

# Genetic parameters of functional and fertility-determined length of productive life in Swedish dairy cattle

*Erling Strandberg and Anki Roth*

*Centre for Reproductive Biology  
Department of Animal Breeding and Genetics,  
Box 7023, Swedish University of Agricultural Sciences,  
SE-750 07 Uppsala, Sweden,*

---

## Abstract

The objective of this study was to estimate genetic parameters and breeding values of sires for different definitions of length of productive life in Swedish Red and White dairy cattle. The data consisted of 534 016 daughters with initial calvings between 1988 and 1996. These cows were daughters of 1266 sires (55 of which were considered proven bulls, and treated as fixed effects) and another 314 bulls were included in the pedigree information. The model for the hazard included: a random time-dependent effect of herd-year-season (*hys*, seasons from Jan-Jul and Aug-Dec), a fixed time-dependent effect of parity by stage of lactation (parities from 1 to 6<sup>+</sup>, stages changed at 0, 60, 180, 270, and 365 days of lactation), a fixed time-dependent effect of the cow's peak yield as deviation from her herdmates in that herd-year, a fixed time-independent effect of age at first calving, and a random effect of sire. The *hys* effect was assumed to follow a gamma distribution and the sire effect a normal distribution with mean zero and variance  $A\sigma_s^2$ , where  $\sigma_s^2$  is the variance among sires. Length of productive life was defined as number of days from first calving to culling or end of data. Two types of length of productive life were studied: 1) PL, all cows that were culled before the end of data were considered as uncensored; 2) Fertility-determined productive life, FPL, only cows that were culled for fertility problems were considered as uncensored. Heritability estimates were 6.94% and 6.13% for PL and FPL, respectively. The approximated genetic correlation estimate between calving and last insemination and PL was 0.47, and the corresponding estimate with FPL was 0.55. In conclusion, it was feasible to use the trait fertility-determined length of productive life, its heritability was almost as large as for PL, and the genetic correlation was higher with fertility than for PL.

---

## 1. Introduction

Length of productive life, usually measured as time from first calving to death, describes the ability of the cow to avoid culling done by the farmer. If the measure is adjusted for within-herd production deviation, we usually call it functional productive life, and this trait describes the cow's ability to avoid involuntary culling.

The reasons for involuntary culling may range from fertility and health problems to severe injuries and accidents. Naturally, the so-defined longevity trait is a conglomerate of all these reasons. Therefore, it is reasonable to believe that one can divide the trait functional productive life into smaller components, each describing a separate trait complex (Strandberg, 1997).

We have defined one such component, namely fertility-determined length of productive life. The basic assumption behind this trait is that we only consider a cow as being dead if she has been culled

for fertility-related problems, otherwise we assume her observation to be censored. Our hypothesis is that this approach should give a clearer definition of the trait complex in question (fertility) and that it therefore should have a higher correlation with fertility, than does overall productive life. This might also lead to a higher heritability of the longevity trait.

The objective of this study was to estimate genetic parameters and breeding values of sires for length of functional and fertility-determined productive life in Swedish dairy cattle, and to estimate genetic correlation between these traits and a commonly used fertility trait.

## 2. Material and methods

The initial data set consisted of 729 135 records of Swedish Red and White (SRB) cows with first calving from 1988 through 1996. Further editing was done as follows. Records from cows were deleted if they had: missing sire identification; missing or

erroneous herd identification; age at first calving outside 18-42 months; or incorrect calving dates. We also deleted records of cows changing herds and records from herds with less than 5 uncensored observations for productive life (censoring is described later). Only records from sires born 1983 or later and with >50 daughters (young sires) or from sires born before 1983 with >1000 daughters (proven bulls) were kept. This editing resulted in 534 016 cows.

These data were from 1211 young bulls and 55 proven bulls. When relationships were included among the random bulls, there were 1525 bulls included in the analysis.

The information consisted of calving dates and exit codes and dates, from the Swedish AI scheme and the official milk recording system, which contains 80-85% of all cows. The cows enter the system either at first service as a heifer or at first calving.

