors of the genetic correlations became larger when two 5-yr intervals were further apart, because the number of sires and maternal grandsires with progeny in both intervals decreased when intervals were further apart (Table 4.1). Pooled over 10 bivariate analyses per model, the average correlation between 5-yr intervals was 0.91 for *surv_12mo*, 0.87 for *surv_1st_lac*, and 0.87 for *surv_12mo* with the threshold model. Adjusting for within-herd production level resulted in a higher average correlation of 0.93 for functional *surv_12mo* and 0.91 for functional *surv_12mo* with the threshold model. Also, the range of genetic correlations was smaller for functional *surv_12mo* (0.86–1.00) than for *surv_12mo* (0.74–0.99). However, based on standard errors of the genetic correlations, it was difficult to conclude whether difference existed between the models or whether survival between 5-yr intervals differed genetically (i.e., genetic correlations below unity).

Table 4.8. Univariate (UN), multitrait (MT), and autoregressive models (AR) models fitted to the full data set for survival with the total number of variance components estimated (#VC), the number of variance components estimated for the genetic (#gen) and residual (#res) components, the value to which the genetic correlations ($r_g = 0, 0.95, 0.99$) or the autoregressive parameter ϕ_1 were constrained or estimated (s was estimated from the data), Akaike’s information criteria (AIC), and Bayesian information criteria (BIC) for each model for surv_12mo, surv_1st_lac, and functional surv_12mo¹

Model ²	# VC	# gen	# res	r_g /AR	AIC			BIC		
					surv_12mo	surv_1st_lac	functional surv_12mo	surv_12mo	surv_1st_lac	functional surv_12mo
UN _{hom}	2	1	1	–	19306.7	9051.8	15428.4	19209.8	8957.2	15331.3
UN _{het}	6	1	5	–	268.6	177.6	60.9	219.2	130.3	11.3
MT _{UC}	10	5	5	0	324.7	364.8	383.6	322.7	364.8	381.4
MT _{0.95}	10	5	5	0.95	2.0	9.7	2.2	0.1	9.7	0.0
MT _{0.99}	10	5	5	0.99	24.6	48.5	6.2	22.6	48.5	4.1
MT _{US}	20	15	5	s	nc ³	nc	nc	nc	nc	nc
AR1	11	6	5	s	0.0	0.4	0.0	9.9	12.2	9.7
AR1 _{0.90}	10	5	5	0.90	2.0	0.0	12.7	0.0	0.0	10.5
AR2	12	7	5	s	1.7	2.1	3.9	23.5	25.7	25.4

¹Model with lowest AIC or BIC was set to 0.00. surv_12mo = survival until mo 12; surv_1st_lac = survival of first lactation.

²Descriptions of models can be found in the Alternative Parameterizations section.

³nc = models that did not converge.

4.3.3 Alternative Parameterizations of Genetic Covariances

Table 4.8 contains the results of the comparison of alternative parameterizations of the genetic covariance structure G . In comparison with the bivariate analyses, all 5-yr intervals were included rather than two 5-year intervals at the time; most parameterizations of G converged, except $MT_{0.5}$. This exception was the main reason why the results of the bivariate analyses were presented above. The parsimonious models gave similar heritability estimates, except for the univariate models. According to the AIC, AR1 was the best model for `surv_12mo` and `functional surv_12mo`, and $AR1_{0.90}$ was the best model for `surv_1st_lac`. According to the BIC, which puts a higher penalty when modeling more variance components, $AR1_{0.90}$ was the best model for `surv_12mo` and `surv_1st_lac` and $MT_{0.95}$ was the best model for `functional surv_12mo`. Ranking models according to both AIC and BIC suggests that, for `surv_12mo` and `surv_1st_lac`, AR1, $AR1_{0.90}$, and $MT_{0.95}$ gave a better fit than the other parameterizations, and for `functional survival` $MT_{0.95}$, AR1, and multivariate model with a correlation of 0.99 gave a better fit.

With AR1 the estimated autoregressive correlations ϕ_1 were 0.94 for `surv_12mo`, 0.92 for `surv_1st_lac`, 0.93 for `surv_12mo` estimated with the threshold model, 0.97 for `functional surv_12mo`, and 0.96 for `functional surv_12mo` estimated with the threshold model; this resulted in genetic correlations between the most extreme intervals of 0.78, 0.73, 0.74, 0.88, and 0.83 respectively (Table 4.9). Compared with the bivariate analyses, standard errors of the genetic correlations were smaller because the full data set was used (all 5-yr intervals simultaneously), resulting in better connectedness due to common sires and a more parsimonious model with fewer variance components. Interestingly, for `functional surv_12mo`, genetic correlations were closer to unity compared with the analyses without adjusting for within-herd production level. This was supported by the model comparison which suggest that models with a fixed genetic correlations of 0.95 and 0.99 were better for `functional survival`. The average genetic correlations from the bivariate model between two 5-yr intervals with zero, one, or two 5-yr intervals in between, were in agreement with the estimates from the autoregressive model (Table 4.9). Both the bivariate analyses and the autoregressive models showed genetic correlations for survival between 5-yr intervals closer to unity for `functional survival` than for `survival`. Therefore, when adjusting for within-herd production level, we found little change in the survival trait over the past 25 yr.

Table 4.9. Genetic correlations between 5-yr intervals estimated with first-order autoregressive function and average genetic correlations from bivariate models for surv_12mo, surv_1st_lac, functional surv_12mo, and surv_12mo and functional surv_12mo estimated with a threshold model (SE in parentheses)¹

No. of 5-year intervals in between ²			Functional	Surv_12mo	Functional
	Surv_12mo	Surv_1st_lac	surv_12mo	(threshold)	surv_12mo (threshold)
Autoregressive					
0	0.94 (0.02)	0.92 (0.02)	0.97 (0.01)	0.93 (0.02)	0.96 (0.01)
1	0.88 (0.03)	0.85 (0.03)	0.94 (0.02)	0.86 (0.03)	0.91 (0.03)
2	0.83 (0.04)	0.79 (0.04)	0.91 (0.03)	0.80 (0.04)	0.87 (0.04)
3	0.78 (0.05)	0.73 (0.05)	0.88 (0.04)	0.74 (0.05)	0.83 (0.05)
Bivariate					
0	0.93 (0.04)	0.91 (0.04)	0.96 (0.02)	0.91 (0.04)	0.95 (0.03)
1	0.88 (0.08)	0.87 (0.07)	0.92 (0.06)	0.86 (0.08)	0.89 (0.07)
2	0.87 (0.14)	0.84 (0.14)	0.91 (0.12)	0.83 (0.15)	0.87 (0.14)
3	0.97 (0.22)	0.80 (0.31)	0.90 (0.18)	0.83 (0.27)	0.90 (0.19)

¹surv_12mo = survival until mo 12; surv_1st_lac = survival of first lactation.

²Combinations with 0 intervals between 5-yr intervals: 1989–1993 with 1994–1998, 1994–1998 with 1999–2003, 1999–2003 with 2004–2008, and 2004–2008 with 2009–2013; combinations with 1 interval between 5-yr intervals: 1989–1993 with 1999–2003, 1994–1998 with 2004–2008, and 1999–2003 with 2009–2013; combinations with 2 intervals between 5-yr intervals: 1989–1993 with 2004–2008, and 1994–1998 with 2009–2013; combinations with 2 intervals between 5-yr intervals: 1989–1993 with 2009–2013.

4.4 Discussion

In the current study, we investigated the changes in survival over the past 25 yr in the Netherlands and used genetic parameters to evaluate whether survival was a different trait according to years of first calving. Raw means for survival of first lactation increased from 79.5 to 86.6% over the past 25 yr, whereas in Australia survival of first lactation appeared to be stable around 85% between 1993 and 2010 (Haile-Mariam and Pryce, 2015); in the United States, between 1980 and 2000, survival of first lactation decreased from 77 to 72%, followed by an increase to 74% (Hare et al., 2006). Although survival to 12 mo increased in our study, the difference between surv_12mo and surv_1st_lac was rather constant, with 4 to 5% between 1989 and 2013. The genetic correlation between surv_12mo and surv_1st_lac was 0.97 when survival of all 25 yr was analyzed as 1 trait. Because of this high genetic correlation, not all results for surv_1st_lac were shown. Of all culled first-parity cows, a larger proportion was culled after 12 mo in more recent years. In Germany, it was shown that culling at the end of first lactation is mainly due to infertility, up to 50% of all culling, whereas main culling reasons until 12 mo

4 Genetic changes of survival traits

were udder diseases, claw and leg disorders, metabolic diseases, and other diseases (Heise et al., 2016). In the United States, infertility was also one of the main culling reasons in first 3 lactations (Tsuruta et al., 2015). Thus, it is likely that relatively more emphasis has gone to culling for fertility in the past 25 yr, which agrees with the decline in fertility observed over the past decades (Veerkamp et al., 2015).

Genetic parameters for different 5-yr intervals were estimated using genetic links (i.e., sires and maternal grandsires) across existing data. Different covariance structures were modeled for the additive genetic effect to investigate whether survival changed genetically across years. Genetic correlations were analyzed with bivariate models using subsets of the full data set or with parsimonious models using the full data set. Bivariate analyses with subsets of the data many years apart proved of limited value to estimate the genetic correlations, as it was difficult to conclude if genetic correlations were different from unity. Although convergence was rather fast with these models, estimated genetic correlations had large standard errors (0.09–0.31). This was not surprising, because a small number of sires were common in 5-yr intervals separated by 2 or more 5-yr intervals. For example, only 1 sire had at least 10 daughters, and 3 sires had at least 10 granddaughters in both 5-yr intervals 1989 to 1993 and 2009 to 2013. To maintain genetic connectedness between the data collected over so many years, the full data set should always be considered and not split over time. A genomic relationship matrix could improve the connectedness between sires born in different years and separated by multiple generations, as was also observed when estimating genetic correlations between countries (Berry et al., 2014) or between traits (Veerkamp et al., 2011), resulting in smaller standard errors. However, estimation of all the genetic correlations between all 5-yr intervals might still be problematic given that genetic correlations are close to unity. Genetic correlations close to unity lead to estimation problems because sampling forces variance components to be outside the parameters space, but software packages try to estimate them within the parameter space. We attempted to estimate the full genetic (co)variance matrix allowing ASReml to get outside of the parameters space, but attempts were unsuccessful; therefore, we chose to analyze the full data set with parsimonious models. Parsimonious models that made a priori assumptions about the correlation structure for survival across the five 5-yr intervals; the autoregressive models required only 1 (or 2) parameters to model the correlation structure, compared with MT_{US} that used 10 covariances to describe the full genetic (co)variance matrix. The combined effect of being able to analyze data in all five 5-yr intervals together and the simplified models was that standard errors

became smaller and convergence improved. Analyses of MT_{U5} resulted in convergence problems and nonpositive definitive matrices. Many other parsimonious models could be used. Another parsimonious model that has been used for estimating the genetic correlation for survival across many years is the random regression model (Tsuruta et al., 2004, Haile-Mariam and Pryce, 2015). A random regression model is more often applied to estimate genetic correlation over a time trajectory on the same animal [i.e., lactation curve (Schaeffer and Dekkers, 1994), weight (Meyer, 2004), or survival (Veerkamp et al., 2001, Van Pelt et al., 2015)], but can also be used to model traits in different animals in different environments (Calus and Veerkamp, 2003). A random regression model appears also very well suited to model longevity over a time span of 25 yr. However, more general solutions exist to fit parsimonious models to a set of highly correlated traits (Thompson et al., 2003, Thompson et al., 2005); these have not been attempted here, as we expect little difference from the autoregressive models.

We saw that genetic variance, residual variance, phenotypic variance, and heritability decreased over 25 yr. Because the mean survival rate, p , increased, a reduction of phenotypic variance, which is $p(1 - p)$, was expected. With an increase in survival rate it was expected that the heritability on the observed scale would decrease with the linear model when the heritability on the underlying scale was constant. However, heritability estimates from the threshold model also decreased across 5-yr intervals, albeit less for functional survival. Across all analyses, heritability in 1989 to 1993 was 2 to 5 times higher than heritability in 2009 to 2013. This continuous reduction in heritability was not found in other studies analyzing survival of first lactation (Haile-Mariam and Pryce, 2015) or productive life (Tsuruta et al., 2004). A difference with those studies was that survival of first lactation or productive life did not improve over time. Though the effects of selection are relatively slow, we considered it important to evaluate the change in genetic parameters over a longer time period; the data from the current study are also included in the genetic evaluation for longevity.

Ducrocq (1987) proposed 2 different trait definitions for survival: survival and functional survival. Survival is defined as the ability to delay culling and functional survival is defined as the ability to delay involuntary culling only. According to this definition, functional survival rather than survival is the trait that should be selected for, and it is the genetic correlation between functional survival and milk yield which is of primary interest for dairy cattle breeders (Essl, 1998). Adjustment of survival for within-herd production level provides an approximation of functional survival, and most countries participating in the Interbull evaluation for longevity (14 out of 19) accounted for production in their national evaluation for survival

4 Genetic changes of survival traits

(Forabosco et al., 2009). In the Netherlands, a genetic evaluation for functional longevity was introduced in 1999, but changed to longevity in 2008. The effect of adjusting for within-herd milk production is shown in several results in the current study. For example, for functional survival heritability was stable (~ 0.04) between 1989 and 2004 (Table 4.5), whereas the heritability for survival decreased from 0.063 to 0.034 (Table 4.3).

Models assuming high genetic correlations between the years were significantly better for functional survival, whereas they were not for survival. More convincingly, genetic correlations between the 5-yr intervals were with all models (bivariate models using subsets of the full data set, with parsimonious models using the full data set, or linear models or threshold models) lower for survival than for functional survival. For functional survival, most genetic correlations between different 5-yr intervals were ≥ 0.90 , indicating that when adjusting for within-herd milk production survival did not change across the 25 yr. For survival analyzed with a linear model for `surv_12mo` and `surv_1st_lac` or with a threshold model for `surv_12mo`, genetic correlations were < 0.90 when 5-yr intervals were separated by at least one 5-yr interval. Other studies found genetic correlations < 0.40 for survival in 2 different periods that were separated by at least 10 yr when using MT models (Tsuruta et al., 2004, Haile-Mariam and Pryce, 2015). Their standard errors of the genetic correlations were larger than in our study, and these low genetic correlations did not significantly differ from unity. Both studies used random regression fitting a Legendre polynomial of first order (intercept and linear) to estimate the genetic covariance structure. Haile-Mariam and Pryce (2015) found genetic correlations between different 2-yr intervals that were > 0.90 , and Tsuruta et al. (2004) found that most genetic correlations between breeding values for different years were > 0.70 . From these studies it was not obvious that survival changed across time, because results from different models were conflicting with each other. We concluded that the trait functional survival (i.e., involuntary culling) did not really change across years, and bulls rank more similar using functional survival in comparison with using survival. After adjusting for the deviation of individual production from herd production level, farmers were still culling animals for roughly the same reasons in, for example, 1995 and 2010. However, voluntary culling did change across years (Ducrocq, 1999), as shown by genetic parameters in our study, which agrees with phenotypic analyses performed on the same data set (Van Pelt et al., 2016). Van Pelt et al. (2016) showed that the effect of culling for within-herd production level changed in the past 25 yr, and in more recent years animals with low production had a lower risk of culling. Therefore, it is important that we consider adjusting for within-herd production level in the genetic

evaluation to have more consistent breeding values, both for selection and for setting up the reference population for genomic prediction.

4.5 Conclusions

Analyses of 25 yr of data on survival in first lactation in Dutch dairy cattle demonstrated that survival increased, whereas genetic variances and heritability decreased. Genetic correlations between 5-yr intervals showed that survival changed over time, whereas genetic correlations for functional survival did not indicate that survival changed. The difference in genetic correlations between survival and functional survival are likely explained by less emphasis of dairy farmers on culling in first lactation for production in more recent years. This suggests that genetic evaluation for longevity using historical data should analyze functional survival rather than survival.

4.6 Acknowledgments

The authors acknowledge funding from the Dutch Dairy Board (PZ; Zoetermeer, the Netherlands), Genetic Evaluation for Sires (GES) and cattle improvement cooperative CRV (Arnhem, the Netherlands) and the Breed4Food project (program “Kennisbasis Dier,” code: BO-12-22.04-011-001-ASG-LR).

