Breeding for longevity and survival in dairy cattle

Erling Strandberg and Johann Sölkner
Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences
PO Box 7023, S-75007 Uppsala, Sweden
and
Department of Livestock Science, University of Agriculture
Gregor-Mendel-Strasse 33, A-1180 Wien, Austria

Abstract

The aim of dairy cattle breeding should be to improve lifetime profit, rather than longevity in itself. It is most likely, though, that traits that determine longevity will be selected for, and thus the ability to live longer will improve. Health and reproductive traits are described as the most important in determining longevity. The main focus of the article is on describing methods used as indicators of longevity: stayabilities, survival scores, and failure time analysis.

Measures of production and involuntary culling (e.g. mastitis, reproduction) are suggested for use as selection criterion, if measured in the population. If not, other indicator traits, such as somatic cell counts, could be used instead. If no or few health and reproduction traits are available in a population, breeding values for length of productive life adjusted for the within-herd milk production deviation, analyzed with failure time analysis, could be used as a complement to milk production.

1. Introduction

The length of life of a dairy cow has substantial impact on the economic performance. The largest effect is probably that a longer life decreases the cost of replacement per year. Also, a longer average life will lead to a higher proportion of cows in later high-producing lactations. An increased length of productive life from about three to four lactations increased milk yield per lactation or profit per year by 11-13% (Renkema & Stelwagen, 1979; Essl, 1984).

There may also be additional benefits if the lengthened life is due to less culling for disease due to lower disease incidence. This would lead to lower treatment costs for diseases. However, given an improvement of disease resistance, it might be more beneficial to increase voluntary culling instead of increasing average length of life. Van Arendonk (1985, 1986) and Rogers et al. (1988) showed that when involuntary culling decreased it was best to increase voluntary culling so that the total effect on productive life was less than expected.

One of the first steps taken when starting a breeding program should be to establish the breeding goal. However, even before that, the overall objective needs to be established. For the dairy farmer it might be to achieve the greatest profit per year possible and for the society it could be to produce food as efficiently as possible, e.g. measured as cost per unit of energy, protein or per kg of milk. These are not breeding goals but objectives on a higher (and phenotypic) level. These overall objectives are sometimes called "profit functions", keeping in mind though, that they need not specify "economic profit".

The breeding goal reflects those parts of the overall objective that are possible to change genetically. The breeding goal itself is not longevity nor does it even have to include longevity as a trait per se as has been shown by Goddard (1989). For more information on how to calculate economic weights when considering longevity see Goddard (1989), Strandberg (1995) and Groen (these proceedings).

Estimates of the genetic correlation between first lactation milk yield and observed productive life have ranged from 0.3 to 0.9 (Strandberg, 1985; Harris et al., 1992; Short & Lawlor, 1992; Visscher & Goddard, 1994). However, several simulation studies have shown that these correlation estimates are severely biased due to the voluntary culling on milk yield and do not correctly describe the expected genetic response to selection (Essl, 1989; Strandberg, 1991; Dekkers, 1993; Strandberg & Håkansson, 1994).

To avoid this problem, length of productive life adjusted for milk yield has been suggested as the trait of choice, often termed functional productive life. When productive life is adjusted for level of milk yield, estimates of genetic correlation with milk yield
have been much lower, usually around zero (Sölkner, 1989; Ducrocq et al., 1988b; Harris et al., 1992; Short & Lawlor, 1992). Therefore, whenever examining an estimate of genetic correlation between milk yield and longevity, one should note whether the longevity measured is adjusted for milk yield or not.

These factors point to some of the complexities surrounding breeding for longevity. Another central problem when contemplating selection for improved longevity is that one has to be able to predict longevity with good precision early in life. Using actual life length measures (when all animals are dead) is impractical, to say the least. Therefore, this article will concentrate on describing the methods available to estimate breeding values for longevity using information early in life in order to minimize the generation interval.

2. Traits associated with longevity

One can distinguish between traits that actually determine longevity, i.e. traits that are part of the culling criteria, and traits that are only, or mainly, indicators of longevity. For another review, and an excellent one, see Dekkers & Jairath (1994).