## 2. 1. Survival analysis

Longevity was measured as time from first calving to culling or end of data collection. Two different definitions of length of productive life were considered, their definition depending on when records were considered to be censored or not. For the first definition, **functional productive life (PL)**, records were considered to be censored if cows were alive at the end of data collection (96-12-31) or if the herd had disappeared from the recording system. In the latter case, if a herd did not have any calvings in a series of years, including 1996, then all exit dates from the year preceding the first missing year and onwards, were considered to be censored observations. After these rules were applied, about 36% of the records were (right-)censored.

In addition to *PL* another longevity trait was defined, **fertility-determined length of productive life (FPL)**. In addition to the censoring rules for *PL*, records were considered to be censored if the exit codes were not related to fertility problems. Expressed differently, we only considered the cow to be dead if it was culled for reproductive reasons, otherwise it was censored. This resulted in about 78% right-censored records.

The model used was:

$$h(t; \mathbf{z}) = h_0(t) \exp\{\mathbf{z}'(t) \boldsymbol{\beta}\} \quad [1]$$

where:

$h_0(t)$  is the baseline hazard,  $\lambda \rho(\lambda t) \rho^{-1}$ , assumed to follow a Weibull distribution with

parameters  $\lambda$  and  $\rho$ , and  $t$  is time in days from first calving,

$\mathbf{z}'(t) \boldsymbol{\beta}$  contains the (possibly time-dependent) covariates affecting the hazard, described in more detail below.

The effects included in the model were as follows:

*hys* random time-dependent effect of herd-year-season class; where seasons were from January-July and August-December;

*ys* fixed time-dependent effect of year-season, same season classes as above;

*p\*s* fixed time-dependent effect of parity by stage of lactation (parities were from 1 to 6<sup>+</sup>, stages changed at 0, 60, 180, 270, and 365 days of lactation);

*peak* a fixed time-dependent effect of the cow's peak test-day yield as deviation from her herdmates in that herd-year. Normalized deviations were calculated using the herd-year mean and the overall phenotypic standard deviation of test-day peak yield within herds (averaged over all herds). The calculations were done separately for first vs. later lactations. These normalized deviations were used to classify the cow in one of 13 classes, the cut-off points chosen such that the classes were expected to contain (starting from lowest yields) 4 x 2.5%, 1 x 5%, 7 x 10% and 1 x 15% of the observations. Before calculations, peak yields were adjusted for day in lactation up to day 60, based on a 4<sup>th</sup> degree polynomial estimated in the same data. Cows with missing information on peak yield were set to average herd-year production (class 9, deviation 0);

*age* a fixed time-independent effect of age at first calving, classified as 18-20, 21, 22, ..., 39, 40-42 months (21 classes);

*sire<sub>f</sub>* a fixed time-independent effect of the fixed sire (if a daughter from a proven bull); and

*sire<sub>r</sub>* a random time-independent effect of the random sire (if a daughter from a young bull).

The *hys*-effect was assumed to follow a gamma distribution with the parameter  $\gamma$  and the sire effect a normal distribution with mean zero and variance  $A\sigma_s^2$ , where  $\sigma_s^2$  is the variance among sires. Sire of

sire and maternal grandsire of sire were included in the relationship matrix A. The inclusion of fixed and random sires was done technically as follows. A dummy  $sire_f$  and a dummy  $sire_r$  were created. If a daughter was from a young bull, it had the dummy  $sire_f$  and its true father as  $sire_r$ . If a daughter was from a proven bull, it had the dummy  $sire_r$  and its true father as  $sire_f$  (Just Jensen, *pers. comm.*, 1998). The  $hys$ -effect was algebraically integrated out in the analysis program.

The analyses were done using the Survival Kit (Ducrocq & Sölkner, 1998). The analyses were done in a step-wise manner to improve convergence and to minimize computer requirements. The  $hys$  variance (or  $\gamma$ ) was estimated in a model without the sire effects included, and the estimate was then used as a fixed value in model [1]. Weibull parameters  $\rho$  and  $\lambda$  were estimated simultaneously with  $\sigma_s^2$ .