4.7 References

- Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. in Proc. Proc. 2nd Int. Symp. Information Theory. Akademiai Kiado, Budapest, Hungary.
- Beaudeau, F., H. Seegers, V. Ducrocq, C. Fourichon, and N. Bareille. 2000. Effect of health disorders on culling in dairy cows: a review and a critical discussion. *Ann. Zootech.* 49:293-311.
- Bergevoet, R. H. M., C. J. M. Ondersteijn, H. W. Saatkamp, C. M. J. van Woerkum, and R. B. M. Huirne. 2004. Entrepreneurial behaviour of Dutch dairy farmers under a milk quota system: goals, objectives and attitudes. *Agric. Syst.* 80:1-21.
- Berry, D. P., M. P. Coffey, J. E. Pryce, Y. de Haas, P. Lovendahl, N. Krattenmacher, J. J. Crowley, Z. Wang, D. Spurlock, K. Weigel, K. Macdonald, and R. F. Veerkamp. 2014. International genetic evaluations for feed intake in dairy cattle through the collation of data from multiple sources. *J. Dairy Sci.* 97:3894-3905.
- Calus, M. P. L. and R. F. Veerkamp. 2003. Estimation of environmental sensitivity of genetic merit for milk production traits using a random regression model. *J. Dairy Sci.* 86:3756-3764.
- CRV. 2015. CRV Jaarstatistieken Nederland 2015. Arnhem, the Netherlands.
- Dempster, E. R. and M. Lerner. 1950. Heritability of threshold characters. *Genetics* 35:212-236.
- Ducrocq, V. 1987. An analysis of productive life in dairy cattle. Cornell University, Ithaca, New York.
- Ducrocq, V. 1999. Two years of experience with the French genetic evaluation of dairy bulls on production-adjusted longevity of their daughters. *Interbull Bulletin* 21:60-68.
- Essl, A. 1998. Longevity in dairy cattle breeding: a review. *Livest. Prod. Sci.* 57:79-89.
- Fetrow, J., K. V. Nordlund, and H. D. Norman. 2006. Invited review: Culling: Nomenclature, definitions, and recommendations. *J. Dairy Sci.* 89:1896-1905.
- Forabosco, F., J. H. Jakobsen, and W. F. Fikse. 2009. International genetic evaluation for direct longevity in dairy bulls. *J. Dairy Sci.* 92:2338-2347.
- Gilmour, A. R., B. Gogel, B. Cullis, S. Welham, and R. Thompson. 2015. ASReml user guide release 4.1 Functional specification. VSN International Ltd, Hemel Hempstead, UK.
- Haile-Mariam, M. and J. E. Pryce. 2015. Variances and correlations of milk production, fertility, longevity, and type traits over time in Australian Holstein cattle. *J. Dairy Sci.* 98:7364-7379.

- Hare, E., H. D. Norman, and J. R. Wright. 2006. Survival rates and productive herd life of dairy cattle in the United States. *J. Dairy Sci.* 89:3713-3720.
- Heise, J., Z. T. Liu, K. F. Stock, S. Rensing, F. Reinhardt, and H. Simianer. 2016. The genetic structure of longevity in dairy cows. *J. Dairy Sci.* 99:1253-1265.
- LTO. 2011. Melkveehouderij: midden in de maatschappij – Visie, LTO Nederland Vakgroep Melkveehouderij. Den Haag, the Netherlands.
- Meyer, K. 2004. Scope for a random regression model in genetic evaluation of beef cattle for growth. *Livest. Prod. Sci.* 86:69-83.
- Miglior, F., B. L. Muir, and B. J. Van Doormaal. 2005. Selection indices in Holstein cattle of various countries. *J. Dairy Sci.* 88:1255-1263.
- Robertson, A. 1966. A mathematical model of culling process in dairy cattle. *Animal Production* 8:95-108.
- Schaeffer, L. R. and J. C. M. Dekkers. 1994. Random regressions in animal models for test-day production in dairy cattle. Pages 443-446 in *Proc. Proc. 5th World Congr. Genet. Appl. Livest. Prod., Guelph, Ontario, Canada.*
- Schwarz, G. 1978. Estimating the dimension of a model. *The annals of statistics* 6:461-464.
- Thompson, R., S. Brotherstone, and I. M. S. White. 2005. Estimation of quantitative genetic parameters. *Philos. Trans. R. Soc. B-Biol. Sci.* 360:1469-1477.
- Thompson, R., B. Cullis, A. Smith, and A. Gilmour. 2003. A sparse implementation of the average information algorithm for factor analytic and reduced rank variance models. *Aust. N. Z. J. Stat.* 45:445-459.
- Tsuruta, S., D. A. L. Lourenco, I. Misztal, and T. J. Lawlor. 2015. Genotype by environment interactions on culling rates and 305-day milk yield of Holstein cows in 3 US regions. *J. Dairy Sci.* 98:5796-5805.
- Tsuruta, S., I. Misztal, and T. J. Lawlor. 2004. Genetic correlations among production, body size, udder, and productive life traits over time in Holsteins. *J. Dairy Sci.* 87:1457-1468.
- Van Arendonk, J. A. M. 1985. Studies on the replacement policies in dairy cattle. II. Optimum policy and influence of changes in production and prices. *Livest. Prod. Sci.* 13:101-121.
- Van Pelt, M. L., G. de Jong, and R. F. Veerkamp. 2016. Changes in the genetic level and the effects of age at first calving and milk production on survival during the first lactation over the last 25 years. *Animal* 10:2043-2050.
- Van Pelt, M. L., T. H. E. Meuwissen, G. de Jong, and R. F. Veerkamp. 2015. Genetic analysis of longevity in Dutch dairy cattle using random regression. *J. Dairy Sci.* 98:4117-4130.

4 Genetic changes of survival traits

- Veerkamp, R. F., S. Brotherstone, B. Engel, and T. H. E. Meuwissen. 2001. Analysis of censored survival data using random regression models. *Anim. Sci.* 72:1-10.
- Veerkamp, R. F., H. A. Mulder, R. Thompson, and M. P. L. Calus. 2011. Genomic and pedigree-based genetic parameters for scarcely recorded traits when some animals are genotyped. *J. Dairy Sci.* 94:4189-4197.
- Veerkamp, R. F., A. M. M. Tenghe, L. Kaal, and A. C. Bouwman. 2015. Genetics and genomics of fertility in dairy cows. *Cattle Pract.* 23:98-102.
- Zijlstra, J., M. Boer, J. Buiting, K. Colombijn-van der Wende, and E.-A. Andringa. 2013. Rapport 668: Routekaart Levensduur; Eindrapportage van het project "Verlenging levensduur melkvee". Wageningen UR Livestock Research, Wageningen, the Netherlands.

5

Genetic evaluation of longevity using a random regression animal model

M.L. van Pelt^{1,2}, G. de Jong¹, R.F. Veerkamp²

¹ CRV BV, Animal Evaluation Unit, PO Box 454, 6800 AL Arnhem, the Netherlands;

² Animal Breeding and Genomics Centre, Wageningen UR Livestock Research, PO
Box 338, 6700 AH Wageningen, the Netherlands

To be submitted

Abstract

Longevity is an important trait for dairy farmers, and genetic evaluation has been available since 1999 in the Netherlands and Flanders. A new model is developed here based on a random regression animal model (RRM), and the impact of this new model on the stability of estimated breeding values (EBV) for longevity was analyzed. The trait analyzed was survival per month, fitted with a fifth-order Legendre polynomial until 72 months after first calving. Three different RRM models were tested, which were stepwise improvements on the current proportional hazards model (PHM), and for functional longevity a RRM was used where survival was adjusted for within-herd production level. The stability of EBV was analyzed by checking the difference between the first EBV a bull had received since 2006 with EBV in 2016. Correlations between the first and later EBV were higher for the RRM (0.60–0.70) than for PHM (0.50), meaning that the ranking of bulls was more consistent using RRM. The standard deviation of the difference between first EBV and EBV-2016, expressed in genetic standard deviations, was similar (0.57–0.63) between the models; however, expressed in months this ranged between 4.1 and 5.5 mo, where the smallest changes were in RRM for functional longevity and the largest changes in PHM. The ranking of bulls and the stability of EBV over sequential runs appear to be better for RRM. Bias due to incomplete daughter information was observed for all models; more bias in EBV was observed with RRM, however, due to the estimation of the genetic trend. Before implementing this new model for genetic evaluation of longevity, more research is needed in order to reduce the bias in the breeding values.

Key words: longevity, random regression, genetic evaluation

5.1 Introduction

Longevity is of economic importance for dairy farmers; in animal breeding, however, selection for longevity is complex. The true longevity of a cow is available only at the end of her life, whereas selection and breeding decisions are made earlier in life. Therefore, censored data is used in genetic evaluations, and expected lifespan is extrapolated from survival data during early life. Various genetic evaluation models are used for the analysis of censored data (Forabosco et al., 2009), and survival analysis using a proportional hazard model (**PHM**) has been popular in genetic evaluations since the software package Survival Kit was introduced (Ducrocq and Sölkner, 1994, 1998, Ducrocq et al., 2010). This package has been implemented for routine genetic evaluation in the Netherlands since August 1999. Until 2007, breeding values for functional longevity predicted survival adjusted for within-herd production level (Vollema et al., 2000); however, from 2008 onwards no adjustment was made for within-herd production level in genetic evaluation (Van der Linde et al., 2007).

In recent years, practical experience has shown that the breeding values for longevity fluctuate greatly between runs, more than could be expected from the change in reliability of the breeding value between runs. One hypothesis put forward for this fluctuation was that the current genetic evaluation model used in the Netherlands was too simplistic a model for the genetic variation in longevity over the life of a cow. Veerkamp et al. (2001) had previously proposed the random regression model (**RRM**) for survival analysis. The advantage of the RRM is that multiple genetic effects can be fitted over time, rather than assuming one genetic effect over the life of an animal; non-unity genetic correlations for survival using different time intervals and heterogeneous genetic variances can thus be modeled over the life of a cow. The RRM was further investigated by Jamrozik et al. (2008), Van Pelt et al. (2015), and Sasaki et al. (2015). Van Pelt et al. (2015) examined the importance of fitting different genetic effects in Dutch Holstein, and found genetic correlations of between 0.25 and 1.00 for survival at different time intervals over the lives of cows, showing that survival was not the same genetic trait over the life of a cow. In other countries, genetic correlations for survival between parities have been estimated in the range 0.33 to 0.96, with an average of 0.77 (Visscher and Goddard, 1995, Brotherstone et al., 1997, Boettcher et al., 1999, Veerkamp et al., 2001, Gengler et al., 2005, Sewalem et al., 2007, Holtsmark et al., 2009).

Another reason for breeding values fluctuating more than expected may be changes in fixed effects over the long time periods in which data is collected; time-dependent fixed effects are required to be calculated by fitting an interaction

between the fixed effect and a time variable. Van Pelt et al. (2016a) demonstrated that survival at first lactation was affected differently by age at first calving (**AFC**) and within-herd production level in the mid-1990s, compared with later decades. AFC and within-herd production, and the interaction between these, have become less important in recent years. This change in culling reasons also affected the ranking of bulls across the 25 years of data (Van Pelt et al., 2016b). Genetic correlations for survival at first lactation measured more than 10 years apart were below 0.90 when within-herd production level was not accounted for. However, when within-herd production level was fitted in the model, genetic correlations were ≥ 0.90 . This suggested that a genetic evaluation for longevity using historical data should take account of the changes in fixed effects over time; in particular, the change in emphasis on milk production suggested that functional survival (adjusted for milk production) should be analyzed rather than survival without adjustment for milk production, at least when the genetic effects are modeled as being constant over time.

The studies above have given rise to the development and investigation of a new prototype for the genetic evaluation of longevity in the Netherlands. The aim of this research was to evaluate the impact of these changes to the model by: 1) fitting multiple genetic effects across the life of a cow; 2) fitting time-dependent fixed effects; and 3) adjusting for milk production on the breeding values for longevity of bulls. Thus, breeding values reflect the true breeding value at an earlier stage in the life of the bull; breeding values therefore become more stable, and the new random regression animal model can be compared with the current genetic evaluation using a PHM.

5.2 Material and Methods

5.2.1 Data

Longevity or length of productive life was defined as the time from first calving to the last test date for milk production before the animal died or was culled for slaughter; this also includes dry periods. For the PHM, total length of productive life was analyzed; for the RRM, the analysis period was length of productive life until 72 mo after first calving.

Data were available from the Dutch/Flemish cattle improvement cooperative CRV (CRV, Arnhem, the Netherlands). The dataset was constructed from records of pedigree, lactations and movements of cows in the Netherlands. Herdbook-registered cows from a dairy breed with a test-day record on or after January 1, 1988 were included. Data up to February 15, 2016 were used, and this study

therefore covered 28 years. Cows were required to be at least 640 d old at first calving. If the date of first calving was unknown, it was assumed that a cow first calved at 800 d of age, the average of the total data set. If the first calving of a cow took place before the starting date of the study, the record was considered to be left-truncated. Records of cows that were still alive at the time of data collection were considered to be right-censored. Records of cows that were moved to another milking herd were also considered to be right-censored, if this herd was not participating in the milk recording scheme.

Records for the RRM were constructed for each month a cow was present in a herd, from first calving up to the month the cow was culled, or 72 mo, or when the cow was censored. A cow culled in month j has $j - 1$ records with score 100 (alive), and record j with score 0 (culled). Monthly records were treated as missing after culling. Additional procedure for RRM was that cows with an AFC of > 40 mo were deleted. The total dataset comprised 357,140,952 records from 10,788,976 animals in 48,272 herds.

The dataset for the PHM contained 10,553,223 animals. The number of animals for PHM was lower than that for RRM since bulls needed to have at least 10 daughters to be included in the dataset. In contrast to the national genetic evaluation no voluntary waiting period of 270 days was applied; the information was included at a time when the cow could have been productive for at least 270 days after the first calving. The voluntary waiting period is applied in the national evaluation as a short-term fix to reduce bias in breeding values. In order to compare PHM and RRM using similar selection criteria, the voluntary waiting period was not applied for PHM.

5.2.2 Statistical Model

Current Proportional Hazard Model. The current model used in the genetic evaluation for longevity is a proportional hazard model, as described in (Van der Linde et al., 2004, Van der Linde et al., 2007). The piecewise Weibull model used to analyze longevity was:

$$h(t) = h_{0,pq}(\tau) * \exp \left\{ \sum_m f_m(t) + hys_i + s_j + 0.5 mgs_k + gg_l \right\},$$

where the hazard function $h(t)$ at time t was modeled with a piecewise Weibull baseline hazard function of general form $h_{0,cp}(\tau) = \lambda\rho(\lambda\tau)^{\rho-1}$ with scale parameter λ and shape parameter ρ . Since the risk of being culled differed between lactations and at various stages of lactation, different baselines (i.e., different scale and shape parameters) were defined for each combination of lactation p (1, 2, ≥ 3) and stage of lactation q (1 to 6), giving 18 different baseline

hazard functions. Changes to the hazard function were assumed to occur at $\tau = 0, 30, 190, 250, 330$, and the dry date of each consecutive lactation, where τ denotes days since last calving. The time scale τ is needed since values of t at the second and later calvings differ between cows; $\sum_m f_m(t)$ is the sum of fixed effects which increase or decrease the instantaneous culling risk. Fixed effects included in the model were parity-year-two-month (this two-month class effect is time-dependent, changing on the 15th of each odd month, starting in January 1988 for 10 parity classes), change in herd size (calculated as the number of cows on January 1 of one year minus the number of cows on January 1 of the next year, divided by the number of cows on January 1 of the next year. Seven classes are distinguished: shrinkage between 90% and 50%; shrinkage between 50% and 30%; shrinkage between 30% and 10%; neither shrinkage nor growth over 10%; growth between 10% and 30%; growth over 30%; and herds that were terminated, i.e., more than 90% shrinkage), AFC (classes of 15 to 30 days, depending on the number of animals per class), heterosis (10 classes in steps of 12.5%: Class 1 for animals without heterosis, and Class 10 for animals with 100% heterosis), and recombination (10 classes in steps of 6.25%: Class 1 for animals without recombination, and Class 10 for animals with 50% recombination). Random effects included in the model were herd-year-season i (assumed to follow a log-gamma distribution with a fixed parameter γ , and with changes occurring at April 1 and October 1 of each year or when a cow is transferred to another herd), sire j + maternal grandsire k + genetic group l (representing the additive genetic contribution of sire j , maternal grand sire k , and genetic group l of the maternal granddam of the cow).

New Random Regression Animal Model. Suggested improvements for the genetic evaluation were the incorporation of non-unity genetic correlations over the life of a cow, the fitting of time-dependent fixed effects for AFC and adjustment for within-herd production level. All suggested improvements were tested using a full RRM (**RR-full**), and three alternative RRMs (**RR-1 to RR-3**). In RR-full, functional survival was analyzed; in RR-1, survival was analyzed; in RR-2, several time-dependencies were excluded compared with RR-1; and in RR-3, a unity genetic correlation between all months was modeled. RR-full was:

$$y_{klmn} = FIXED + \sum_{i=0}^5 hc_{ik} t_l + \sum_{i=0}^5 animal_{im} t_l + \varepsilon_{klmn} ,$$

where y is the observation for the survival of animal m in a herd x five-year period k in month l (periods of five years were created by counting backwards from the last test date; if the oldest five-year class contained less than three years of data, the

5 Genetic evaluation of longevity

class was merged with the next five-year class in time); FIXED included fixed effects for herd-year-season of observation (seasons were winter, 1 January to 31 March; spring, 1 April to 30 June; summer, 1 July to 30 September; autumn, 1 October to 31 December), parity x month in lactation x five-year class (for month in lactation classes of 1 mo between 1 and 17 mo; ≥ 18 mo; and a separate class for dry periods), parity x year x season of calving, year x month of observation, AFC x five-year (classes of 1 mo between 21 and 34 mo, and ≥ 35 mo), parity x within-herd production level x five-year (within-herd production level was a ranking of animals within a herd by five-year interval for predicted or realized age-corrected 305-day yield of combined kg fat and protein. Animals were ranked into seven classes from worst to best, with (1) 1% to 5%, (2) 6% to 20%, (3) 21% to 40%, (4) 41% to 60%, (5) 61% to 80%, (6) 81% to 95% and (7) 96% to 100%), heterosis and recombination; hc_{ik} is the i th-order random regression coefficient of the fifth-order Legendre polynomial in time interval t_i for the herd curve effect of herd x five-year period k ; $animal_{im}$ is the i th-order random regression coefficient of the fifth-order Legendre polynomial in time interval t_i for the additive genetic effect of animal m ; and ϵ_{klmn} is the residual effect. Residual variances were fitted for each month (mo 1 to 72) separate.