2.1. Determinants of longevity

2.1.1. Health traits

Health traits (diseases and disorders) are naturally one of the main groups of traits affecting longevity of the dairy cow. Due to the lack of health trait registration in many countries there is a shortage of studies showing exactly how important diseases are in impairing longevity. One way to get an approximation is to study the culling reasons reported by the farmers. Shook (1989) concluded in a review that mastitis is the third most common reason for culling, after milk yield and reproduction. Dentine et al. (1987a) reported that of all cows culled, about one-third were culled for diseases, one-third of which were culled for mastitis. In Sweden 15-25% of those culled are culled for mastitis and another 15% are culled for other diseases (Svensk Husdjursskötsel, 1992).

The economic values of health traits have been reviewed by Rogers (1994) and by Groen (these proceedings). Most health traits have economic values of around 10% of that for milk yield, if expressed per genetic standard deviation. Mastitis is the exception with an economic value of about 25% of that for milk yield. This value includes cost for drugs, labour, veterinary treatment, discarded milk, and premature culling. Heuven (1987) found a higher value for mastitis in later lactations (70% for that for milk yield).

Even if mastitis is the trait included in breeding goal, selection could be on somatic counts (SCC) instead, if measurements of clinical mastitis are not available. In addition, heritabilities of SCC have usually been somewhat higher than those for clinical mastitis, even if the latter is correctly treated as a binomial trait (Emanuelson, 1988; Lin et al., 1989; Shook, 1989; Lyons et al., 1992). Including mastitis or SCC in the selection index improved the total response in the breeding goal by only 0.4-0.8% (Strandberg & Shook, 1989; Duval & Colleau, 1993). However, the breeding goal consists only of milk yield and mastitis traits, not longevity.

Dystocia (calving difficulty) is most common in first-calf heifers (Philippson et al., 1979; Erb et al., 1985). Heritability estimates have usually been higher for first calvings (0.03-0.20) compared to late calvings (0.00-0.08) (Philippson et al., 1979). Due to the commonly found negative genetic correlation between direct and maternal effects, it is recommendable to evaluate bulls both as sires and as maternal grandsires (e.g. Philippson, 1976; Thompson et al., 1981; Philippson, these proceedings).

Milk fever is very rare in first lactation cows, but when it occurs it may increase the risk of culling by more than five times (Gröhnl et al., 1986a). Heritability estimates for milk fever have been close to zero (Emanuelson, 1988) and for the first parities very difficult to estimate (Gröhnl et al., 1986b). Ketosis is also more frequent in later lactations and usually heritability estimates are low. For both traits, the low heritability and the low frequency in first lactation limits the use of these traits in selection.

Several of these diseases are unfavourably genetically correlated with milk yield. Mastitis and SCC are extensively studied and almost invariably a genetic antagonism with milk yield has been found (Emanuelson, 1988; Banos & Shook, 1990; Simianer et al., 1991; Boettcher et al., 1992; Wellner et al., 1992; Welp & Freeman, 1992). A similar situation seems to exist for ketosis (Emanuelson, 1988).

2.1.2. Reproduction

Reproduction is the largest or second largest culling reason in dairy cattle. 15-40% of those culled are culled for inadequate reproductive performance (Shook, 1989; Svensk Husdjursskötsel, 1992). The economic value for days open is approximately 10%.
2.2.3.1. Failure time analysis

Failure time analysis has been used before in the fields of medicine and engineering, e.g. to study the effect of various medical treatments on relapse, recovery or survival (e.g. Kalbfleisch & Prentice, 1980). This approach models the actual survival times or usually (because it is more convenient) the hazard, the risk of failure (culling) at a certain time given that the individual survived up to that time. The observations used are the failure times (e.g. death) or the censoring times (e.g. still alive at the end of data collection), combined with an indicator of whether the measure is censored or not.

Probably the most commonly used model is the proportional hazards model (e.g. Kalbfleisch & Prentice, 1980). The hazard \( \lambda(t; z) \) for time \( t \) and a set of known covariates \( z \):

\[
\lambda(t; z) = \lambda_0(t) e^{z\beta}
\]  

where the hazard is the product of a time-dependent term \( \lambda_0(t) \), called the baseline hazard, and a time-independent term \( e^{z\beta} \). The baseline hazard is related to the general aging process and the term \( e^{z\beta} \) depends on the covariates in \( z \), e.g. effect of treatment, animal etc. Two of the most common assumptions about the baseline hazard are a) that it is constant \( (\lambda_0(t) = \lambda) \), in which case the survival times follow an exponential distribution, and b) that it is \( \lambda_0(t) = \lambda \cdot e^{z\beta} \), for some values of the parameters \( \lambda \) and \( \beta \), in which case the survival times follow a Weibull distribution. Examples of the two distributions are given in Fig. 1.