Heritability ( $h^2$ ) was estimated as  $4\sigma_s^2/(\pi^2/6+\sigma_s^2)$ , where  $\pi^2/6$  is the variance of the standard extreme value distribution (corresponding to residual variation in a normal distribution model). Variance of  $hys$  was the second moment (trigamma) of the log-gamma distribution (Lawless, 1982). For calculation of precision of sire breeding values, the heritability on the "original" scale ( $h_o^2$ ) was calculated by dividing  $h^2$  by  $\{\exp(-v/\rho)\}^2$ , where  $v$  is Euler's constant (0.5772) (Ducrocq, *pers. comm.*, 1998).

## 2. 2. Mixed linear model

The analysed fertility trait was interval in days between calving and last insemination in firstlactation (CLI). This analysis is described in more detail in Roth *et al.* (1998) and is based on the same data set as the current study. In the current analysis the predicted sire breeding values and the number of daughters from Roth *et al.* (1998) were used. The following mixed linear sire model was used:

$$y_{ijklmn} = hy_i + y_j + m_k + a_l + S_m + e_{ijklmn}$$

where:

- $y_{ijklmn}$  is the observation for CLI
- $hy_i$  is the random effect of the  $i$ th herd-year,  $IND(0, \sigma_{hy}^2)$ ,
- $y_j$  is the fixed effect of the  $j$ th year of calving,
- $m_k$  is the fixed effect of the  $k$ th month of calving,
- $a_l$  is the fixed effect of the  $l$ th age of calving,
- $S_m$  is the effect of the  $m$ th sire, random  $ND(0, A\sigma_s^2)$  if young bull and fixed if proven bull, and
- $e_{ijklmn}$  is the random residual for this observation,  $IND(0, \sigma_e^2)$

Sire of sire and maternal grandsire of sire were included in the relationship matrix A. (Co)variance components were estimated by use of restricted maximum likelihood (REML) in the DMUAI module of the DMU-program by Jensen & Madsen (1994).

Records in herd-year subclasses with less than 5 observations were discarded to ensure reasonable data structure for the contemporary groups and to avoid confounding with other effects. Sires born from 1983 to 1992 with at least 50 daughters were regarded as young bulls, whereas sires born before 1983 with more than 1000 daughters were considered as selected, proven sires. The proven bulls were treated as fixed effects, whereas young bulls were treated as random. Daughters of sires not fitting into these two categories were deleted.

## 2. 3. Genetic correlations

Genetic correlations between PL, FPL and CLI were approximated from calculated correlations between predicted breeding values for these traits according to Blanchard *et al.* (1983):

$$r_{gij} = r_{EBVij} [(\Sigma b_i)(\Sigma b_j)]^{0.5} / (\Sigma b_i b_j)$$

where  $b_i$  is the precision of the predicted breeding value of a sire for trait  $i$ , ( $r_{TT}^2$ ) calculated as

$$N/(N+(4-h_o^2)/h_o^2)$$

where  $N$  is the number of uncensored daughters of the sire and  $h_o^2$  is the heritability on the original scale, if  $i$  is PL or FPL, or where  $N$  is number of daughters and  $h_o^2$  is heritability, if  $i$  is CLI.

## 3. Results

The  $\gamma$  parameter was estimated to 1.10 and 1.26, for PL and FPL, respectively. These values correspond to  $hys$ -variances of 1.433 and 1.187, respectively.

The Weibull parameter  $\rho$  was estimated to 0.79 and 0.32 for PL and FPL, respectively. Because the transformation of  $h^2$  to  $h_o^2$  was very sensitive to values of  $\rho$  below unity, a value of unity was used. When the analysis for PL was run with  $\rho$  set to unity, the objective function was changed by 0.0289%.

The effect of parity-stage of lactation together with the baseline hazard is shown in Figure 1. The effects of stage of lactation (separately for each parity), year-season, peak test-day yield deviation, and age at first calving are given in Figures 2-5.

Sire variances for PL and FPL were estimated to 0.02904 and 0.02561, respectively. This corresponds to heritabilities of 6.94% and 6.13%, respectively.

Genetic correlations between CLI on one hand and PL or FPL on the other, were estimated to 0.47 and 0.55, respectively.