In RR-full the fixed effect within-herd production level was fitted in the model to correct for culling due to low production, which is assumed to be the major source of voluntary culling. To compare the stability of breeding values of RR-full with PHM, the statistical model of RR-full was changed in three steps (RR-1 to RR-3) to observe which change had the highest impact in obtaining more stable breeding values. In RR-1, compared with RR-full, the fixed effect within-herd production level was excluded from the model. In RR-2, compared with RR-1, the fixed effects for AFC and parity x month in lactation were made time-independent by removing the interaction with five-year interval. In RR-3, compared with RR-2, a (close to) unity genetic correlation was fitted between all 72 mo, to mimic one genetic effect over the lifetime of a cow, but allowing for heterogeneous genetic variances. With these three adjustments, RR-3 was most similar to the PHM, mainly due to the use of genetic correlations close to unity between all 72 mo and the similarity of the fixed effects used in the model.

5.2.3 Genetic Parameters

Variance components were estimated for RRM using the procedure described by Van Pelt et al. (2015) with three more years of data and a fifth-order Legendre polynomial for the genetic effect and for the effect of herd curve. Functional longevity was modelled in RR-full, and within-herd production level was fitted as an

extra effect compared to analyzing longevity in RR-1 to RR-3. The data set for the parameter estimation consisted of 71,404 cows with 2,565,405 observations for monthly survival from January 1988 to December 2015, originating from 89 herds and 711 sires and maternal grandsires; where sires had (i.e. as sire or as maternal grandsire or combined) at least 20 progeny that could have been productive for at least 36 months after first calving, (2) sires had progeny in at least two herds and (3) every herd-year-month class had at least 45 observations. Pedigree information of sires was traced back 6 generations, resulting in a pedigree file with 1,040 sires. Variance components were estimated with ASReml (Gilmour et al., 2009). The partial derivatives of a profit function were used to transform variance components on the survival scale to those for longevity, using the phenotypic means for survival per month (Van Pelt et al., 2015).

5.2.4 Breeding Value for Longevity

Survival rate per month, or hazard rate, was analyzed in the RRM. The total estimated breeding value (EBV) for longevity can be written as an index with 72 monthly EBVs for survival in that month (Van Pelt et al., 2015). The index predicting longevity is: $I_{LON} = v_1g_1 + v_2g_2 + \dots + v_{71}g_{71} + v_{72}g_{72}$, where v_i is the weight month i , and g_i is the genetic value for surviving month i . Equivalent to the derivation of the economic weights in a non-linear profit equation, the weighting factors \mathbf{v} for each \mathbf{g} were derived from partial derivatives from the equation using mean survival per month after first calving up to 72 mo. If $LON = p_1 + p_1p_2 + p_1p_2p_3$, then $v_1 = 1 + p_2 + p_2p_3$, $v_2 = p_1 + p_1p_3$, and $v_3 = p_1p_2$, where p_i is the survival in month i . Genetic variance of longevity over 72 mo was calculated using $\mathbf{v}'\Phi'\mathbf{G}\Phi\mathbf{v}$, where Φ is an $k \times l$ matrix whose elements are the Legendre coefficients for k time intervals and order l of the Legendre polynomial, \mathbf{G} is an $l \times l$ matrix containing the estimated covariance components for the additive genetic effect, and \mathbf{v} is a vector with k partial derivatives.

5.2.5 Calculation of Reliability

In the RRM, the reliabilities of the EBV for survival in each month were estimated using the method of Liu et al. (2004). This reliability approximation method is based on the concept of the multiple-trait effective daughter contribution for the RRM. The reliability of the EBV of (functional) longevity was derived using the weights in \mathbf{v} in a selection index. The reliability of the EBV in PHM was calculated using the method of (Yazdi et al., 2002), and the number of observations was the number of culled daughters + 0.5 x number of culled maternal granddaughters (Meyer, 1989).

5.2.6 Validation

In total, five models were examined: PHM, RR-full, RR-1, RR-2, and RR-3. To evaluate the stability of breeding values for bulls, subsets of data for all five models were used in the genetic evaluations by deleting the most recent years. The national dataset contained data up to February 15, 2016. First a genetic evaluation was carried out using all data (full run, EBV-2016). Following this, 10 historic runs were carried out by deleting one year of data for each subsequent run. The first historic run involved data from January 1, 1988 to February 15, 2015 (EBV-2015); the tenth run involved data to February 15, 2006, (EBV-2006). Per model, a bull could have a maximum of 11 breeding values for longevity, one from each evaluation (i.e., the full run and 10 historic runs), depending on when the bull received the first proof based on progeny.

For black and white Holstein bulls born since 2000 and tested in the Netherlands and Flanders, the first EBV for both PHM and RRM was defined in the evaluation run where ≥ 15 daughters were at 4 mo after first calving and ≤ 50 daughters were ≤ 18 mo after first calving. It should be noted that currently for PHM a waiting period of 9 mo is applied before including an animal, and therefore a bull with 15 daughters at 4 mo after first calving will not yet receive an EBV. For comparing PHM with RRM, these requirements were set to define first EBV from all models. We defined test bulls and proven bulls for the validation. Test bulls were defined as bulls with ≤ 250 daughters in any of the evaluation runs, and proven bulls as those with > 250 daughters. Most of the test bulls only had daughters born in a single year, and these form the so-called first crop. The majority of proven bulls first had an EBV based on first-crop daughters; when these were used again, new batches of daughters were born, the so-called second crop. The mean change in EBV, the standard deviation of the mean change in EBV, and the correlation between EBVs of sequential runs were calculated separately for test bulls and proven bulls. These runs were not sequential in terms of yearly evaluations, but represented the sequential run after the first evaluation of a bull.

5.3 Results

5.3.1 Genetic Parameters

Estimates of genetic parameters for longevity for all five models are presented in Table 5.1. For PHM, genetic parameters from the current national evaluation were used, with a heritability of 0.12 and genetic standard deviation of 9.0 mo. For RRM, the heritability of longevity ranges between 0.12 and 0.15, and genetic standard deviations between 7.1 and 8.0 mo. RR-full analyzed functional longevity, whereas

the other RRM models analyzed longevity, i.e., not corrected for within-herd production level. RR-full showed the smallest genetic standard deviation, with 7.1 mo. RR-3 showed genetic correlations close to unity over all 72 mo, and PHM modeled one genetic effect; both of these models showed the largest genetic standard deviation (8.0 and 9.0 mo). Correcting for within-herd production level resulted in different genetic correlations for survival between months as well as a lower genetic standard deviation. For longevity, genetic correlations between months ranged between 0.51 and 1.00, and for functional longevity between 0.06 and 1.00 (Table 5.2). Genetic correlations between longevity and cumulative time intervals up to 6 mo after first calving were comparable, with 0.78 for longevity and 0.80 for functional longevity (Figure 5.1). However, the genetic correlation with longevity was >0.95 when records were available after 30 mo or more. For functional longevity, the genetic correlation was >0.95 only after 42 mo.

Table 5.1. Heritability (h^2) and genetic standard deviation (σ_g) in months of longevity for five different models (RR-full, RR-1, RR-2, RR-3, and PHM)

Model	h^2	σ_g
RR-full	0.12	7.1
RR-1 and RR-2	0.14	7.6
RR-3	0.15	8.0
PHM	0.12	9.0

Table 5.2. Genetic correlations between one-month intervals for survival, estimated using a random regression model with a fifth-order Legendre polynomial for the genetic effect for longevity (below diagonal) and functional longevity (above diagonal)

Mo	6	12	18	24	30	36	42	48	54	60	66	72
6		0.96	0.85	0.70	0.58	0.49	0.39	0.29	0.19	0.11	0.07	0.06
12	0.92		0.90	0.73	0.63	0.56	0.48	0.38	0.27	0.18	0.13	0.14
18	0.89	0.96		0.95	0.89	0.84	0.78	0.71	0.63	0.56	0.52	0.49
24	0.86	0.91	0.98		0.98	0.95	0.91	0.86	0.81	0.76	0.73	0.69
30	0.84	0.87	0.96	0.99		0.99	0.97	0.93	0.89	0.84	0.81	0.80
36	0.81	0.85	0.94	0.98	0.99		0.99	0.97	0.93	0.88	0.86	0.87
42	0.77	0.82	0.91	0.95	0.98	1.00		0.99	0.96	0.91	0.89	0.92
48	0.71	0.78	0.88	0.93	0.96	0.98	0.99		0.99	0.95	0.94	0.96
54	0.65	0.72	0.85	0.91	0.94	0.96	0.97	0.99		0.99	0.98	0.97
60	0.59	0.65	0.80	0.87	0.90	0.92	0.94	0.96	0.99		1.00	0.95
66	0.55	0.59	0.75	0.83	0.87	0.90	0.92	0.95	0.98	1.00		0.96
72	0.51	0.53	0.65	0.74	0.81	0.86	0.90	0.92	0.93	0.92	0.94	

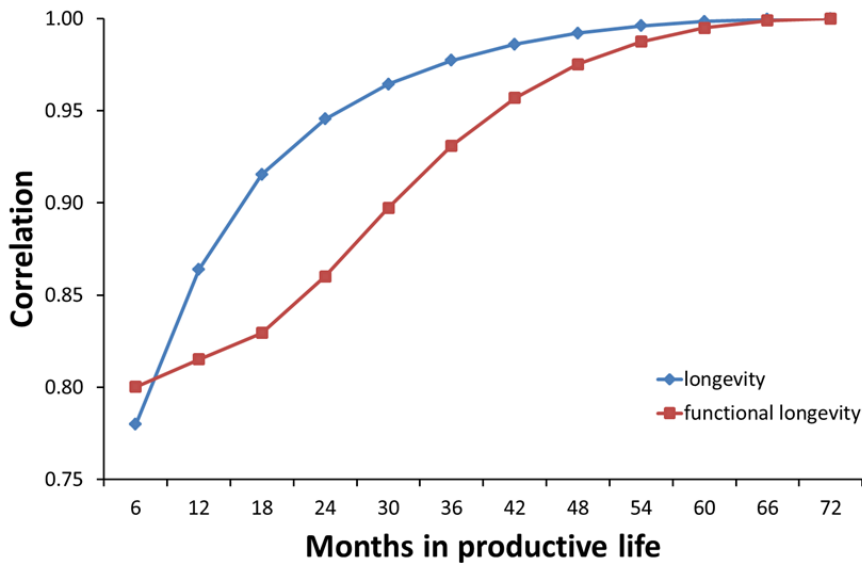


Figure 5.1. Genetic correlations between survival over cumulative intervals (in steps of 6 months) and EBV for longevity (RR-1 and RR-2) and EBV for functional longevity (RR-full).

5.3.2 Reliabilities

Number of bulls, number of progeny and the reliability per model are shown in Table 5.3 for test bulls and in Table 5.4 for proven bulls from the first to 11th EBV. Selected bulls were born between 2000 and 2010, and all 3,149 (432) test (proven) bulls had at least the first EBV, while 432 (45) test (proven) bulls had 11 EBVs. First EBV was based on an average of 108 daughters for test bulls and 175 daughters for proven bulls, and reliability ranged between 42% and 53%. Mean reliability for test bulls ranged from 76% to 79% for the sixth EBV, while proven bulls with a larger progeny group had a mean reliability between 87% and 91%. When bulls had 11 EBVs, the mean reliability increased further, to 80%–83% for test bulls and ≥95% for proven bulls.

5 Genetic evaluation of longevity

Table 5.3. Number of bulls, number of daughters, and reliability of EBV for RR-full, RR-1 to RR-3, and PHM for test bulls from first to 11th EBV

EBV no.	No. of bulls	No. of daughters	Reliability (%)			
			RR-full	RR-1 and RR-2	RR-3	PHM
1 st	3,149	108	43	44	49	42
2 nd	2,924	124	54	57	61	53
3 rd	2,719	126	65	68	70	63
4 th	2,551	126	72	74	75	69
5 th	2,340	127	75	77	78	73
6 th	2,081	127	76	78	79	76
7 th	1,805	127	77	79	80	77
8 th	1,488	129	78	79	81	79
9 th	1,135	133	78	80	81	80
10 th	779	142	80	81	82	82
11 th	432	149	80	82	83	83

Table 5.4. Number of bulls, number of daughters, and reliability of EBV for RR-full, RR-1 to RR-3, and PHM for proven bulls from first to 11th EBV

EBV no.	No. of bulls	No. of daughters	Reliability (%)			
			RR-full	RR-1 and RR-2	RR-3	PHM
1 st	271	175	46	47	53	43
2 nd	264	349	59	63	69	57
3 rd	245	374	71	74	77	68
4 th	210	537	79	81	83	75
5 th	181	1,336	84	86	88	82
6 th	157	2,579	88	90	91	87
7 th	132	3,779	91	92	93	90
8 th	105	3,870	93	94	94	92
9 th	85	4,392	94	95	95	94
10 th	66	4,765	94	95	95	94
11 th	45	5,049	95	95	96	95

5.3.3 Correlations between Models

Correlations between the EBV-2016 from the five models are shown in Table 5.5. RR-full had the lowest correlations with the EBV from the other four models, ranging from 0.73 to 0.77, since RR-full gave the EBV adjusted for within-herd milk production (functional longevity) whereas the other models gave the EBV for longevity. PHM had correlations with RR-1 to RR-3 of 0.93. RR-1 to RR-3 correlated very highly (≥ 0.99).

Table 5.5. Correlations between EBV from current PHM and four different RRM

Model	RR-full	RR-1	RR-2	RR-3
RR-1	0.77			
RR-2	0.75	1.00		
RR-3	0.75	0.99	1.00	
PHM	0.73	0.93	0.93	0.93

5.3.4 Correlations between first EBV and later EBV for test Bulls and proven Bulls

The correlation between first EBV and later EBVs are shown per model in Table 5.6 for test bulls (with only first-crop information) and in Table 5.7 for proven bulls (with first- and second-crop information). A total of 2,924 test bulls, born between 2000 and 2010, had more than 15 daughters in two genetic evaluations, and 432 test bulls born in 2000 and 2001 had 11 EBVs. The calculation of the correlations between first EBV and later EBVs was based on a different number of bulls, depending on the nth EBV compared with the first EBV. The correlation between the first and second EBV of a bull ranged between 0.79 for PHM and 0.86 for RR-full. For later EBVs, more information on daughters was included in the last evaluation, and the correlation decreased to ~ 0.50 for PHM in the eighth to the 10th EBV, ~ 0.60 for RR-1 to RR-3 in the sixth to the 10th EBV, and ~ 0.67 for RR-full in the fifth to the 10th EBV. For test bulls, the correlation did not drop further after the fifth evaluation with the RRM. For all models, the correlation between first EBV and 11th EBV increased compared to the 10th EBV; however, only two birth years of bulls were included in the calculation of this correlation. For proven bulls, the drop in correlation between first EBV and later EBV followed a similar pattern to that of the test bulls up to the sixth EBV; however, when the information on the second-crop daughters was included in the EBV of proven bulls, the correlation continued to decline. A comparison of the first with the 11th EBV correlation was lowest for PHM (0.18), and highest for RR-full (0.43).

Table 5.6. Correlations between first EBV and later EBV (second to 11th) for five models for test bulls (born between 2000 and 2010, with fewer than 250 daughters)

EBV no.	RR-full	RR-1	RR-2	RR-3	PHM
2 nd	0.86	0.83	0.83	0.82	0.79
3 rd	0.75	0.71	0.71	0.71	0.68
4 th	0.69	0.66	0.65	0.66	0.62
5 th	0.67	0.63	0.63	0.63	0.58
6 th	0.66	0.62	0.62	0.61	0.55
7 th	0.67	0.61	0.60	0.60	0.53
8 th	0.66	0.59	0.59	0.58	0.50
9 th	0.68	0.60	0.60	0.59	0.50
10 th	0.68	0.62	0.62	0.61	0.49
11 th	0.72	0.65	0.65	0.64	0.55

Table 5.7. Correlations between first EBV and later EBVs (second to 11th) for five models for proven bulls (born between 2000 and 2010, with 250 daughters or more)

EBV no.	RR-full	RR-1	RR-2	RR-3	PHM
2 nd	0.88	0.83	0.83	0.81	0.75
3 rd	0.78	0.73	0.73	0.72	0.65
4 th	0.70	0.67	0.68	0.67	0.60
5 th	0.64	0.64	0.65	0.64	0.54
6 th	0.60	0.61	0.61	0.60	0.54
7 th	0.56	0.57	0.57	0.57	0.54
8 th	0.54	0.51	0.51	0.49	0.41
9 th	0.44	0.41	0.42	0.38	0.28
10 th	0.41	0.31	0.31	0.26	0.14
11 th	0.43	0.35	0.37	0.28	0.18

5.3.5 Difference between first EBV and all other EBV

The differences between EBV-2016 and all previous EBVs of a bull are shown in Figure 5.2 for test bulls and in Figure 5.3 for proven bulls. For all five models, the EBVs of earlier runs were overestimated compared to EBV-2016. The smallest difference in terms of genetic standard deviations (gen.sd) from EBV-2016 was for PHM (0.31 gen.sd for test bulls and 0.25 gen.sd for proven bulls), and the largest difference from EBV-2016 was for RR-3 (0.53 gen.sd for test bulls and 0.50 gen.sd for proven bulls). Expressed on the absolute scale in months, the smallest difference for test bulls was for RR-full (2.5 mo) and the largest for RR-3 (4.0 mo). For proven bulls, the smallest difference was for PHM (2.2 mo) and the largest difference for RR-3 (3.8 mo). The difference from EBV-2016 decreased when more information was added, and the difference of the fourth EBV from EBV-2016 was

5 Genetic evaluation of longevity

0.10–0.20 gen.sd. When test bulls had the sixth or seventh EBV, the information on the daughters should be complete, since most daughters have at this point been productive for 72 mo; in addition, the correlation with the final proof became constant (Table 5.6). For PHM, the difference from EBV-2016 was close to zero for the seventh and later EBVs, whereas the mean EBV in all RRM models continually decreased.