Cox (1972) simplified [1] by showing that \( \lambda_0(t) \) need not be specified. Smith (1983) and Smith & Quasas (1984) studied age at disposal using an extension of the Cox’s regression model:

\[
\lambda_{jk}(t) = \lambda_0(t) \exp \{ h_k + g_l + s_{lm} \}
\]  

where: \( \lambda_0(t) \) is a piecewise constant baseline hazard, stratified by year of birth \( j \); \( h_k \) is the \( k \)th herd effect nested within the \( j \)th year of birth; \( g_l \) is the \( l \)th genetic group; and \( s_{lm} \) is the \( m \)th sire within the \( l \)th group. In contrast to traditional failure time analysis where all effects are considered fixed, here both herd and sire were assumed to be random.

Another new feature of this model was the stratification, i.e. they divided the material into strata, here based on year of birth. Each stratum was allowed to have a different baseline hazard. This method of stratification can be used if a certain factor does not seem to act multiplicatively on the hazard, as it would if added to the model component in an exponent (e.g. Kalbfleisch & Prentice, 1980).

The next large study using failure time analysis was the thesis of Ducrocq (1987; summarized Ducrocq et al., 1988a, b). Smith (1983) and Smith & Quasas (1984) used age as the time scale. Because there is virtually no culling for around two years, it is almost impossible for any parametric distribution to fit that survival or hazard function well. To avoid this problem, Ducrocq (1987) and Ducrocq et al. (1988b) used length of productive life (time from first calving to disposal). Also, instead of deciding a priori to use an exponential distribution, they estimated the parameters \( p \) and \( \lambda \) of the Weibull distribution (of which the exponential is a special case) to get the best fit possible.

Another major extension was the inclusion of time-dependent effects. In equation [2] the stratification is according to herd within year of birth and herd effect \( (h_k) \) is constant throughout the life of the cow. However, Ducrocq (1987) argued that it is more likely that each production year has a specific effect on the hazard of all cows alive at that time. For instance, perhaps the herd size is increasing during a few years. The hazard of cows alive during that period, regardless of when they were born, is lower than for cows during other periods. To accommodate this, Ducrocq (1987) included a herd effect that was piecewise constant, changing every new year.

To account for a changing herd size, Ducrocq (1994) also included a time-dependent effect which was a combination of season and class of change in herd size the previous year. The effects were assumed to change at 1st March and 1st December each year. The date 1st March was chosen because the quota period ended 31st March. The risk of culling was higher in the months just before the end of the quota period compared with during the rest of the year. Also, in herds of decreasing size, the hazard was much higher than in herds of increasing size.

The effect of within-herd class of milk yield was included in an attempt to estimate length of productive life adjusted for voluntary culling, termed by Ducrocq (1987) “functional productive life”. The correlation between sire estimates for length of productive life (actually relative risk of culling) and estimated breeding values for milk yield changed from favourable (-0.28) to slightly unfavourable (0.13) with the adjustment for within-herd class of milk yield. This result was not repeated in the study by Ducrocq (1994), both correlations were -0.40. One possible explanation put forward was that the breed used in Ducrocq (1994), Normande, is a dual-
of that of milk yield (Rogers, 1994; Groen, these proceedings).

Estimates of heritability for reproductive measures have been low, below 10% (Schaeffer & Henderson, 1972; Berger et al., 1981; Philipsson, 1981; Hansen et al., 1983a; b; Freeman, 1984, 1986; Strandberg & Danell, 1989; Lyons et al., 1991; Fuerst & Sölkner, 1994). Nevertheless, there is a substantial genetic coefficient of variation, 3-18% (Philipsson, 1981; Strandberg & Danell, 1989), which indicates that it is possible to distinguish between sire progeny groups. Furthermore, there is evidence of genetic antagonism with milk yield (Berger et al., 1981; Philipsson, 1981; Hansen et al., 1983a; Strandberg & Danell, 1989; Oltenacu et al., 1991) suggesting that reproduction (and thus longevity) will deteriorate if selection is on milk yield only.

2.2. Indicators of longevity for genetic evaluation

Type traits (conformation, workability) are sometimes useful indicator traits for longevity. However, because these will be discussed elsewhere in these proceedings (Rogers; Vollema; Weigel), in this article we will concentrate on indicator traits more directly associated with survival and longevity.