#### 4. Discussion

The relative risk of culling increased consistently throughout the lactation (Figure 1) for both longevity traits. The small decrease in risk during a given stage is due to the estimated value of  $\rho$  being below unity. It is not clear to us whether such a  $\rho$  is reasonable or not. Most other studies (e.g. Ducrocq, 1994) have found  $\rho$  above unity and have interpreted that as that risk should naturally increase over life. It seems that in our material, the parity\*stage of lactation effect in itself takes account of that increased risk, even to the extent of making  $\rho$  come out below unity.

Risk also increased with increasing parity (Figure 2). For a given stage, risk was always higher in a later parity. The difference between parities were larger at the later stages of lactation. Even considering the time-dependent effect of the baseline hazard, this was true (Figure 1).

There is a difference in risk patterns over the lactation between PL and FPL. For FPL there is hardly any risk of culling at the beginning of the lactations (up to 180 days). This is reasonable, as during early lactation cows are not expected to be pregnant and thus it is highly unlikely that they will be culled for that same condition. The fact that these estimates are so low is in a way a validation of the exit codes: if we had found high risk in the beginning of the lactations, some cows marked as culled for fertility would most likely been culled for other reasons.

The hazard for FPL is also lower than for PL, at any given time, which seems reasonable, and the difference becomes larger with increasing parity.

The estimates of fixed year-season effects showed a decreasing trend over the time studied, with the exception of the first years (Figure 3). The pattern of the estimates were the same for both traits. There was a clearly higher risk of culling during the season August to December compared with the rest of the year. This may be related to the beginning of the indoor housing season.

The reasons for the decreasing trend over time is not quite obvious to us. However, the interpretation might be the other way around, that culling is rather

constant during the last 5-6 years but was higher for some reason during the early years, possibly due to political decisions at the time. The quota system started in 1995 and has therefore hardly affected our data.

The estimated effects of risk of culling for cows in different production classes showed a very distinct pattern for PL (Figure 4). Only the very lowest class had a distinctly higher risk of culling, then the risk decreased substantially and was more or less independent of production. There was a clear benefit of dividing the classes "unevenly" with respect to the deviation, to get a finer distinction between production classes in the lower part of the spectrum. The trait used, peak yield, which is description of early production, is probably better to use than total lactation production. The decision to cull on production (and not to inseminate) is probably made at an early stage, before total production is known. If the cow is kept for a whole lactation, the actual 305-day production may become higher because she is not pregnant and the deviation from her herdmates may become smaller than that which actually contributed to her being culled.

The pattern is a bit more complex for FPL. There is a general downward trend, but the next to lowest production class has the highest risk of culling. The very lowest producers do not run the highest risk of being culled for fertility (then they are culled for production instead). However, if they are low producers they also have higher risk of being culled for fertility.

Increased age at first calving contributed to increased risk of culling for PL (Figure 5). However, the increase was lower around the period of normal ages of calving (26-33 months). For FPL, the increase in risk was lower and more constant over the whole period.

The genetic variation and heritability did not decrease very much when changing from PL to FPL. Perhaps one reason is that there is a clearer description of the reasons for culling for FPL than for PL.

The approximated genetic correlation estimate between FPL and the fertility trait CLI was somewhat stronger than that between PL and CLI. This was, of course, in accordance with our expectations, that a fertility-determined longevity measure should be more correlated to fertility than an overall measure of longevity. It was not possible to test whether the difference in estimated correlations was significant or not, but judging from experience of standard errors on genetic correlations, most likely it was not. The procedure of calculation genetic