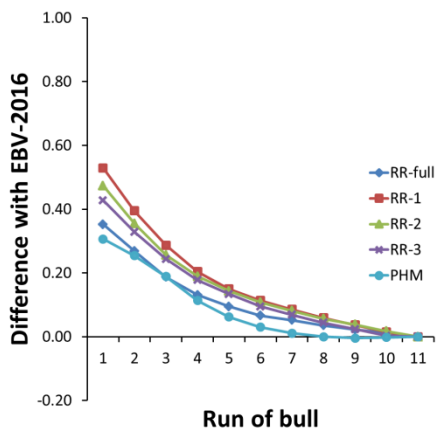


Figure 5.2. Mean difference of EBV in gen.sd between EBV-*i* (*i*=2006–2015) and EBV-2016 for test bulls.

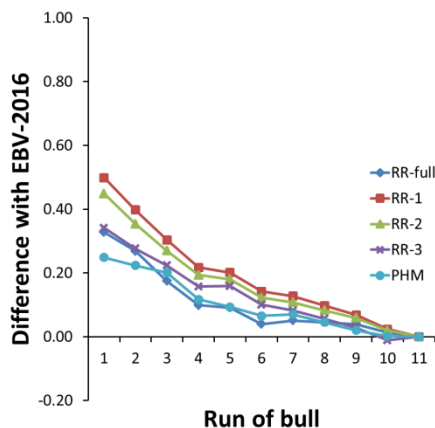


Figure 5.3. Mean difference of EBV in gen.sd between EBV-*i* (*i*=2006–2015) and EBV-2016 for proven bulls.

A measure of the variability in the EBV for bulls (i.e., the stability) is the standard deviation of the differences between EBV-2016 and previous EBVs, as shown in Figure 5.4 for test bulls and Figure 5.5 for proven bulls. Overall, RR-full showed the lowest standard deviation of differences, and PHM the largest, meaning that EBV for bulls changed less with RR-full than with PHM.

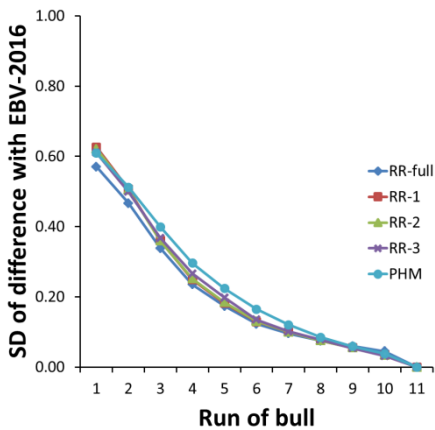


Figure 5.4. Standard deviation of differences of EBV between EBV- i ($i=2006-2015$) and EBV-2016 for test bulls.

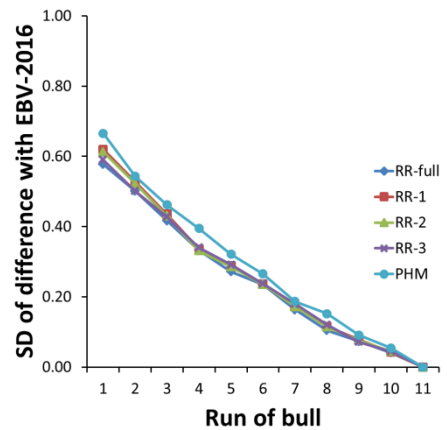


Figure 5.5. Standard deviation of difference of EBV between EBV- i ($i=2006-2015$) and EBV-2016 for proven bulls.

Genetic trends of bulls for four different runs using data up to February 15, 2007, 2010, 2013, and 2016 are shown for PHM, RR-3, and RR-full in Figures 5.6 to 5.8. The genetic trends of RR-1 and RR-2 were almost identical to those of RR-3, and are not presented. A sharp increase can be observed in the genetic trend for the bulls that received their first EBV; however, with later runs this sharp increase disappeared. For example, for RR-3 for run EBV-2010, the EBVs of bulls born in 2005 were 0.54 sd higher than bulls born in 2004, and with EBV-2016 this increase was reduced to 0.17 sd. The decline in mean EBV from first to later EBVs was observed for all models. For example, mean EBV for bulls born in 1997 was reduced by 0.24 sd for RR-3 and 0.14 sd for RR-full, and only differed by 0.01 sd for PHM. The difference in genetic standard deviations based on EBV-2016 between bulls born in 1985 and 2011 was 1.80 sd (16.2 mo) for PHM, 2.60 sd (19.8 mo) for RR3, and 0.95 sd (6.8 mo) for RR-full. All models showed that longevity has improved genetically; however, the responses were not equal in terms of genetic standard deviations, and there appears to be a strong bias in the youngest generation of bulls.

5 Genetic evaluation of longevity

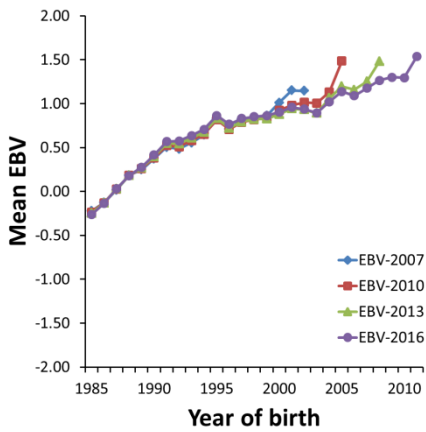


Figure 5.6. Estimates of genetic trend for longevity with PHM from runs with data up to 15 February 2007, 2010, 2013 and 2016 for Holstein bulls born between 1985 and 2011.

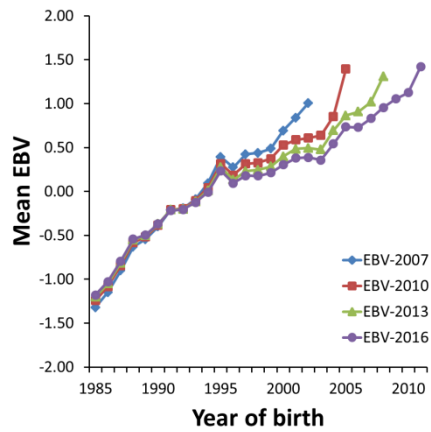


Figure 5.7. Estimates of genetic trend for longevity with RR-3 from runs with data up to 15 February 2007, 2010, 2013 and 2016 for Holstein bulls born between 1985 and 2011.

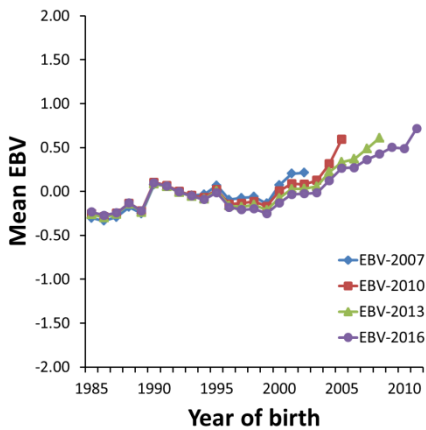


Figure 5.8. Estimates of genetic trend for functional longevity with RR-full from runs with data up to 15 February 2007, 2010, 2013 and 2016 for Holstein bulls born between 1985 and 2011.

5.4 Discussion

A random regression animal model was developed for the genetic evaluation of longevity, and breeding values for longevity were compared with the current genetic evaluation using a PHM. The differences from PHM were that in the RRM, multiple genetic effects over the life of a cow were fitted, extra time-dependent fixed effects were included, and an adjustment for milk production was included in the model. We expected that these changes would improve the breeding values over time, i.e., they would become closer to the final breeding values more quickly when all the information was available. The results showed that the correlations between first and later EBV were higher for the RRM than for PHM, meaning that the ranking of bulls was more consistent using RRM. The EBVs were biased initially for both RRM and PHM. In an unbiased breeding value estimation, the mean change in EBV between sequential runs for a group of bulls should be zero; our results showed a decrease in the mean EBV over sequential runs and, as a group, the EBVs of the bulls dropped further over time using RRM. When investigating individual bulls, the variability in EBV (i.e., the difference between the first EBV and EBV-2016) expressed in genetic standard deviation units was similar (0.57–0.63) for all models. The variability expressed in months ranged between 4.1 and 5.5 mo, with the smallest variation for RR-full and the largest for PHM. In terms of the ranking of bulls, the stability of individual bulls' EBVs over sequential runs appear to be better for RR-full and to a lesser extent for RR-1 to RR-3; however, as a group the EBVs of bulls fell over time.

In the current genetic evaluation for longevity in the Netherlands and Flanders a waiting period of 9 mo after first calving is applied to include data of a cow. However, in this study was chosen not to apply this waiting period, and investigate how EBV evolve for both PHM and RRM based on similar selection criteria. The correlations in Table 5.6 and 5.7 gave the impression that the first published EBV are not a very good reflection of the EBV including most available information; in practice, the users never see these EBV at this stage. Applying a voluntary waiting period before including data helps to improve the correlation between first published EBV, and a longer waiting period will help to improve this correlation (Figure 5.1), and also reduce bias in EBV.

The bias, as presented in Figures 5.2 and 5.3, can be divided into two types. The first type of bias occurs when more information is added; that is, existing daughters are getting older, more daughters are becoming available or a combination of both of these effects, providing more information for the total breeding value for 72 mo of survival. This bias is seen in both PHM and RRM, and occurs until the seventh

5 Genetic evaluation of longevity

EBV; at the seventh EBV, all first-crop daughters have been productive for 72 mo, and therefore nearly all information is available at this point. The second type is a bias in the genetic trend. For RRM, the mean EBV was still falling after the seventh EBV, and this bias was also visible in the genetic trends between the sequential genetic evaluations presented for RR-full and RR-3. However, for PHM, the bias in genetic trend seemed to be smaller, as the mean change in EBV with the eighth and later EBVs was close to zero, and genetic trends between sequential genetic evaluations overlapped.

The first type of bias, arising from the addition of more information, contributes the most to the total bias; this effect was shown for both PHM and RRM. This bias in EBV, expressed in genetic standard deviations, was larger using RRM as compared with using PHM. The bias was greatly reduced, to the level of PHM, when functional longevity was estimated with RRM. To identify the factors that may be causing differences in EBV, RR-full was compared with PHM, as well as three models that changed in steps from RR-full towards PHM. We expected that fitting a unity genetic correlation between all months (RR-3) would be close to modeling a single genetic effect, as in PHM. However, the results from RR-1 to RR-3 were nearly identical, and all models overestimated first EBV. Hence, the first type of bias appears to be intrinsic to the data, and none of the models takes this into account.

The second bias, arising from the genetic trend, is stronger for RRM than for PHM. There were certain differences between RR-3 and PHM: heterogeneous genetic and residual variances were fitted in RR-3, whereas in PHM one genetic and residual variance is fitted, and an animal model was used for all RRM whereas for PHM a sire-MGS model was used. With the sire-MGS model, only 31.25% of the animal variation is explained; with the animal model, the complete animal effect is explained including Mendelian sampling (Jenko et al., 2013). Moreover, a curve for the genetic effect was estimated for each animal with RRM; this may be somewhat ambitious for a binary trait, since an animal culled in month j has $j-1$ records with score 100 (alive) and only record j with score 0 (culled). Young animals will predominantly have a score of 100. In a sire model, the average of the progeny is used to estimate a survival curve, and it may therefore be easier to estimate and give more consistent breeding values based on young animals. For this reason, a more complete validation of the model should also investigate the EBV of cows. The way in which the EBV of a cow evolves from being alive to being culled is of interest. A stepwise test should be carried out to determine whether the difference in bias in EBV between RR-3 and PHM is related to fitting an animal model rather

than a sire-MGS model, or whether this is related to fitting heterogeneous genetic variances rather than a single genetic variance.

Another explanation for the bias may be that the parent average is overestimated, or that the parent average is not a good starting point for the EBV of offspring. The parent average makes a large contribution to the first EBV of a bull, and is reduced when more information is added to the EBV. A quick method of testing whether the dams of bulls are overestimated would be to set the dam as 'missing' in the pedigree, or to fit a sire-MGS model (as in PHM). The bias of bull dams is a well-known phenomenon in dairy cattle breeding; however, in terms of longevity there may be an additional mechanism behind this phenomenon. Genetic selection in dairy cattle has been primarily focused on milk production, and longevity has also improved genetically as a correlated response. In addition, from a phenotypical perspective, milk production has been shown to have a large effect on survival at first lactation (Van Pelt et al., 2016a). If the dam has good milk production, she has a higher probability of surviving than her contemporaries. However, when she is tested among the contemporaries of her daughters (the next generation), her milk production will be lower than these contemporaries; therefore, she may have a lower probability of survival. However, in the parent average, the effect of her higher milk production in the previous generation is included. The same holds for the other voluntary selection criteria that farmers use to cull animals, and where there is a genetic trend in the population. It is also demonstrated here that the inclusion of within-herd production level (RR-Full) reduces this bias. However, test bulls were still overestimated by 0.35 sd in their first EBV compared to EBV-2016 with a correction for milk production. It may therefore be necessary to model farmers' preferences for certain specific characteristics in order to overcome the bias in breeding values for longevity.

The correlation between the first EBV and > sixth EBV was expected to be around 0.80, based on the genetic parameters estimated by Van Pelt et al. (2015). In addition, survival in the first year after first calving was not the same genetic trait over the years (Van Pelt et al., 2016b). Correlations of below 0.80 were therefore expected between first EBV and > sixth EBV, since correlations of around 0.60 for longevity were found for test bulls. The authors above showed that functional survival was genetically a more similar trait over the years, and from RR-full we obtained EBV for functional longevity with correlations between first EBV and > sixth EBV that were around 0.70 for test bulls. The adjustment for within-herd production level helped to improve the ranking of bulls over time, although a considerable re-ranking was still taking place; this was expected from the genetic parameters and the change in reliability.

The re-ranking of proven bulls was observed to be more prominent than the re-ranking of test bulls. To some extent this larger change was expected, since the change in reliability from first to last breeding value is also clearly larger for proven bulls than for test bulls. Other reasons for this may be that the proven bulls are a group selected from all tested bulls, and the correlation is likely influenced by this selection. Bulls that initially ranked high for longevity were more likely to become proven bulls, since longevity was included in the national index. In addition, daughters of proven bulls are given more preference by farmers, since these have been shown to be good bulls. The number of proven bulls with eleven EBVs was limited, and the reduction in correlation between first and later EBVs was continuous and consistent over all models. A reason for this continuous reduction may be that second-crop daughters of proven bulls were culled differently from the first-crop daughters. In the progeny testing system, in the pre-genomics era, test bulls were mated randomly. The first-crop daughters were not treated preferentially by farmers; only if a test bull was promoted to a proven bull and used heavily would farmers have been aware that they had a first-crop daughter of this particular bull and kept her longer. Therefore, the first batch of second-crop daughters may have been treated differently than later batches of second-crop daughters of this bull. The first batch of second-crop daughters was based on a limited group of first-crop daughters (<200), and the expectations for these bulls were high. Farmers were aware that they had the daughters of this bull; if the bull turned out to be good, they used the bull again. The later batches were treated without extra attention, unless the bull was exceptionally good or bad and the farmer was prejudiced in a decision to keep or cull the progeny of this bull. Since it is possible that farmers culled first-crop daughters differently from second-crop daughters, an investigation should be carried out into whether a fixed effect for the type or age of the bull could remove the bias we observed in the breeding values. For conformation, this type of fixed effect is included in genetic evaluation in Germany, Italy and the Netherlands (Interbull GE-forms, 2016). In the case of the Netherlands, an interaction between the type and age of the bull is included in the statistical model, together with an effect for the age of the dam. Including these two effects in the model helped to reduce the bias in the breeding values for the conformation of bulls from different sequential evaluations.

The suggestions put forward above, including extra fixed effects and testing of the differences in EBV between a sire-MGS model and an animal model, should be tested and analyzed to determine the impact on the bias of EBV. When the bias in the RRM is reduced or removed completely, we expect that the standard deviation of differences in EBV will decrease. This could potentially lead to an even higher

correlation between first and later EBVs than that found in this study. However, before the introduction of the new genetic evaluation for longevity based on random regression, Interbull validation tests must be successfully passed, as described by Boichard et al. (1995).

5.5 Conclusions

A new genetic evaluation model for longevity is developed here based on a random regression animal model. Based on the correlation between first EBV and later EBVs, the ranking of bulls was shown to be more stable for RRM than for PHM. Bias in EBV was observed, due to incomplete daughter information. Adjusting for within-herd production level reduced this bias greatly in RR-full, to a level similar to that for PHM. Before implementing this new model for genetic evaluation, the cause of this bias needs to be further investigated. Possible options for investigation are the fitting of an extra effect for type of bull or the fitting of a sire-MGS model.

5.6 Acknowledgments

The authors acknowledge funding from the Dutch Dairy Board (PZ, Zoetermeer, the Netherlands), Genetic Evaluation for Sires (GES), cattle improvement cooperative CRV (Arnhem, the Netherlands) and the Breed4Food project (program “Kennisbasis Dier”, code: BO-12-22.04-011-001-ASG-LR).