2.2.1. Stayability

Stayability is probably the most commonly used indicator of longevity. This trait is measured as the survival (0 or 1) to start a certain lactation (Robertson & Barker, 1966; Schaeffer & Burnside, 1974), up to a certain age (Everett et al., 1976a, b; Hudson & Van Vleck, 1981; Delorenzo & Everett, 1982; Van Doormaal et al., 1985) or up to a certain time after first calving (Van Doormaal et al., 1985). The ages used have been 36, 48, 60, 72 and 84 months, probably chosen assuming 24 months at first calving and a 12-month calving interval. Van Doormaal et al. (1985) used a different approach to determine the time when to measure stayability. They studied the hazard (the relative risk of culling) during the cow's life and chose the periods of low risk of culling as the time to measure stayability. This resulted in stayability up to 42, 54, 66, and 78 months of age, and up to 17, 30, 43, and 55 months from first calving. Delorenzo and Everett (1986) chose 41 and 54 months, being the average ages in the middle of the dry period in the first two calving intervals, respectively.

Although stayability is a binary trait most studies have used linear models. One exception is DeLorenzo and Everett (1986) who used a logistic linear model. They found quite substantial differences in ranking of sires between the logistic model and the linear model; rank correlations were 0.6-0.7. Nevertheless, nobody seems to have used this model further, perhaps due to computational limitations.

Heritability estimates for stayability have been low, usually less than 0.05 (Strandberg, 1985; Dentine et al., 1987b; Brotherstone & Hill, 1991; Short & Lawlor, 1992; Visscher & Goddard, 1994). Unfortunately for selection purposes, there was also a tendency to lower heritabilities for earlier measures, the ones mainly used in selection. The only exception was DeLorenzo and Everett (1986) who estimated heritabilities of around 0.27 for survival to 41 or 54 months of life using a logistic linear model.

2.2.2. Survival within each lactation

Instead of using stayability Madgwick and Goddard (1989) used a series of survival scores, \( S_i \), where \( S_i = 1 \) if the cow survived from year \( i \) to year \( i+1 \) after first calving and \( S_i = 0 \) otherwise.

Heritabilities for survival scores were low, the highest were between 0.028 and 0.053 (for \( S_2 \)). They found very high correlations between sire solutions based on linear analysis of survival scores and a nonlinear analysis. Genetic correlations among survival scores were around 0.8 for \( S_0 \) to \( S_2 \). The genetic correlations between production in first lactation and survival scores were highest for \( S_3 \) to \( S_5 \), between 0.26 and 0.63, depending on the production trait. One of the "problems" with their analysis was that it was performed on Australian cattle, which have very long productive lives: the average in their study was 5.8-6.6 years. The very low culling from one lactation to the next might have affected the performance of the method. Visscher and Goddard (1994) also found low heritabilities for survival scores in Australian Friesian, but somewhat higher in Jerseys (0.07-0.08).

2.2.3. Length of productive life

Length of productive life itself is not a useful measure for selection because of the long time before it is realized. For many animals one would only know that length of life is at least as long as a certain time. Such records are called censored. Different methods have been used to accommodate censoring.
2.2.4. Properties of the different methods of genetic evaluation

One important property of methods used for genetic evaluation of longevity is the ability to account for censoring. For failure time analysis, censoring is accounted for directly within the method but not in the other methods.

In the first two methods the problem with censoring is at least partially avoided by the definition of the trait. For genetic evaluation using stayability a short period is chosen, e.g. 17 months of productive life and all animals are given the opportunity to live that long and no animals are censored. Although this procedure avoids censoring it means that not all information in the data is utilized. Cows culled 1 day or 1 year before the limit are treated as identical. Danner et al. (1993) and Egger-Danner (1993) compared sire breeding values estimated using a Cox's regression model and sire breeding values for stayability to different ages using an ordinary linear model. In both models relationships were included, and adjustment for within-herd milk yield deviation was carried out. The full data set was divided according to year of first calving and then truncated to simulate different amounts of censoring. The rank correlation between breeding values estimated on the full data set or on the censored data set using the Cox's regression model were higher than when applying the corresponding stayability model. The difference was largest when censoring was high, i.e. for stayability to 16 months of productive life and its corresponding censored data set including two years of data.