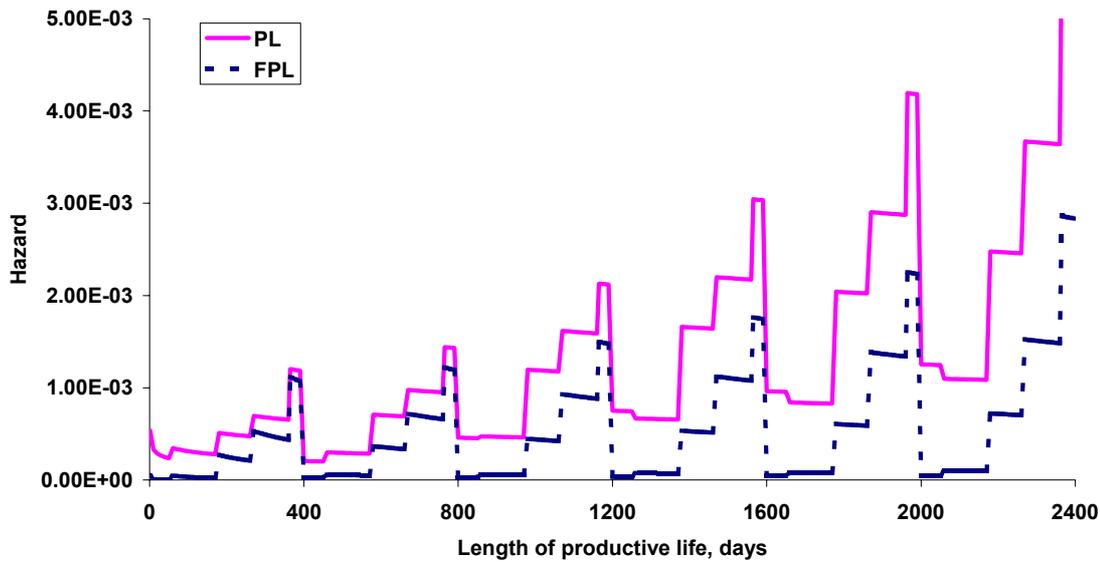
correlations from correlations between predicted breeding values is not the ideal one, and better methods will hopefully be developed and some attempts are discussed in other papers at the workshop.

The large proportion of censored records did not make computational efforts much more demanding. However, we have not estimated precision of the estimated parameters, but we would expect precision to be lower for FPL parameters, because we have fewer number of uncensored records available. The average  $r^2_{TI}$  was 0.76 for PL and 0.52 for FPL.

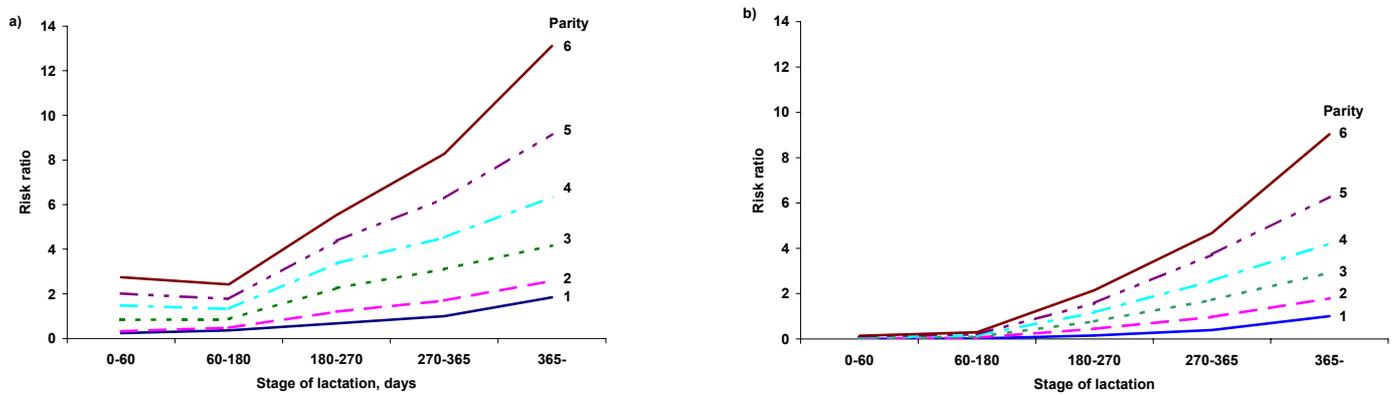
These preliminary results show that it seems to be possible to define a fertility-determined longevity measure which can be used for selection, if ordinary fertility measures are unavailable. A similar approach could be tried with e.g., udder-health-determined longevity.