5.7 References

- Boettcher, P. J., L. K. Jairath, and J. C. W. Dekkers. 1999. Comparison of methods for genetic evaluation of sires for survival of their daughters in the first three lactations. *J. Dairy Sci.* 82:1034-1044.
- Boichard, D., B. Bonaiti, A. Barbat, and S. Mattalia. 1995. Three methods to validate the estimation of genetic trend for dairy cattle. *J. Dairy Sci.* 78:431-437.
- Brotherstone, S., R. F. Veerkamp, and W. G. Hill. 1997. Genetic parameters for a simple predictor of the lifespan of Holstein-Friesian dairy cattle and its relationship to production. *Anim. Sci.* 65:31-37.
- Ducrocq, V. and J. Sölkner. 1994. "The Survival Kit", a FORTRAN package for the analysis of survival data. Pages 51-52 in Proc. Proc. 5th World Congr. Genet. Appl. Livest. Prod., Guelph, Ontario, Canada.
- Ducrocq, V. and J. Sölkner. 1998. The Survival Kit—V3.0: A Package for Large Analyses of Survival Data. Pages 447–450 in Proc. Proc. 6th World Congr. Genet. Appl. Livest. Prod., Armidale, Australia.
- Ducrocq, V., J. Sölkner, and G. Mészáros. 2010. Survival Kit v6 - a software package for survival analysis. in Proc. Proc. 9th World Congr. Genet. Appl. Livest. Prod., Leipzig, Germany.
- Forabosco, F., J. H. Jakobsen, and W. F. Fikse. 2009. International genetic evaluation for direct longevity in dairy bulls. *J. Dairy Sci.* 92:2338-2347.
- Gengler, N., S. Vanderick, P. Mayeres, A. Gillon, and C. Croquet. 2005. Genetic evaluation of cow survival using a lactation random regression model. *Interbull Bulletin* 33:176.
- Gilmour, A. R., B. Gogel, B. Cullis, R. Thompson, D. Butler, M. Cherry, D. Collins, G. Dutkowski, S. Harding, and K. Haskard. 2009. ASReml user guide release 3.0. VSN International Ltd, Hemel Hempstead, UK.
- Holtsmark, M., B. Heringstad, and J. Ødegård. 2009. Predictive abilities of different statistical models for analysis of survival data in dairy cattle. *J. Dairy Sci.* 92:5730-5738.
- Jamrozik, J., J. Fatehi, and L. R. Schaeffer. 2008. Comparison of models for genetic evaluation of survival traits in dairy cattle: a simulation study. *J. Anim. Breed. Genet.* 125:75-83.
- Jenko, J., G. Gorjanc, M. Kovac, and V. Ducrocq. 2013. Comparison between sire-maternal grandsire and animal models for genetic evaluation of longevity in a dairy cattle population with small herds. *J. Dairy Sci.* 96:8002-8013.

- Liu, Z., F. Reinhardt, A. Bünger, and R. Reents. 2004. Derivation and calculation of approximate reliabilities and daughter yield-deviations of a random regression test-day model for genetic evaluation of dairy cattle. *J. Dairy Sci.* 87:1896-1907.
- Meyer, K. 1989. Approximate accuracy of genetic evaluation under an animal model. *Livest. Prod. Sci.* 21:87-100.
- Sasaki, O., M. Aihara, A. Nishiura, H. Takeda, and M. Satoh. 2015. Genetic analysis of the cumulative pseudo-survival rate during lactation of Holstein cattle in Japan by using random regression models. *J. Dairy Sci.* 98:5781-5795.
- Sewalem, A., F. Miglior, G. J. Kistemaker, P. Sullivan, G. Huapaya, and B. J. Van Doormaal. 2007. Short communication: Modification of genetic evaluation of herd life from a three-trait to a five-trait model in Canadian dairy cattle. *J. Dairy Sci.* 90:2025-2028.
- Van der Linde, C., G. de Jong, and A. Harbers. 2004. Using a piecewise Weibull mixed model in the genetic evaluation for longevity. *Interbull Bulletin* 32:157.
- Van der Linde, C., A. Harbers, and G. de Jong. 2007. From functional to productive longevity in the Netherlands. *Interbull Bulletin* 37:203.
- Van Pelt, M. L., G. de Jong, and R. F. Veerkamp. 2016a. Changes in the genetic level and the effects of age at first calving and milk production on survival during the first lactation over the last 25 years. *Animal* 10:2043-2050.
- Van Pelt, M. L., V. Ducrocq, G. de Jong, M. P. L. Calus, and R. F. Veerkamp. 2016b. Genetic changes of survival traits over the past 25 years in Dutch dairy cattle. *J. Dairy Sci.* 99:9810-9819.
- Van Pelt, M. L., T. H. E. Meuwissen, G. de Jong, and R. F. Veerkamp. 2015. Genetic analysis of longevity in Dutch dairy cattle using random regression. *J. Dairy Sci.* 98:4117-4130.
- Veerkamp, R. F., S. Brotherstone, B. Engel, and T. H. E. Meuwissen. 2001. Analysis of censored survival data using random regression models. *Anim. Sci.* 72:1-10.
- Visscher, P. M. and M. E. Goddard. 1995. Genetic parameters for milk yield, survival, workability, and type traits for Australian dairy cattle. *J. Dairy Sci.* 78:205-220.
- Vollema, A. R., S. dan der Beek, A. G. F. Harbers, and G. de Jong. 2000. Genetic evaluation for longevity of Dutch dairy bulls. *J. Dairy Sci.* 83:2629-2639.
- Yazdi, M. H., P. M. Visscher, V. Ducrocq, and R. Thompson. 2002. Heritability, reliability of genetic evaluations and response to selection in proportional hazard models. *J. Dairy Sci.* 85:1563-1577.

6

General discussion

6.1 Introduction

Longevity is of economic importance for dairy farmers. Cows that live longer help to increase the profitability of the dairy farm, since (1) the annual costs of replacing cows are reduced; (2) the average herd yield is increased through an increase in the proportion of cows in the higher-producing age groups; (3) a given acreage is better used by reducing the number of replacements to be reared, thereby allowing an increase in the size of the milking herd; and (4) voluntary culling for low production can be increased (Rendel and Robertson, 1950). Genetic improvement is an important option for increasing the potential longevity of dairy cows. However, selection for better longevity is complex, because the true longevity of a cow is available only at the end of her life, whereas selection and breeding decisions are made earlier in life. Most daughters are still alive when a bull is of interest to the market, and right-censored data need to be used in genetic evaluations. Breeding values for lifespan are therefore an extrapolation from survival data during early life. The use of survival records adds to the complexity of the statistical models since survival is a binary trait. This led to the development of many models in the mid-1990s, and several different models are currently used for genetic evaluation; for an overview see Forabosco et al. (2009).

The aim of this thesis was to investigate the genetics of longevity and to develop a genetic evaluation system for the prediction of breeding values for longevity, with the objective that breeding values should reflect the true breeding value more quickly during early life and therefore become more stable. The components of an improved model were investigated in Chapters 2 to 4 of this thesis. These components led to the definition of a new model for genetic evaluation of longevity in the Netherlands and Flanders; software for this model was developed and the new model was implemented and tested in Chapter 5. In this section, the proposed model is first discussed within the context of these specifications and other models used worldwide. Following this, longevity and its association with other traits are examined, and finally, practical considerations for the implementation of a routine genetic evaluation of longevity based on the final model described in Chapter 5 are put forward.

6.2 Modeling longevity in genetic evaluation

The genetic evaluation model should ideally be able to take into account all the information available on dairy cows (right-censored records from the early life of young heifers); it should allow for fixed effects to vary over time, to account for different reasons for culling during the lifespan of animals (fitting more than one

6 General discussion

genetic effect over life), to handle the non-normal distribution of survival, and ideally to include information on genomics and other predictor traits collected during early life. Currently, none of the models used in four major dairy breeding countries (France, United States, Canada and the Netherlands) fits all of these specifications (Table 6.1). How the research in this thesis, and the proposed model (NLD⁵ in Table 6.1) takes into account these specifications will be discussed.

Table 6.1. Specifications of genetic evaluation for longevity in four different countries (France, United States, Canada, the Netherlands; current and proposed model)

Requirement	FRA ¹	USA ²	CAN ³	NLD ⁴	NLD ⁵
Use of all data available: trait definition	✓	✓		(✓)	✓
Use of time-dependent effects	✓		✓/0	✓	✓
Fitting multiple genetic effects across the cow's life			✓		✓
Fitting an animal model to large data sets		✓	✓		✓
Accounting for non-normal distribution of data	✓			✓	
Include genomic information					(✓)
Use of predictor traits					(✓)

¹S-MGS PHM; ²ST AM; ³MT AM; ⁴current genetic evaluation model S-MGS PHM; ⁵proposed genetic evaluation model RR (MT) AM

6.2.1 Use of all data available: trait definition

To obtain the breeding values of a bull as quickly as possible, and to avoid bias, the data of living and culled cows should be used simultaneously in the breeding value estimation. When only data of culled cows is used, for example by analyzing the actual longevity of cows, the information on the cows which are still alive is ignored, and bulls will have inaccurate and biased breeding values (Veerkamp et al., 2001). The trait definition and statistical model determine if all the information available can be used in the genetic evaluation. Several different trait definitions are used by countries worldwide when estimating breeding values for longevity; for example the length of productive life, calculated as the number of days between first calving and last known test date; survival over the first three lactations, survival over successive lactations; or survival rate at second calving (Forabosco et al., 2009). In this thesis, months one to 72 after first calving were investigated in Chapter 2, and survival per month was proposed in Chapter 5 for the final random regression model (**RRM**) for estimating breeding values. In this proposed model, the observation of survival was modeled as a record for each time interval in which

a cow was 'alive', followed by one record in which it was 'dead', at the time interval the cow was culled; this was followed by records marked 'missing' up to the end of the observed period, following Veerkamp et al. (2001). Thus, in principle, all of these records can be used in the genetic evaluation. This approach stands in contrast to methods which consider survival over the first three lactations only, and where survival over later lactations is ignored. The proportional hazards model (PHM) currently implemented in the Netherlands is also able to use all records available; however, a voluntary waiting period of 270 days is implemented to avoid the fluctuation in breeding values arising when the first records of the second-crop daughters of a bull become available. In Chapter 2, evidence was presented that those early records should indeed be given a lower weight for the prediction of lifespan, because the genetic correlation between survival until 6 mo and lifespan is 0.80. This lower correlation between early survival and lifespan is taken into account; thus, although the proposed model has not yet been formally tested, it can be expected that where the records are weighted by the genetic correlation, a voluntary waiting period may be less important, and a breeding value may therefore be available earlier in the life of a bull. In Chapter 5, no voluntary waiting period was applied for the estimation of breeding values, and the random regression curves of the animals were based solely on the first month. With only one data point and the parent average, the estimation of the regression curve may be difficult; thus, bulls need to have at least 15 daughters which are at least 120 days into first lactation in order to have their breeding value for milk production published in the Netherlands and Flanders. However, all information is used in the genetic evaluation for milk production, whereas in the currently used model for longevity based on PHM, a waiting period of nine months is applied before including the information on early life. This means that a cow is required to have been alive for at least nine months after first calving at the most recent date in the dataset for genetic evaluation. In the model proposed here, a waiting period is not applied; however, criteria are recommended to be set at bull level, for example a requirement for at least 15 progeny with at least four months of survival data, in order to obtain a more reliable estimate of the regression curve as compared with publishing the breeding value with only one daughter in the first month. When discussing trait definition, it should be noted that there are two different ways in the literature in which survival is modeled as time-series data. In the proposed approach, observations of survival were recorded as 'alive', 'dead' and 'missing' as described above. In this way, the hazard of a cow to be culled within a time interval is modeled, and this definition is close to the hazard function in the PHM. Schaeffer (2004), however, proposed the use of the same number of records

6 General discussion

up to a certain endpoint for each uncensored cow; after culling, the records for a cow are coded as ‘dead’ up to a certain endpoint, rather than ‘missing’. This definition is close to the survival function in the PHM. Both of these definitions are illustrated in Figure 6.1. Jamrozik et al. (2008) analyzed survival up to 100 mo in simulated data; survival up to eight parity in beef cattle was also analyzed by Jamrozik et al. (2013) using the definition of Schaeffer (2004). With Schaeffer’s trait definition, (co)variances for survival can be summed using standard matrix algebra, similar to deriving lactation breeding values for milk yield traits (Jamrozik et al., 1997). Although both trait definitions are not compared here, the definition used in Chapter 2 and 5 for modeling the hazard rate appears to offer a more appropriate way to model the contemporary groups. When survival is defined up to a certain period, contemporary groups are formed by grouping animals calving for the first time in the same herd-year-season (Schaeffer, 2004). However, when survival is defined in the way proposed in this work, contemporary groups are formed from the animals present within the same herd-year-season. In the first type of contemporary group, culled animals are still part of the contemporary group, although farmers cannot in fact consider those animals for culling. This last type of contemporary group is closer to actual practice, as we used it, in which the farmer decides which animals to cull from the current herd composition.

a)	1	2	3	4	5	6	7	8	9	10	11	12
	1	1	1	1	1	1	1	1	0	-	-	-

b)	1	2	3	4	5	6	7	8	9	10	11	12
	1	1	1	1	1	1	1	1	0	0	0	0

Figure 6.1. Trait definition for survival according to a) Veerkamp et al. (2001), as used in this thesis, and b) Schaeffer (2004), with an example in which 12 time intervals are defined and the animal is culled in the ninth time interval.

6.2.2 Use of time-dependent effects

Some of the fixed effects in the genetic evaluation can be considered to have the same effect on survival over the whole life of the cow, while others may vary over time. For example, the definition of herd x year x season at birth assumes that there is a fixed effect influencing the risk of culling over the entire life of an animal. In contrast, the contemporary group of animals in the herd during the evaluated survival period, i.e., herd x year x season of the observation, could be used. This latter definition of the contemporary group assumes that effects vary over time, and these are therefore treated as time-dependent fixed effects. Two types of

time-dependency are distinguished: (1) during the life of a cow; and (2) over calendar time. One time-dependent effect during the life of a cow is parity x lactation stage. Year x season and age of first calving x 5-year interval of first calving are time-dependent effects that vary across calendar time. PHM and RRM are able to handle both of these types of time-dependent effects, since these models can analyze time-series data. One primary advantage of time-series data is that effects which change over life can be modeled more precisely when time intervals are short, as compared with single-trait models which use an observation measured over a longer time interval. This is the case for parity x lactation stage, where it is possible to fit a separate effect for the dry period. We know that if the cow has a dry period, she will not be culled. Therefore the hazard of culling is zero during the dry period, as opposed to the other classes for stage of lactation. The fitting of this dry period is not possible when time intervals are defined over time periods longer than 1 mo. Therefore, time intervals of 1 mo are chosen in this work, and this allows the modeling of time-dependent fixed effects that change every month during life.

6.2.3 Fitting multiple genetic effects over the cow's life

Based on earlier literature and the observation that breeding values fluctuate more than expected on the basis of changes in the reliability of the breeding value, an investigation was carried out in Chapter 2 into whether survival was a genetically different trait between different time intervals over life, in order to determine the complexity needed for the (co)variance matrices for genetic evaluation. By analyzing survival per time interval, the (co)variances between independent time intervals could be estimated. RRM thus allows the estimation of genetic correlations across life, whereas with PHM or a single-trait model this is not possible. Even when genetic correlations between different periods of life are close to unity, RRM can fit heterogeneous genetic variances. The parameterization proposed by Schaeffer (2004) yields part-whole correlations. The parameterization proposed here allows the investigation of the genetic correlation structure between different moments in the life of a cow, for example between 3 mo and 36 mo after first calving. Moreover, with the definition used in this work it was possible to calculate part-whole correlations for survival up to different endpoints. Genetic correlations between time intervals ranged from 0.25 to 1.00, as shown in Chapter 2; in general, genetic correlations decreased when time intervals were further apart. In previous studies, genetic correlations of survival between different (part) lactations were in the range 0.33 to 0.96, where the lowest correlations were estimated between more distant lactations (Visscher and Goddard, 1995,

Brotherstone et al., 1997, Boettcher et al., 1999, Veerkamp et al., 2001, Sewalem et al., 2007, Holtsmark et al., 2009). The range of genetic correlations in the current study was slightly wider, due to the higher number of time points analyzed, with shorter time intervals, and because the analyzed period was longer. In Chapter 5, genetic parameters were re-estimated for both longevity and functional longevity using a fifth-order Legendre polynomial to fit survival per month. For longevity, genetic correlations between survival at various months after first calving confirmed the results of Chapter 2, albeit they were somewhat higher (0.51–1.00); however, when survival was adjusted for within-herd production level, genetic correlations showed a larger range (0.06–1.00). The genetic correlation structure indicates that survival early in productive life should be treated differently in genetic evaluations than survival later in life by fitting multiple genetic effects with nonunity correlations.

6.2.4 Fitting an animal model to large data sets

In 1949, the method of best linear unbiased prediction (BLUP) was first described by Henderson, although due to limitations in computing power and hardware the first application, known as the sire model, was not used until 1970 (Mrode, 2014). BLUP has evolved towards more complex models used for genetic evaluations. The first application of the animal model was described in 1989, and in 2000 the random regression model was implemented, facilitated by the development of computer hardware and increases in sizes of memory (Grosu et al., 2013). Genetic evaluation for longevity was introduced in 1999 in the Netherlands, and at that time the optimum choice was the implementation of the Survival Kit software (Ducrocq and Sölkner, 1994). This software package allowed the analysis of large datasets, although only using a sire or a sire-MGS model. Although it is possible to use an animal model with the Survival Kit software, memory requirements are very large, even for a small population (Jenko et al., 2013). The genetic evaluation model developed in Chapter 5 was a random regression animal model, and this showed acceptable performance in terms of computer memory usage and CPU time. Hence, it is computational possible to use an animal model for the genetic evaluation of longevity.