The ability to account for systematic environmental effects is also different for the described methods. In the analysis of stayability and extended censored records of longevity, the environmental effects included usually pertain to first calving and are assumed constant regardless of the actual longevity. Both with failure time analysis and the traits survival within each lactation (Madgwick & Goddard, 1989), it is possible to have time-dependent environmental effects, which should improve the possibility to account for a changing environment.

The genetic model is somewhat different for the various methods. In failure time analysis the implicit assumption is that survival is the same throughout life. In the other methods, it is possible to study each measure (stayability, survival through different lactations) as separate genetic traits which, at least conceptually, is an advantage.

The adjustment for milk yield to achieve a measure of functional productive life should be possible in all methods, either by preadjustment or by inclusion of milk yield in the model. Because culling is within herd and year, the milk yield should deviated from the herd-year mean.

3. Conclusions

It is important to remember that our aim should not be to improve longevity in itself. Our aim should be to improve the overall objective, which may be expressed as lifetime profit, efficiency or some other measure of utility. In doing so, we will probably also improve the animals' ability to live longer and improving traits that determine longevity. However, as pointed out before, the actually observed longevity may not change at all or may not change as much expected from the changes in the other traits.

It is not quite clear which approach is the most beneficial for decreasing reasons for involuntary culling and thus improving longevity - using the actual determinants of longevity as selection criteria, i.e. traits like mastitis, conception rate etcetera, or using indicator traits such as type traits, stayability, or longevity measures using failure time analysis. Of course, a combination of these approaches is possible. However, one should be somewhat cautious of including all measured traits in the selection criterion. If the true parameters are known, the precision of evaluation always increases with increasing number of traits. However, parameters are estimated with error. As pointed out by Sales & Hill (1976) and Visscher (1994), the precision might even decrease if traits are included that have weak true correlations to the goal but that have estimates (with large error variances) that indicate otherwise.

Naturally, the availability of records is important. Therefore, it might be necessary to use indicator traits rather than the actual determinant traits. For instance, if somatic cell counts are available one could use that trait: it might be as good or better for predicting mastitis as mastitis itself. For prediction of reproduction in the breeding goal, it might be valuable to use both calving to first insemination interval and number of inseminations, because they describe two, sometimes opposing, sides of days open.

In our minds, one cannot speak of a breeding programme aimed at decreasing involuntary culling and improving longevity unless it contains at least one criterion for mastitis and one for reproduction. If no or few health and reproduction traits are recorded in the population, one solution could be to use evaluations of longevity as an overall measure of health, reproduction, and all other involuntary culling. However, these longevity measures should be adjusted for the within herd-year production deviation. It seems beneficial to use failure time
purpose breed with voluntary culling for other traits as well.

Estimating heritabilities using failure time analysis is not as straightforward as for ordinary linear models. Due to the methods and the censoring it is not possible to estimate the phenotypic variance in an unambiguous way. However, making some simplifying assumptions, Ducrocq et al. (1988b) estimated a "pseudo-heritability" of 0.085.

The future developments in this area should focus on incorporating the relationship between sires into the model in a similar manner as for the ordinary mixed models. Theoretically, it should be possible to set up an animal model also for this type of model, however, the predictive power and use of such individual animal breeding values might be limited. Ducrocq and Sölkner (1994) are currently working on a set of computer programs to be used for failure time analysis of animal breeding data. A Cox regression model is available for model testing and analysis purposes. For prediction of breeding values and estimation of variance components the parametric Weibull model may be better suited due to lower computational demands. Both programs may incorporate relationships between individuals.

2.2.3.2 Extension of censored records

VanRaden and Klaaskate (1993) tried to accommodate censored records while avoiding the complexity of the nonlinear methods used in failure time analysis. Extension of incomplete (censored) records is a common method when studying 305-day milk yield, and the idea here was to use a similar procedure for length of productive life. The predicted records were expanded to give the same variance as for completed records and the predicted records were given a lower weight in the mixed model analysis later. The R²-values for predicting their longevity measure from earlier ages were quite low except at 72 months of age. However, the genetic correlations between longevity from completed records only and longevity from combined projected and completed records were all quite high, above 0.92.


analysis because of its better ability to use the information in lifetime data, and to account for censoring and changing systematic environmental effects.

References


Everett, R.W., Keown, J.F. and Clapp, E.E. 1976a. Relationships among type, production and