## References

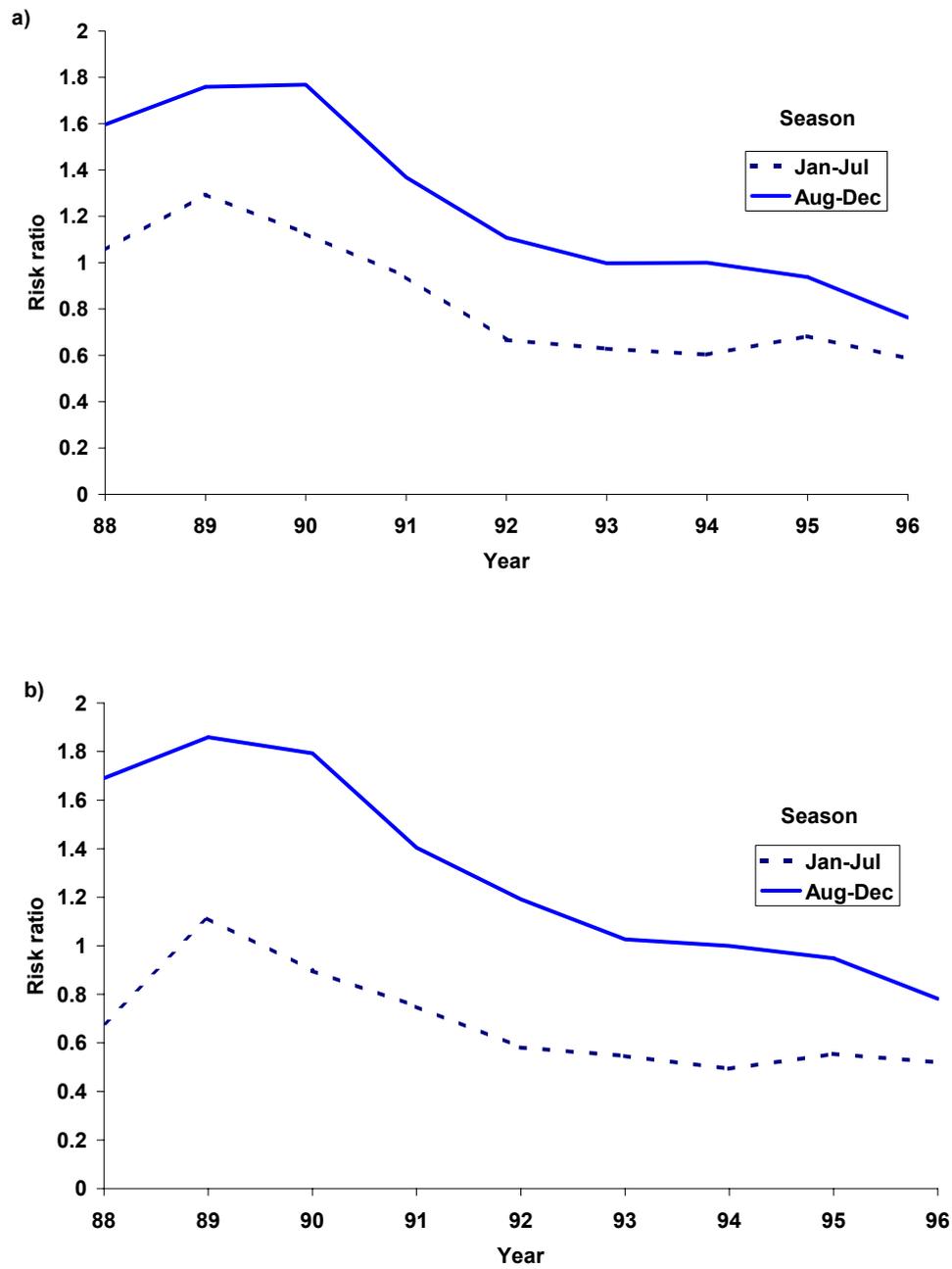
- Blanchard, P.J., Everett, R.W., and Searle, S.R. 1983. Estimation of genetic trends and correlations for Jersey cattle. *J. Dairy Sci.* 66:1947-1954.
- Ducrocq, V.P. 1994. Statistical analysis of length of productive life for dairy cows of the Normande breed. *J. Dairy Sci.* 77:855-866.
- Ducrocq, V.P. & Sölkner, J. 1998. "The Survival Kit V3.0"- a package for large analyses of survival data. Proceedings of 6<sup>th</sup> World Congr. Genet. Appl. Livestock Prod., 27:447-450.
- Jensen, J. & Madsen, P. 1994. A user's guide to DMU - A package for analysing multivariate mixed models. National Institute of Animal Science Research Centre, Foulum, Denmark.
- Lawless, J. 1982. Statistical methods for survival data analysis. John Wiley & Sons, New York, NY.
- Roth, A., Strandberg, E., Berglund, B., Emanuelson, U. & Philipsson, J. 1998. Genetic correlations among female fertility traits and milk production in different parities in Swedish dairy cattle. Paper pres. at BSAS/GIFT/BCVA workshop on Metabolic stress in dairy cows, Edinburgh, October 28-30.
- Strandberg, E. 1997. Breeding strategies to improve longevity. Paper pres. 48th ann. meet. EAAP.