6.2.5 Accounting for non-normal distribution of data

Survival is coded as 1 (survived) or 0 (culled), making this a categorical trait. Linear and non-linear models have been applied for the genetic analysis of categorical traits with the assumption of an underlying normally distributed liability (Mrode, 2014). Usually, the non-linear (threshold) models are more complex and have

higher computing requirements. The advantage of the linear model is the ease of implementation, as software used for the genetic evaluation of quantitative traits can be utilized without any modifications. Meijering and Gianola (1985) showed that a linear model is able to perform just as well as a threshold model in terms of correctly ranking dairy sires. For survival, no or little re-ranking of bulls was expected in other studies with correlations of 0.98 and higher between estimated breeding values (**EBV**) from a linear model and EBV from a threshold model (Boettcher et al., 1999, Holtsmark et al., 2009), and also for clinical mastitis (Carlén et al., 2006). In the proposed model, little re-ranking was found for bulls with correlations of 0.97 and higher between EBV from a linear model and EBV from a threshold model estimated with RRM and different interval lengths for survival, as described in Chapter 2. Although the proposed model does not fully take account of the binary distribution of survival data, little loss in terms of accuracy of the breeding value is expected.

6.2.6 Include genomic information

Genomic enhanced breeding values (**GEBV**) for longevity in the Netherlands and Flanders are currently based on a post-processing step (VanRaden et al., 2009), where conventional EBVs are integrated with direct genomic values (**DGV**) from de-regressed proofs and SNP information (Stoop et al., 2011). A drawback of this method is that the genomic information of an individual does not influence the breeding value of non-genotyped relatives. For all other traits, genomic and conventional data are combined in a single genetic evaluation, using a method based on Mäntysaari and Strandén (2010). In this method, a two-step approach is used in which: 1) a DGV is calculated in a genomic evaluation (De Roos et al., 2009); and 2) the resulting DGV of an animal is transformed into a pseudo-observation (PSR) on absolute scale for a pseudo-trait with a heritability of 0.999, which is subsequently included in the national evaluation as a correlated trait (Stoop et al., 2013). For all traits except longevity, PSRs are used with a multiple-trait animal model using the PCG-solving algorithm, which also can handle random regression models, resulting in GEBV for males and females. The current genetic evaluation for longevity is a single trait sire-MGS PHM; it cannot easily be expanded to a multiple-trait model or use genomic information directly. Although a two-step approach is currently used to obtain GEBV, single-step approaches have been developed which directly combine genomic and conventional information into a GEBV (Misztal et al., 2009, Fernando et al., 2014). The expectation is that future genetic/genomic evaluations in the Netherlands and Flanders will use a single-step approach. To enable the single step analysis, a change from the PHM model to the proposed

RRM model is required, with the added advantage that the genetic evaluations for all traits use the same solving algorithm and software.

6.2.7 Use of predictor traits

Predictor traits, measured early in life, can improve the reliability of the breeding value due to the genetic correlation with longevity. Most countries participating in the Interbull evaluation for longevity combine breeding values of predictors and direct longevity in a selection index to have a more reliable breeding value when only few cows are culled. Conformation traits, production traits, somatic cell count, calving traits, fertility traits and workability are all used as predictors (Forabosco et al., 2009). Instead of combining the predictor traits and longevity in a two-step approach with a selection index, the predictor traits can be used directly in the genetic evaluation as correlated traits in a multiple trait model. The software developed for the proposed model for longevity is already able to perform multi-trait evaluations, by fitting a linear model for the predictor traits and an RRM for survival. This model allows for genetic correlations to differ between the predictor traits and survival at different time intervals, as shown by Heise et al. (2016), where, for example, the correlation between EBV for longevity and a fertility EBV ranged between 0.04 and 0.46 between early and late lactation. Further investigation is required to identify the genetic correlations between predictor traits and survival per mo, and which traits should be added as predictors in the genetic evaluation for longevity.

6.3 Relationship of longevity with other traits

The lifespan of a cow is affected by many factors, and there are numerous reasons for culling (Beaudeau et al., 2000, Fetrow et al., 2006), for example fertility problems, mastitis, lameness, health disorders, low production or other reasons. In the mid-1990s there was a great deal of interest in developing breeding values for health and fertility traits; however, phenotypic data were not readily available for genetic evaluation of many of these traits. In the Netherlands, no extra effort was needed to collect data, as information about longevity was derived from existing information sources on milk production and information on the movement of animals between herds. The breeding value for longevity was perceived to be an all-inclusive trait, i.e. a reflection of an overall breeding goal focused on improving production, health and fertility traits to obtain high-producing, trouble-free cows. Animals with poor production, health and fertility have a higher culling risk compared to animals that perform well in these traits, since farmers cull animals for these reasons and consequently the good animals automatically grow older.

Many more breeding values have been developed since the 1990s, and it is important to discuss the breeding value for longevity within the context of the genetic correlation with these other traits.

6.3.1 Genetic correlations

From the analysis in Chapter 5, breeding values for longevity and functional longevity were available; it is of interest to understand their relationships with other traits in the breeding goal. Traits were chosen from the total merit index (NVI) of the Netherlands and Flanders. Genetic correlations between (functional) longevity and the traits related to milk production, conformation, and health and fertility were estimated using bivariate analyses based on the method described by Schaeffer (1994). The genetic correlation with claw health was also estimated, since this is a health trait expected to be related to one of the main reasons for culling. Estimated genetic correlations between all traits and heritabilities used are shown in Table 6.2. Differences exist between longevity and functional longevity in their genetic correlations with the other traits. The correlation with production traits changes from positive for longevity (0.34 to 0.44) to a negative or zero correlation for functional survival (-0.27 to 0.04). For fertility traits, the correlation changes from no correlation with longevity (0.02 and 0.10) to a positive correlation with functional survival (0.25 and 0.30). It was expected that functional survival would have a correlation close to zero with production traits, because functional survival was adjusted for within-herd production level. Consequently, fertility traits show positive correlations with functional survival. The other traits show a difference in genetic correlation of 0.10 or smaller.

When ranking the traits based on the genetic correlation with (functional) longevity, the three trait groups with the strongest relationship with longevity were udder health, claw health and production; with functional longevity these were udder health, claw health and fertility. Udder health and claw health are thus the best predictors for both longevity and functional longevity. This is as expected, since these traits are related to culling for mastitis and lameness, two primary reasons for culling. The third best predictor is production for longevity and fertility for functional longevity. The other traits in the NVI also showed a positive genetic correlation with (functional) longevity, and aid in improving longevity when bulls are selected based on the NVI.

6 General discussion

Table 6.2. Genetic correlations between traits and heritabilities (in bold) for the traits used in the selection index to predict (functional) longevity

	lon	fun	fat	pro	lac	udd	f&l	udh	ifl	cin	dce	mce	dlv	mlv	clw
lon	0.12														
fun	0.94	0.14													
fat	0.34	-0.27	0.58												
pro	0.44	-0.09	0.61	0.50											
lac	0.35	0.04	0.34	0.83	0.55										
udd	0.23	0.33	-0.11	-0.14	-0.11	0.34									
f&l	0.22	0.17	0.07	0.06	0.04	0.18	0.17								
udh	0.52	0.59	-0.05	-0.13	-0.08	0.29	0.06	0.089							
ifl	0.10	0.30	-0.26	-0.34	-0.40	0.07	0.01	0.26	0.08						
cin	0.02	0.25	-0.31	-0.39	-0.43	0.09	0.11	0.27	0.97	0.15					
dce	0.27	0.26	0.05	0.02	-0.05	0.02	0.04	0.16	0.28	0.23	0.068				
mce	0.24	0.30	-0.03	0.03	-0.01	0.10	0.10	0.19	0.44	0.39	0.79	0.048			
dlv	0.19	0.18	0.04	0.05	0.01	-0.02	0.05	0.08	0.16	0.13	0.42	0.34	0.038		
mlv	0.10	0.11	0.01	0.00	-0.03	-0.07	0.00	0.06	0.32	0.23	0.14	0.53	0.12	0.085	
clw	0.52	0.42	0.17	0.11	-0.08	0.12	0.69	0.14	0.22	0.30	0.35	0.21	0.17	0.08	0.17

lon = longevity, fun = functional longevity, fat = kg fat, pro = kg protein, lac = kg lactose, udd = udder, f&l = feet and legs, udh = udder health, ifl = interval between first and last insemination, cin = calving interval, dce = direct calving ease, mce = maternal calving ease, dlv = direct livability, mlv = maternal livability, clw = claw health

6.3.2 Accuracy of selection

If the trait of longevity is the breeding goal, it is of interest to determine the accuracy of a selection for longevity when breeding values for the predictor traits are available. Selection index theory can be used for this purpose to calculate the accuracy of longevity based on correlated traits. The two traits for prediction as breeding goals are longevity and functional longevity, as presented in Chapter 5. For simplicity, the traits in the NVI plus claw health were chosen. Claw health was added to the selection index calculations, since this was the only other trait with a moderate genetic correlation (>0.40) with both longevity and functional longevity. Accuracies were calculated for four types of bulls and 14 scenarios, for both longevity and functional longevity. The four types of bulls were chosen in order to simulate the amount of information available at a certain age of the bull, and these were: 1) a young bull with genomic information and parent average; 2) a test bull with 125 daughters with milk production records from the test period; 3) a heavily used genomic bull, with 2,000 daughters with milk production records; and 4) a highly reliable proven bull with 50,000 daughters with milk production information. The genomic information for a young bull was simulated as 15 progeny for all traits in the selection index. The number of progeny for a test bull was not the same for all traits: these were 125 for milk production, udder health, and interval between first and last insemination; 100 for calving interval; 75 for conformation and maternal calving ease; 200 for direct calving ease; 400 for direct livability; 140 for maternal livability; and 50 for claw health. The same ratio was used for the number of progeny of production traits and the other traits for the heavily used genomic bull and the proven bull. The 14 scenarios were chosen in order to gain an impression of how well other traits can predict longevity, and to simulate the evolution of the NVI. The scenarios were: 1) production only; 2) conformation only; 3) health only; 4) claw health only; 5) production and conformation; 6) health and claw health; 7) production, conformation and health; 8) production, conformation, health and claw health; 9) (functional) longevity only; 10) all traits in the NVI; 11) all traits; 12) conformation and health; 13) conformation, health and claw health; and 14) conformation, health, claw health, and longevity.

6 General discussion

Table 6.3. Accuracy of prediction of longevity using different traits or combinations of traits for young genomic bulls (YG), test bulls (Test), highly used genomic bulls with progeny information (G+P), and highly reliable proven bulls (Proven)

Index	prod	conf	hlth+fer	claw	lon	YG	Test	G+P	Proven
1	X					0.41	0.44	0.44	0.45
2		X				0.23	0.27	0.29	0.29
3			X			0.36	0.48	0.58	0.61
4				X		0.37	0.44	0.51	0.51
5	X	X				0.48	0.53	0.55	0.55
6			X	X		0.54	0.67	0.82	0.90
7	X	X	X			0.62	0.74	0.81	0.83
8	X	X	X	X		0.69	0.73	0.92	0.99
9					X	0.68	0.90	0.99	1.00
10	X	X	X		X	0.77	0.92	0.99	1.00
11	X	X	X	X	X	0.79	0.93	0.99	1.00
12		X	X			0.44	0.57	0.67	0.70
13		X	X	X		0.56	0.68	0.85	0.94
14		X	X	X	X	0.74	0.92	0.99	1.00

prod = production traits, conf = conformation traits, hlth+fer = health and fertility traits, claw = claw health, lon = longevity

Table 6.4. Accuracy of prediction of functional longevity using different traits or combinations of traits for young genomic bulls (YG), test bulls (Test), highly used genomic bulls with progeny information (G+P), and highly reliable proven bulls (Proven)

Index	prod	conf	hlth+fer	claw	fun	YG	Test	G+P	Proven
1	X					0.26	0.30	0.30	0.31
2		X				0.29	0.33	0.35	0.35
3			X			0.42	0.55	0.62	0.63
4				X		0.30	0.36	0.41	0.42
5	X	X				0.38	0.44	0.46	0.46
6			X	X		0.51	0.64	0.74	0.78
7	X	X	X			0.55	0.68	0.75	0.77
8	X	X	X	X		0.63	0.76	0.90	0.99
9					X	0.70	0.91	0.99	1.00
10	X	X	X		X	0.76	0.93	0.99	1.00
11	X	X	X	X	X	0.77	0.93	0.99	1.00
12		X	X			0.49	0.61	0.68	0.70
13		X	X	X		0.55	0.66	0.77	0.82
14		X	X	X	X	0.76	0.93	0.99	1.00

prod = production traits, conf = conformation traits, hlth+fer = health and fertility traits, claw = claw health, fun = functional longevity

Table 6.3 shows accuracies of predictions of longevity, and Table 6.4 shows those for functional longevity. When the separate trait groups (indices 1 to 4) are used to predict (functional) longevity the accuracies are limited: 0.23 to 0.63 for longevity and 0.26 to 0.65 for functional longevity. These also depend on the amount of information available for the bull. However, health and claw health have higher accuracies, especially when progeny information is available. Combining production and conformation traits (index 5), as was usual in several countries until the late 1990s (Miglior et al., 2005), results in a maximum accuracy of 0.55 for proven bulls. During these years, most countries carried out genetic evaluations for these traits; it was therefore worthwhile to develop a breeding value for longevity and include this in the selection index. Later, several other health traits were developed. A combination of all health traits (index 6) resulted in better prediction of (functional) longevity compared to index 5 for all types of bulls, from 0.74 to 0.90 for bulls with at least 2,000 daughters for milk production. Today, the NVI includes production, conformation, health and fertility traits (index 7), and this index has a slightly lower accuracy than index 6 for bulls with at least 2,000 daughters for milk production; however, accuracy was higher for the young genomic bulls and the test bulls since production traits have higher reliabilities. Adding claw health to index 6 (index 7) resulted in higher accuracies for all types of bulls. For proven bulls with a large progeny group it is possible to predict (functional) longevity almost completely, and there is no added value from including (functional) longevity in the selection index. It also appears that claw health is able to predict the remaining part of longevity, from a comparison of indices 6 and 7. Udder health, fertility, and claw health are the most related to culling reasons, and it is therefore logical that accuracy improves from ~0.80 to 0.99 when claw health is added (index 8). Not all bulls will obtain highly accurate breeding values for all traits, especially for health traits with low heritability, and accuracy is limited for test bulls, with 0.73 for longevity and 0.76 for functional longevity. When only direct information for (functional) longevity is used (index 9), accuracy increases to 0.90 for test bulls. When the breeding goal includes most or all trait groups and longevity, accuracy is 0.76 or higher (indices 10 and 11), and this is slightly higher than considering only longevity. Indices 12 to 14 are selection indices which do not include production in the index, as a reflection of the current situation where production is not used as predictor for longevity. Compared to the scenarios with production in the index (index 7 vs. 12, 8 vs. 13, and 11 vs. 14), accuracies are lower, especially for longevity for young bulls and test bulls.

From the selection index calculations it is shown that it is possible to predict both longevity and functional longevity with high accuracy when breeding values are

reliable and all health traits are included in the selection index. However, it is well-known that accuracies tend to be overestimated when more traits are added into the selection index (Hayes and Hill, 1981), and the observation here that the variance of (functional) longevity can be completely explained by other traits indeed seems to indicate that this accuracy is overestimated. Still, several conclusions can be drawn from the selection index calculations. For all scenarios in which claw health was included, the accuracy of prediction of (functional) longevity increased. It is therefore recommended that claw health should be added to the NVI in the Netherlands and Flanders, particularly because this is an important reason for culling. For young bulls and test bulls, which have less reliable breeding values, the increase in accuracy still justifies the inclusion of (functional) longevity in the selection index and therefore to have a genetic evaluation for longevity or functional longevity.

6.3.3 Longevity, functional longevity or residual longevity

Breeding values for longevity are estimated mainly in terms of longevity or functional longevity in national genetic evaluations (Forabosco et al., 2009). If longevity is not corrected for milk production, the ability of delaying any culling is modeled; with functional longevity, the ability of delaying involuntary culling (assuming voluntary culling is for yield) is modeled (Ducrocq et al., 1988). Residual longevity is used to all variation left when longevity is corrected for all traits taken into account by a farmer when culling cows by adjusting for other traits that influence longevity (e.g. health and fertility traits). Longevity, adjusted or not, is now included in the selection index in most countries worldwide together with production and health traits (Miglior et al., 2005, Egger-Danner et al., 2015). A selection index in which one trait is adjusted for one or more other traits in the index is equivalent to an index that includes the unadjusted traits (Kennedy et al., 1993), as long as proper economic weights are used. Therefore, it is not important if longevity should be adjusted or not, as other traits related to longevity are included in the breeding goal.

Still it is important to consider the differences between adjusted and unadjusted longevity. This is also shown in Chapter 3, where genetic trends for longevity and functional longevity were completely different till 1999, and became comparable since 1999, i.e. the moment longevity was included in the selection index. Also, the interpretation of functional longevity and even more for residual longevity would be difficult for farmers. Longevity is the reflection of what farmers do and what they can see, as all culling is included; in adjusted longevity (functional or residual) one or more culling reasons are corrected for, making it less clear for farmers what

part of longevity will be selected for. In Chapter 4 it was also shown that for the genetic evaluation it matters which trait is considered; the trait definition of longevity changed over time, and genetic correlations became as low as 0.78 for an interval of 15 years. When survival was adjusted for milk yield, the genetic correlation was 0.88. This can be explained by the results in Chapter 3 where it was shown that farmers paid less attention to milk yield when keeping heifers to the second lactation. Thus, despite that it should not matter if longevity, functional longevity or residual longevity is weighted in a selection index, for predicting breeding values, it is advised to analyze functional longevity in order to have more stable breeding values more quickly during early life.

6.3.4 Culling reasons

From a biological point of view it is of interest to know for which reason a cow is culled and how this is related with other reasons of culling and with other traits, for example milk production and fertility, and by analyzing the separate culling reasons more insight can be gained in the underlying traits affecting longevity. Though, the reported culling reason is a subjective score of the farmer, and culling of a cow is often done for multiple reasons (Fetrow et al., 2006). In addition, voluntary culling is assumed a (economic) decision of the farmer; expectation of a new heifer is higher than expectations for the culled cow. Tsuruta et al. (2015) analyzed survival for five different culling reasons (sold or culled for dairy production, poor production, reproductive problems, and died at the farm) and survival over the first three lactations. Heritabilities for these six survival traits ranged between 0.03 and 0.11. To understand the biological mechanisms the genetic correlations between these six survival traits among them and with other traits would have been helpful, but these were not estimated. For animal breeding, however, the aim is to select cows that can live longer by delaying the moment of culling for any reason. Therefore, it may not yield a higher genetic gain to estimate breeding values for specific culling reasons rather than analyzing overall survival as analyzed in this thesis.

In the US the culling reason 'died at the farm' is analyzed separately by estimating a breeding value for cow livability in addition to the current breeding value for productive life (longevity). The use of separate genetic evaluations for cow livability and productive life was justified, because about 20% of all culled cows died at the farm, the heritability of cow livability was 0.013 and the correlation with EBV of productive life was 0.70 (Wright and VanRaden, 2016). Moreover, the economic impact of cows dying on the farm is large. The current lost income from the death of US cows has a value of about 2.2 billion USD (20% x 9.2 million cows x 1,200

USD) or 800 million USD annually assuming an average productive life of 2.75 years (Norman et al., 2016). At cow level, this represents an annual loss of income of about 87 USD. In the Netherlands, however, the economic loss per cow per year is smaller; of all culled cows, about 15% had a culling code of 'dead' rather than 'slaughter'. The average slaughter price in the Netherlands is about 700 euros, and the average productive life is 3.5 years. At cow level this represents a loss of 30 euros per year ($15\% \times 700 \text{ euro} / 3.5 \text{ years}$). Compared to the annual loss per cow of 78 euro for mastitis (Huijps et al., 2008) and 75 USD for claw health (Bruijnis et al., 2010), the economic loss for cows dying on the farm is rather limited. It is therefore less important to separate the breeding value for cow livability next to the breeding value for longevity, as cow livability should increase due to the correlated response with longevity.

6.3.5 Survival before first calving

This thesis has examined only the productive life of the cow, i.e., from first calving until the final test date; however, it is also important to improve survival at the earlier stages of life. Breeding values for calf livability over the first 24 hours and calf survival from days three to 365 are already available in the Netherlands. By selecting for these two traits together with longevity there is potential for improving overall survival over the entire life of a cow, that is from birth until the last test date. An idea would be to combine the survival in early life with survival after the first calving. However, to combine calf livability, calf survival and longevity in one statistical model may be too ambitious; for calf livability, the direct effect of the calf and the maternal effect of the dam are included as genetic effect in the model (Van Pelt and de Jong, 2011). For calf survival, data from female calves as well as male calves are included in the genetic evaluation (Van Pelt et al., 2012). Genetic correlations between longevity and calf livability are low (Table 6.2), and the genetic correlation between longevity and calf survival is also likely to be low or moderate, and the traits are genetically different. It is therefore more effective to combine longevity, livability and calf survival in one selection index, rather than analyzing total life in a single model for genetic evaluation.

6.4 Publication of breeding value for longevity

From a practical point of view, it is important to decide what the preferred publication scale of the breeding value for longevity should be. The models used in this thesis offer many opportunities to define different longevity traits, from survival probabilities at any moment in life to the expected lifespan in days or months. The EBV in Chapter 5 were presented in months, whereas the current

breeding value for longevity in the Netherlands uses a scale in days. This was purposely chosen, because farmers perceive a change in EBV of 60 days as a large change. However, 60 days represents a change of less than 0.25 genetic standard deviation in the current evaluation, and therefore presenting the EBV with such a high precision in days, is not justified. Also, the perception of such a change may be different if the breeding value is presented on a scale of months instead of days. Therefore in Chapter 5 it was decided to present the breeding value in months.

Another discussion point to present the breeding value for longevity is whether longevity is presented per lactation or over the entire lifespan. In Chapters 2 and 5, survival per month until 72 months after first calving was examined by fitting a genetic curve over the life of a cow, resulting in a breeding value for each month. In Chapter 2 an equation was derived for combining the EBV per month into one overall EBV for longevity. Another approach is to model genetic curves within a lactation, as carried out by Sasaki et al. (2015) for cumulative survival rate; this enables the modeling of lactation-specific culling. In both studies, nonunity genetic correlations were found between survival of different periods. The EBV of Sasaki et al. (2015) enables to select for improved survival at lactation level, whereas the total EBV in this thesis enables to select for improved total productive life at 72 months after first calving. The longevity breeding values however, can be presented on both scales (lactation or lifespan), depending on the assumptions made about the length of the lactation period. In terms of modelling, it is arbitrary to model survival at the level of lactation or survival over the entire lifespan, as long as nonunity genetic correlations between survival of different periods are taken into account.

6.5 Conclusions

In this thesis, a new model for the genetic evaluation of longevity was developed. From a genetic point of view, the trait longevity changed over the years, whereas the trait functional longevity, i.e., adjusting for within-herd production level, changed less. In the new model, survival up to 72 months after first calving is modeled with a random regression animal model using a fifth-order Legendre polynomial to allow for nonunity genetic correlations for survival between different months after first calving and heterogeneous genetic variances. The genetic correlation between survival up to six versus 72 months after first calving is ~ 0.80 . The ranking of bulls in the new model was more stable compared to the existing model. Bias was shown to be present in the EBV, mainly due to incomplete daughter information. Adjusting for within-herd production level reduced this bias

6 General discussion

greatly, to a level similar to that of the PHM model. The cause of the bias needs to be further investigated before this new model is implemented. Predictor traits can add accuracy to the EBV for longevity, especially for young genomic bulls. In terms of future work, the next step is to estimate genetic parameters for longevity and predictor traits, and to implement the genetic model combining data on both longevity and predictor traits.

6.6 References

- Beaudeau, F., H. Seegers, V. Ducrocq, C. Fourichon, and N. Bareille. 2000. Effect of health disorders on culling in dairy cows: a review and a critical discussion. *Ann. Zootech.* 49:293-311.
- Boettcher, P. J., L. K. Jairath, and J. C. W. Dekkers. 1999. Comparison of methods for genetic evaluation of sires for survival of their daughters in the first three lactations. *J. Dairy Sci.* 82:1034-1044.
- Brotherstone, S., R. F. Veerkamp, and W. G. Hill. 1997. Genetic parameters for a simple predictor of the lifespan of Holstein-Friesian dairy cattle and its relationship to production. *Anim. Sci.* 65:31-37.
- Bruijnis, M. R. N., H. Hogeveen, and E. N. Stassen. 2010. Assessing economic consequences of foot disorders in dairy cattle using a dynamic stochastic simulation model. *J. Dairy Sci.* 93:2419-2432.
- Carlén, E., U. Emanuelson, and E. Strandberg. 2006. Genetic evaluation of mastitis in dairy cattle using linear models, threshold models, and survival analysis: A simulation study. *J. Dairy Sci.* 89:4049-4057.
- De Roos, A., C. Schrooten, E. Mullaart, S. van der Beek, G. de Jong, and W. Voskamp. 2009. Genomic selection at CRV. *Interbull Bulletin*:47.
- Ducrocq, V., R. L. Quaas, E. J. Pollak, and G. Casella. 1988. Length of productive life of dairy cows. 1. Justification of a Weibull model. *J. Dairy Sci.* 71:3061-3070.
- Ducrocq, V. and J. Sölkner. 1994. "The Survival Kit", a FORTRAN package for the analysis of survival data. Pages 51-52 in *Proc. Proc. 5th World Congr. Genet. Appl. Livest. Prod.*, Guelph, Ontario, Canada.
- Egger-Danner, C., J. B. Cole, J. E. Pryce, N. Gengler, B. Heringstad, A. Bradley, and K. F. Stock. 2015. Invited review: overview of new traits and phenotyping strategies in dairy cattle with a focus on functional traits. *Animal* 9:191-207.
- Fernando, R. L., J. C. M. Dekkers, and D. J. Garrick. 2014. A class of Bayesian methods to combine large numbers of genotyped and non-genotyped animals for whole-genome analyses. *Genet. Sel. Evol.* 46:13.
- Fetrow, J., K. V. Nordlund, and H. D. Norman. 2006. Invited review: Culling: Nomenclature, definitions, and recommendations. *J. Dairy Sci.* 89:1896-1905.
- Forabosco, F., J. H. Jakobsen, and W. F. Fikse. 2009. International genetic evaluation for direct longevity in dairy bulls. *J. Dairy Sci.* 92:2338-2347.
- Grosu, H., J. Carvalheira, C. Draganescu, and S. L. Lungu. 2013. History of genetic evaluation methods in dairy cattle.
- Hayes, J. F. and W. G. Hill. 1981. Modification of estimates of parameters in the construction of genetic selection indexes (bending). *Biometrics* 37:483-493.

- Heise, J., Z. T. Liu, K. F. Stock, S. Rensing, F. Reinhardt, and H. Simianer. 2016. The genetic structure of longevity in dairy cows. *J. Dairy Sci.* 99:1253-1265.
- Holtsmark, M., B. Heringstad, and J. Ødegård. 2009. Predictive abilities of different statistical models for analysis of survival data in dairy cattle. *J. Dairy Sci.* 92:5730-5738.
- Huijps, K., T. Lam, and H. Hogeveen. 2008. Costs of mastitis: facts and perception. *Journal of Dairy Research* 75:113-120.
- Jamrozik, J., J. Fatehi, and L. R. Schaeffer. 2008. Comparison of models for genetic evaluation of survival traits in dairy cattle: a simulation study. *J. Anim. Breed. Genet.* 125:75-83.
- Jamrozik, J., S. McGrath, R. A. Kemp, and S. P. Miller. 2013. Estimates of genetic parameters for stayability to consecutive calvings of Canadian Simmentals by random regression models. *J. Anim. Sci.* 91:3634-3643.
- Jamrozik, J., L. R. Schaeffer, and J. C. M. Dekkers. 1997. Genetic evaluation of dairy cattle using test day yields and random regression model. *J. Dairy Sci.* 80:1217-1226.
- Jenko, J., G. Gorjanc, M. Kovac, and V. Ducrocq. 2013. Comparison between sire-maternal grandsire and animal models for genetic evaluation of longevity in a dairy cattle population with small herds. *J. Dairy Sci.* 96:8002-8013.
- Kennedy, B. W., J. H. J. van der Werf, and T. H. E. Meuwissen. 1993. Genetic and statistical properties of residual feed-intake. *J. Anim. Sci.* 71:3239-3250.
- Mäntysaari, E. A. and I. Strandén. 2010. Use of bivariate EBV-DGV model to combine genomic and conventional breeding value evaluations. in *Proc. 9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany.*
- Meijering, A. and D. Gianola. 1985. Linear versus nonlinear methods of sire evaluation for categorical traits - a simulation study. *Genet. Sel. Evol.* 17:115-131.
- Miglior, F., B. L. Muir, and B. J. Van Doormaal. 2005. Selection indices in Holstein cattle of various countries. *J. Dairy Sci.* 88:1255-1263.
- Misztal, I., A. Legarra, and I. Aguilar. 2009. Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. *J. Dairy Sci.* 92:4648-4655.
- Mrode, R. A. 2014. *Linear models for the prediction of animal breeding values.* 3 ed. Cabi.
- Norman, H. D., J. R. Wright, and P. M. VanRaden. 2016. Genetic evaluation for cow livability. Accessed: November 30, 2016. https://www.cdc.us/reference/GENETIC_EVALUATION_FOR_COW_LIVABILITY.pdf.
- Rendel, J. M. and A. Robertson. 1950. Some aspects of longevity in dairy cows. *Empire Journal of Experimental Agriculture* 18:49-56.

- Sasaki, O., M. Aihara, A. Nishiura, H. Takeda, and M. Satoh. 2015. Genetic analysis of the cumulative pseudo-survival rate during lactation of Holstein cattle in Japan by using random regression models. *J. Dairy Sci.* 98:5781-5795.
- Schaeffer, L. R. 1994. Multiple-country comparison of dairy sires. *J. Dairy Sci.* 77:2671-2678.
- Schaeffer, L. R. 2004. Application of random regression models in animal breeding. *Livest. Prod. Sci.* 86:35-45.
- Sewalem, A., F. Miglior, G. J. Kistemaker, P. Sullivan, G. Huapaya, and B. J. Van Doormaal. 2007. Short communication: Modification of genetic evaluation of herd life from a three-trait to a five-trait model in Canadian dairy cattle. *J. Dairy Sci.* 90:2025-2028.
- Stoop, W. M., H. Eding, M. L. van Pelt, L. C. M. de Haer, and G. de Jong. 2013. Using pseudo-observations to combine genomic and conventional data in the Dutch-Flemish national evaluation. *Interbull Bulletin*.
- Stoop, W. M., H. Eding, M. L. van Pelt, and G. de Jong. 2011. Combining genomic and conventional data in the Dutch national evaluation. *Interbull Bulletin*.
- Tsuruta, S., D. A. L. Lourenco, I. Misztal, and T. J. Lawlor. 2015. Genotype by environment interactions on culling rates and 305-day milk yield of Holstein cows in 3 US regions. *J. Dairy Sci.* 98:5796-5805.
- Van Pelt, M. L. and G. de Jong. 2011. Genetic evaluation for direct and maternal livability in the Netherlands. *Interbull Bulletin* 44.
- Van Pelt, M. L., H. Eding, P. Vessies, and G. de Jong. 2012. Developing a genetic evaluation for calf survival during rearing in the Netherlands. *Interbull Bulletin* 46.
- VanRaden, P. M., C. P. Van Tassell, G. R. Wiggans, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and F. S. Schenkel. 2009. Invited review: Reliability of genomic predictions for North American Holstein bulls. *J. Dairy Sci.* 92:16-24.
- Veerkamp, R. F., S. Brotherstone, B. Engel, and T. H. E. Meuwissen. 2001. Analysis of censored survival data using random regression models. *Anim. Sci.* 72:1-10.
- Visscher, P. M. and M. E. Goddard. 1995. Genetic parameters for milk yield, survival, workability, and type traits for Australian dairy cattle. *J. Dairy Sci.* 78:205-220.
- Wright, J. R. and P. M. VanRaden. 2016. 0368 Genetic evaluation of dairy cow livability. *J. Anim. Sci.* 94:178-178.

Summary

Longevity is an important trait both from an economical and welfare point of view. Improving longevity helps to increase the profit of the farmer, and it is seen as an important measure of improved animal welfare and sustainability of the sector. Longevity is a complex trait, since the true longevity of a cow is available at the end of her life, whereas the selection decisions are made earlier in life. Breeding values for longevity have been published since the introduction in 1999 in the Netherlands, and have become very important due to the inclusion of the breeding values in the national selection index. It is necessary that breeding values are accurately estimated and will remain stable for the rest of life. However, current breeding values of bulls seem to fluctuate more than can be expected. This perception was amplified by some famous bulls that were highly favorable for longevity initially, but dropped considerable in their breeding value when more information became available. The main aim of this thesis was to revisit the genetics of longevity and develop a genetic evaluation model for the prediction of breeding values for longevity, with the objective that breeding values reflect the true breeding value quicker during early life and therefore breeding values become more stable.

In Chapter 2, genetic parameters were estimated for survival across the life up to 72 months after first calving using different survival models. Both linear and threshold models were used, and 4 different lengths of time intervals for survival were evaluated. The aim was to investigate whether genetic evaluation models need to take account of survival being genetically a different trait across the entire lifespan of a cow. A random regression model (RRM) with second-order Legendre polynomials was fitted for the additive genetic effect. Alternative parameterizations were (1) different trait definitions for the length of time interval for survival after first calving (1, 3, 6, and 12 mo); (2) linear or threshold model; and (3) differing the order of the Legendre polynomial. The partial derivatives of a profit function were used to transform variance components on the survival scale to those for lifespan. Survival rates were higher in early life than later in life (99 vs. 95%). Results indicated that survival is genetically not the same trait across the entire lifespan after first calving because genetic correlations differ from unity between different time intervals, especially when intervals are further apart. Heritability for survival after first calving within a time interval (1, 3, 6, and 12 mo) were low. However, heritability for lifespan across the entire period of 72 mo after first calving was higher and ranged from 0.115 to 0.149. To estimate breeding values for lifespan, the length of the time interval used for survival is important. An improved model for genetic evaluation should treat survival as different traits during the lifespan by splitting lifespan in time intervals of 6 mo or less to avoid

overestimated reliabilities and changes in breeding values when daughters are getting older.