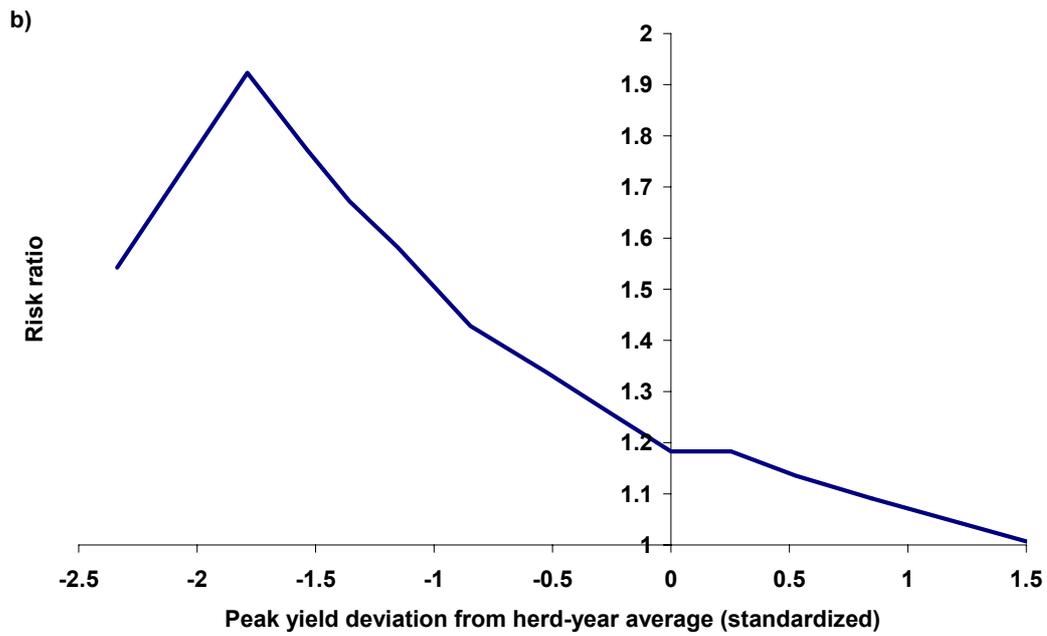
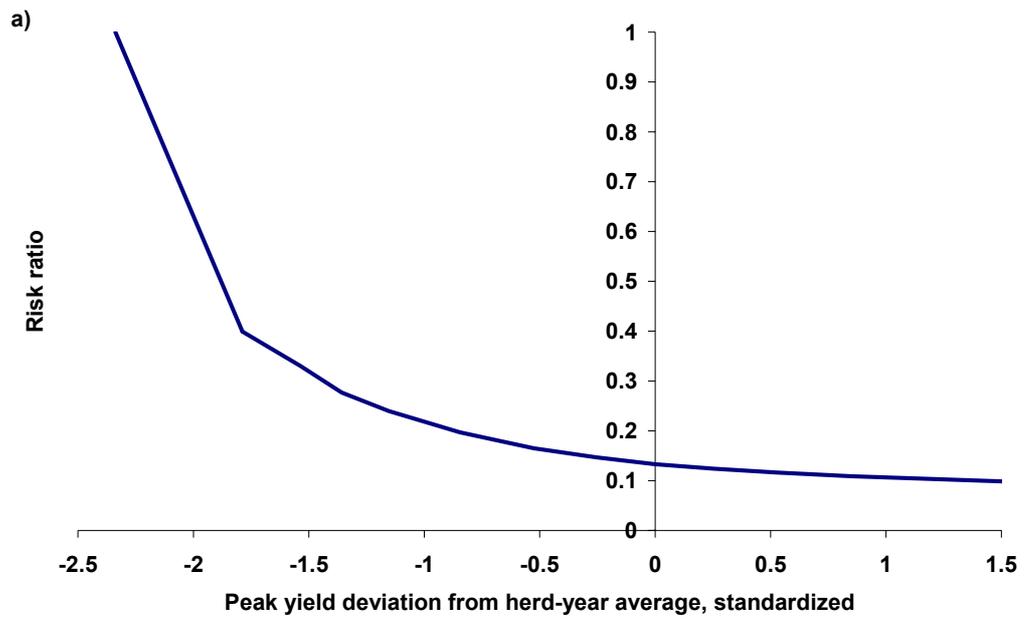
**Figure 1.** Estimated hazard for a cow with 400 day calving interval, average peak yield who was 32 months at first calving in the fall of 1994. Hazard shown both for productive life (PL) and for fertility-determined productive life (FPL).



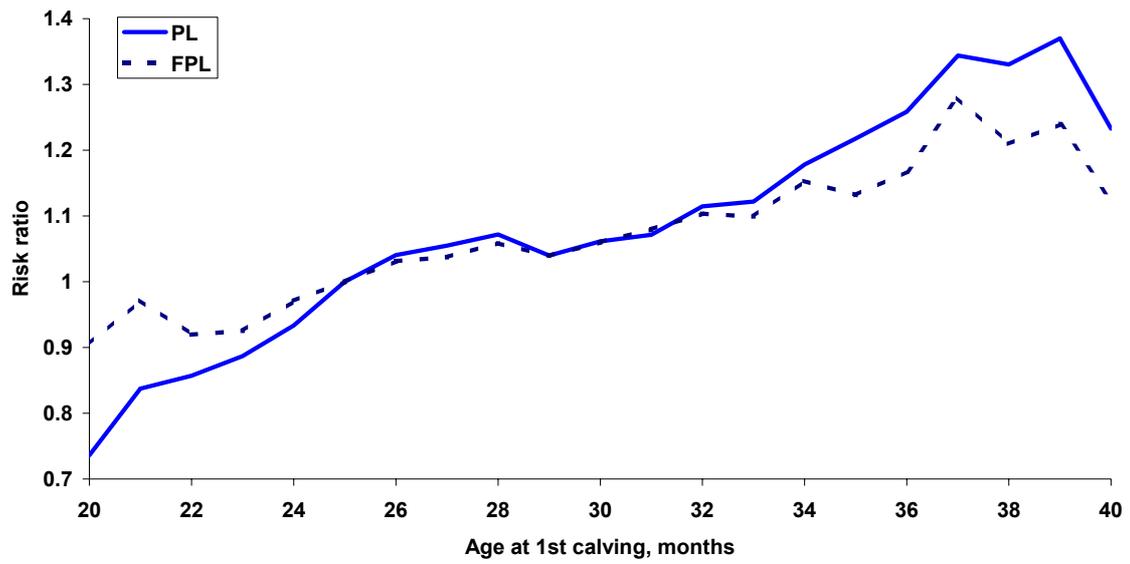
**Figure 2.** Estimates of the relative risk of culling for various stages of lactation within parities (1-6+): a) for productive life (PL) and b) for fertility-determined productive life (FPL).



**Figure 3.** Estimates of the relative risk of culling for various years and seasons within year (January-July and August-December): a) for productive life (PL) and b) for fertility-determined productive life (FPL).



**Figure 4.** Estimates of the relative risk of culling for various classes of peak yield deviations from herd-year average: a) for productive life (PL) and b) for fertility-determined productive life (FPL).



**Figure 5.** Estimates of the relative risk of culling for various classes age at first calving (months), for productive life (PL) and for fertility-determined productive life (FPL).