In Chapter 3, survival during the first year after first calving was investigated over the last 25 years (1989–2013) as well as the association of survival with season of calving, age at first calving (AFC) and within-herd production level over that period. Survival and functional survival were analyzed. Functional survival was defined as survival adjusted for within-herd production level. Survival rate increased by 8% in the last 25 years up to 92% in 2013. Genetically, survival increased 3% to 4% but functional survival did not increase over the 25 years. An interesting difference was found between the genetic trends for survival and functional survival for bulls born between 1985 and 1999. Survival still increased, but was negative for functional survival. Since 1999, genetic trend picked up again for both survival and functional survival. AFC, season of calving and within-herd production level affected survival, and there was an interaction between these effects. The average decrease in survival rate per extra month AFC was 0.6% for survival and 1.5% for functional survival between AFC of 24 and 32 months. Within herd, low-producing cows had a lower survival rate than high-producing cows. However, the effect of AFC and production level became less important during the recent years. Based on survival the optimum AFC is around 24 months, but based on functional survival it is better to have an AFC < 24 months. Overall, survival rate of heifers has improved considerably in the past 25 years, initially due to the focus on a high milk production. More recently, the importance of a high milk production has been reduced towards attention for functional survival.

In Chapter 4, genetic correlations and heritabilities for survival were investigated over a period of 25 years to evaluate if survival in first lactation has become a different trait in this period. The analyses demonstrated that mean survival increased, whereas genetic variances and heritability decreased. Genetic correlations between 5-yr intervals were below one, showing that survival changed over time, whereas genetic correlations for functional survival did not indicate that survival changed. The difference in genetic correlations between survival and functional survival are likely explained by less emphasis of dairy farmers on culling in first lactation for production in more recent years. This suggests that genetic evaluation for longevity using historical data should analyze functional survival rather than survival.

In Chapter 5, the findings from Chapter 2 to 4 were used to develop a new genetic evaluation for longevity based on a RRM in comparison with the currently used proportional hazards model (PHM). The aim was to evaluate the impact of these changes to the model by: 1) fitting multiple genetic effects across the life of a cow;

2) fitting time-dependent fixed effects; and 3) adjusting for within-herd production level on the stability of breeding values for longevity of bulls. The trait analyzed was survival per month, fitted with a fifth-order Legendre polynomial until 72 months after first calving. The stability of breeding values was analyzed by checking the difference between the first breeding value a bull had received since 2006 with his breeding value in 2016. Based on the correlation between first breeding value and later breeding values, the ranking of bulls was shown to be more stable for RRM than for PHM. Bias in breeding value was observed, mainly due to incomplete daughter information. Adjusting for within-herd production level reduced this bias in the breeding values greatly, to a level similar to that for PHM. Before implementing this new model for genetic evaluation, the cause of this bias needs to be further investigated.

In the general discussion, Chapter 6, the proposed model was first discussed within the context of requirements for a new genetic evaluation model and other models used worldwide. Secondly, longevity and its association with other traits were examined, and thirdly, practical considerations for the implementation of a routine genetic evaluation of longevity based on the model described in Chapter 5 were put forward. The genetic evaluation model should ideally be able to take into account all the information available on dairy cows (right-censored records from the early life of young heifers); it should allow for fixed effects to vary over time, to account for different reasons for culling during the lifespan of animals (fitting more than one genetic effect over life), to handle the non-normal distribution of survival, to use an animal model to large data sets, and ideally to include information on genomics and other predictor traits collected during early life. The developed model fulfils most of these specifications, though including genomic and predictor information needs further research. Genetic correlations between (functional) longevity and predictor traits showed that udder health and claw health are the best predictors for both longevity and functional longevity. The third best predictor is production for longevity and fertility for functional longevity. Selection index calculations indicated that predictor traits can add accuracy to the breeding value for longevity, especially for young genomic bulls. It is advised to present the breeding value on a different scale than currently used. The current breeding value is presented in days, but presenting it in months is preferred, to avoid the suggestion that breeding values are very accurate and precise, and feeding the perception of farmers what a big change in breeding value is. Moreover, for publishing breeding values for longevity, it is arbitrary to model survival at the level of lactation or survival over the entire lifespan, as long as nonunity genetic

Summary

correlations between survival of different periods are taken into account to reduce bias in the estimation of breeding values.

Curriculum Vitae

About the author

Mathijs van Pelt was born on July 11, 1980 in Zeven, Germany. He was raised in Ugchelen and obtained his high school diploma in 1999 at Chirstelijke scholengemeenschap Sprengeloo in Apeldoorn. In that same year he started his studies Animal Sciences at Wageningen University. He fulfilled both his major and minor thesis at the Animal breeding and genetics group. His major thesis was entitled "Relationship between somatic cell score and clinical mastitis across and within parities in Dutch Holstein cattle", and his minor thesis was entitled "Estimation of variance components in selected data". Mathijs spent his internship in 2003 at Ambreed in Hamilton, New Zealand, where he worked on "Accuracy of selection of Ambreed's young bulls". After finishing his master studies, he continued working at the Animal breeding and genomics center, and under supervision of Johan van Arendonk he developed a new total merit index (NVI) to rank dairy cattle in the Netherlands and Flanders. In 2006, Mathijs started to work at the animal evaluation unit of CRV and is currently senior researcher. In this capacity, he has worked on updating, development and implementation of genetic evaluations for calving traits, calf survival during the first year, claw health and conformation traits. In 2013, Mathijs was accepted as a PhD-candidate at the Animal breeding and genomics center of Wageningen University. During his PhD, Mathijs investigated longevity in dairy cattle and developed a new genetic evaluation for the prediction of breeding values for longevity, as published in this thesis. Currently, Mathijs is working on the implementation of the new genetic evaluation for longevity in the Netherlands and Flanders.

Publications

Peer reviewed publications

- Van Pelt, M. L.**, V. Ducrocq, G. de Jong, M. P. L. Calus, and R. F. Veerkamp. 2016. Genetic changes of survival traits over the past 25 yr in Dutch dairy cattle. *Journal of Dairy Science* 99:9810-9819.
- Van Pelt, M. L.**, G. de Jong, and R. F. Veerkamp. 2016. Changes in the genetic level and the effects of age at first calving and milk production on survival during the first lactation over the last 25 years. *animal* 10:2043-2050.
- Van Pelt, M. L.**, T. H. E. Meuwissen, G. de Jong, and R. F. Veerkamp. 2015. Genetic analysis of longevity in Dutch dairy cattle using random regression. *Journal of Dairy Science* 98:4117-4130.

Manuscript in preparation

- Van Pelt, M. L.**, G. de Jong, and R. F. Veerkamp. Genetic evaluation of longevity using a random regression animal model.

Other publications and abstracts

- Van Pelt, M. L.** and R. F. Veerkamp. 2015. Evaluation of survival in the first year after calving across years and seasons. in Proc. ADSA-ASAS Joint Annual Meeting, Orlando, Florida.
- Van Pelt, M. L.** and R. F. Veerkamp. 2014. Genetic Analysis of Longevity in Dutch Dairy Cattle using Random Regression. in Proc. 10th World Congress on Genetics Applied to Livestock Production. Vancouver, Canada.
- Van Pelt, M. L.**, R. F. Veerkamp, and G. de jong. 2013. Predicting lifespan of dairy cows: phenotypic and genetic change during life. in Proc. 64th Annual meeting of the European Federation of Animal Science, Nantes, France.
- Van Pelt, M. L.**, H. Eding, P. Vessies, and G. de Jong. 2012. Developing a genetic evaluation for calf survival during rearing in the Netherlands. *Interbull Bulletin* 46:61-65.
- Van Pelt, M. L.**, E. P. C. Koenen, C. van der Linde, G. de Jong, and H. G. Geertsema. 2012. Selection for longevity in Dutch dairy cattle. in Proc. 63rd Annual meeting of the European Federation of Animal Science, Bratislava, Slovakia.
- Van Pelt, M. L.** and G. de Jong. 2011. Genetic evaluation for direct and maternal livability in the Netherlands. *Interbull Bulletin* 44:235-239.

Publications

- Van Pelt, M. L.** and G. de Jong. 2011. Genetic analysis of calf livability in Dutch dairy cows. in Proc. 62nd Annual meeting of the European Federation of Animal Science, Stavanger, Norway.
- Van Pelt, M. L.**, H. Eding, and G. de Jong. 2010. Reduction of bias in breeding values of bull dams for conformation traits. in 9th World Congress on Genetics Applied to Livestock Production. Leipzig, Germany.
- Van Pelt, M. L.**, G. de Jong, H. Eding, and J. E. Roelfzema. 2009. Analysis of calving traits with a multitrait animal model with a correlated direct and maternal effect. *Interbull Bulletin* 40:138-141.
- Van Pelt, M. L.** and G. de Jong. 2008. Robustness: breeding for optimum traits. in Proc. 59th Annual meeting of the European Association for Animal Production, Vilnius, Lithuania.
- Van Pelt, M. L.**, C. van der Linde, A. G. F. Harbers, and G. de Jong. 2006. Consistency of conformation trait definitions across time. *Interbull Bulletin* 35:159-163.

Co-authored publications

- Van der Laak, M., **M. L. van Pelt**, G. de Jong, and H. A. Mulder. 2016. Genotype by environment interaction for production, somatic cell score, workability, and conformation traits in Dutch Holstein-Friesian cows between farms with or without grazing. *Journal of Dairy Science* 99:4496-4503.
- Manzanilla-Pech, C. I. V., R. F. Veerkamp, R. J. Tempelman, **M. L. van Pelt**, K. A. Weigel, M. VandeHaar, T. J. Lawlor, D. M. Spurlock, L. E. Armentano, C. R. Staples, M. Hanigan, and Y. de Haas. 2016. Genetic parameters between feed-intake-related traits and conformation in 2 separate dairy populations-the Netherlands and United States. *Journal of Dairy Science* 99:443-457.
- Ouweltjes, W., J. J. Windig, **M. L. van Pelt**, and M. P. L. Calus. 2015. Genotype by environment interaction for livability of dairy calves from first parity cows. *Animal* 9:1617-1623.
- Stoop, W. M., H. Eding, **M. L. van Pelt**, L. C. M. de Haer, and G. de Jong. 2014. Using Pseudo-observations to combine Genomic and Conventional Data in the Dutch National Evaluation. in 10th World Congresson Genetics Applied to Livestock Production. Vancouver, Canada.
- Stoop, W. M., H. Eding, **M. L. van Pelt**, L. C. M. de Haer, and G. de Jong. 2013. Using pseudo-observations to combine genomic and conventional data in the Dutch-Flemish national evaluation. *Interbull Bulletin* 47:106-110.

- Stoop, W. M., H. Eding, **M. L. van Pelt**, and G. de Jong. 2011. Combining genomic and conventional data in the Dutch national evaluation. *Interbull Bulletin* 44:169-172.
- De Haer, L. C. M., G. de Jong, **M. L. van Pelt**, and J. E. Roelfzema. 2010. Investigation of interval first to last insemination for dairy cattle in the Netherlands. *Interbull Bulletin* 42:100-103.
- Stoop, W. M., G. de Jong, **M. L. van Pelt**, and C. van der Linde. 2010. Implementation of a claw health index in the Netherlands. *Interbull Bulletin* 42:95-99.

Training and supervision plan

Training and supervision plan



The Basic Package (3 ECTS)

WIAS Introduction Course	2013
Ethics and Philosophy in Life Sciences	2015

Scientific exposure (10.6 ECTS)

International conferences

64 th annual meeting of European Federation of Animal Science (EAAP), Nantes, France	2013
7 th International Cattle Breeders Round Table (ICBRT), Odalgården, Sweden	2014
10 th World Congress on Genetics Applied to Livestock Production (WCGALP), Vancouver, Canada	2014
Interbull meeting and Joint Annual Meeting ADSA-ASAS, Orlando, Florida, USA	2015
Intebull meeting and meeting of International Committee of Animal Recording (ICAR), Puerto Varas, Chile	2016

Presentations

Oral presentation at 64 th EAAP	2013
Oral presentation at 7 th ICBRT	2014
Oral presentation at 10 th WCGALP	2014
Oral presentation at Joint Annual Meeting ADSA-ASAS	2015

In-Depth Studies (9 ECTS)

Disciplinary and interdisciplinary courses

Statistics for the Life Sciences (Fortran 90), Lelystad, the Netherlands	2007
Synbreed Summer school "Quantitative traits: Advanced topics in plant and animal breeding", Herrsching, Germany	2013
Genetic analysis using ASReml4.0, Wageningen, the Netherlands	2014
In-depth Course Genotype by environment interaction, uniformity and stability, Wageningen, the Netherlands	2015

Training and supervision plan

PhD students' discussion groups

Quantitative Genetics Discussion Group (QDG), Wageningen, the Netherlands 2013-2016

Professional Skills Support Courses (8.2 ECTS)

NCOI Projectmanagementvaardigheden, Utrecht, the Netherlands 2010

Techniques for writing and presenting a scientific paper, 2014

Wageningen, the Netherlands

Workshop Societal Impact, Wageningen, the Netherlands 2015

WGS PhD Workshop Carousel, Wageningen, the Netherlands 2016

Research Skills Training (6 ECTS)

Preparing own PhD research proposal 2013

Management Skills Training (1.5 ECTS)

Organisation of seminars and courses

Organizing Quantitative Genetics Discussion Group, Wageningen, the Netherlands 2014-2016

Training and Education total

39 ECTS

Acknowledgments

About five years ago I was still under the impression that I would never do a PhD. How wrong was I? Well, it all changed during a dinner with Yvette when she asked if I would consider doing a PhD. I quickly said no, but after some thoughts the idea became more appealing. At that time, I was already working for seven years in the Animal Evaluation Unit of CRV, developing new genetic evaluations for dairy cattle. For me, the added value in that work that it is in between science and practice; the developments in genetic evaluations are science-driven, and the created output is used by AI-companies and farmers to breed for better cows. The project I worked on for the past four years was also focused to develop an application to be used in practice. Although I am still working on the application, I have learned a lot during these four years about science, genetics, as well as personally.

I would like to thank all the people that helped and supported me during my PhD. First of all I want to thank Roel and Ate for selecting me for this project and for your trust in me. Secondly I want to thank the members of my supervisor committee – Roel, Gerben, Mario, Vincent and Theo – for the many constructive discussions and to get a better understanding on what is happening with longevity. Roel, as my daily supervisor and promoter, I had even more discussions with you. You have been a great stimulator for me to dig deep in this topic and broaden my knowledge, as well as improving my writing skills. Gerben, I am grateful for the opportunity you gave me to work outside CRV for some years. You have always been interested in what I was doing and moreover I could pick your brain on little details that no one else would remember. Mario, Vincent and Theo, thank you for the time you invested in reading the manuscripts, discussing results, and proposing other directions. You, being experts in the field of genomics or longevity, I was a bit frightened in the beginning. However, your input has helped me to gain more insight in longevity.

Now we are one group ABGC working in Radix, but when I started we were still in two separate buildings Triton for Livestock Research and Zodiac for Animal Breeding and Genetics. The time in Triton has been a good time and the mixture with other disciplines was refreshing. The move to Radix was necessary and being in a bigger group has advantages as well. The group got too big to thank everyone by name, but I thank everyone – staff, secretaries (Ada, Lisette and Maya) and PhD-candidates – of ABGC working in Radix. Still I would like some colleagues by name: firstly the Triton-ladies Amabel, Coralia, Rianne, Sonia, and Yvonne for the fun we had at Triton. Amabel, I was honored to share the office with you, having so many interesting conversations and I am grateful to have you as a paranymph. Sonia, you

Acknowledgments

are the coolest French girl I know and I am grateful to have you as a paranymph as well. Yvette, thank you for many nice dinners and personal talks. Further thanks go to all the friends I made during the time in Radix: André, Claudia, Ewa, Floor, Gabriel, Gareth, Hadi, Hamed, Jeremie, Juan, Juanma, Kasper, Pascal, Mandy, Marcos, María, Mathieu, Merina, Mirte, Sabine, Sanne, Tessa, and Zih-Hua. Foosball table and players, thank you for relaxing my brain especially during the writing process. Moreover, special thanks to Katrijn and Susan for all the nice dinners together, and hopefully we can keep on continuing.

Further I want to thank my colleagues – Herwin, Irma, Jorien, Lydia, Marianne en Pedro – from the Animal Evaluation Unit at CRV for the interest in my research and the patience you had to miss me four days a week. I appreciated to go to Arnhem one day a week to stay connected to what was happening in practice.

Thank you mum and dad for believing in me, showing interest and supporting me to follow my dream to work in cattle genetics. Gertjan, Viola and the children, seeing your family helped to take distance from the PhD. Also a special thanks to my oom Jaap and tante Riet van Leeuwen, thank you for all the holidays working on the farm and being on the beginning of my passion for the dairy sector.

All my friends in the Netherlands, with some of you contact has been less, but I hope to see you again. I want to thank some people specifically: Jolien and Tim, for all the dinners and not only talking about cows. Noortje and Nico, for the nice dinners with homegrown vegetables and milking at your farm to stay connected with real farm life. Peter and Julius, for long evenings talking about bulls.

Coralia, my love, it must have been destiny to have met you at the beginning of my PhD. It feels so normal to be with you. We share the same interest of travelling the world and together we have seen a lot of nice places. Now I have favorite place for holidays: Mexico!!! I am grateful to know your family and Byron, they have welcomed me from the start. Together we make a great team, hopefully for many more years to come!

Mathijs

The research described in this thesis was financially supported by the Dutch Dairy Board (ZuivelNL, Den Haag, the Netherlands), Genetic Evaluation for Sires (GES) and cattle improvement cooperative CRV (Arnhem, the Netherlands). This thesis was carried out as part of the Breed4Food project (program “Kennisbasis Dier,” code: BO-12-22.04-011-001-ASG-LR).

The data used in this thesis was provided by CRV (Arnhem, the Netherlands).

The cover of this thesis was designed by André Fris and Rogier van der Weiden.

This thesis was printed by Digiforce | Proefschriftmaken.nl, De Limiet 26, 4131 NC, Vianen, the Netherlands.