E-19
Breeding value for longevity (LON)

▪ Introduction

The lifespan of a dairy cow indicates how long a cow has been able to prevent being replaced due to a shortcoming. In other words, how satisfied the farmer is with her. A farmer considers many traits when determining his culling policy. Roughly, these traits can be divided into production, fertility, health and workability. As cows are kept for the purpose of producing milk, replacement based on an unsatisfactory production is called voluntary culling, and replacement based on other traits is called involuntary culling. In practice, the reason that a cow is culled is seldom due to only one of these two. A cow that is not in calf following insemination and with a low milk production is more likely to be culled than another cow that is also not yet in calf, but that has a higher milk production.

The actual lifespan of a dairy cow depends upon the dairy farmer's decision to cull the cow. Assuming that a constant number of cows are kept in a herd throughout the years, one can easily calculate the actual lifespan on an operation using '1/replacement rate'. The replacement rate is the number of heifers divided by the total number of dairy cows. In this way, the lifespan is expressed in terms of the number of lactations. The same applies on a national level. According to the Annual Statistics, the replacement rate has been just above 30% for years now. This means that the average realised productive lifespan was about 3.4 lactations since 2006. The average number of milking days of all of the lactations, including incomplete records, is 320 days. Figure 1 presents information concerning the lifespan of the cows included in the milk recording system that have been culled from 1990 up to and including 2016. The productive lifespan, this is the period between the first calving date and last milk test date, is expressed in number of days.
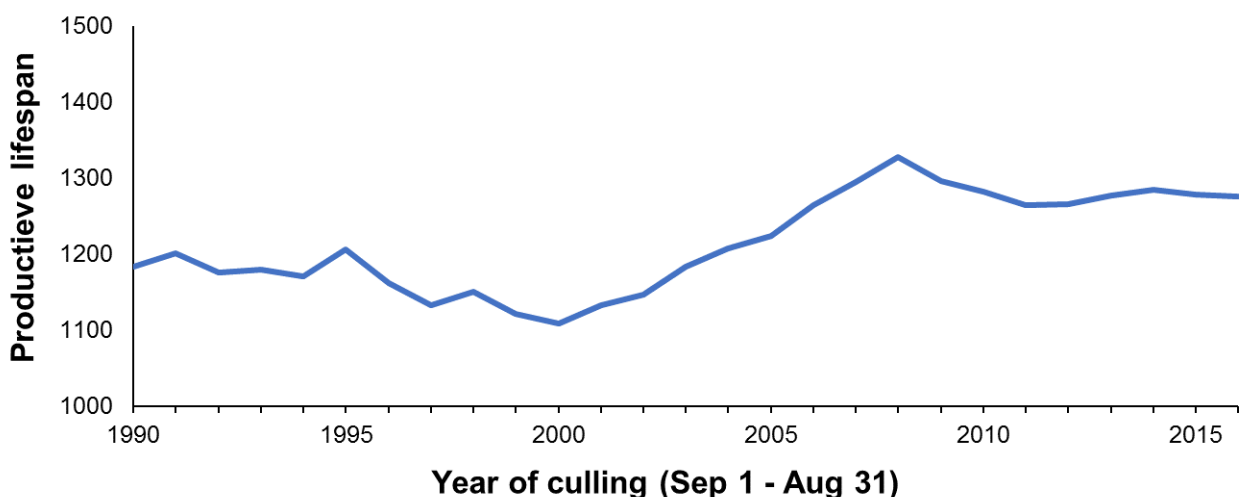

Figure 1. The productive lifespan of herdbook registered cows included in the milk recording system that were culled in the period 1990 – 2016.

Because the farmer determines the lifespan of a dairy cow, partly determined by legislation, it is difficult to visualise the results of selection for longevity. As it is, the dairy farmer continues to select cows in his herd, which is mainly determined by the percentage of young heifers, compared to his dairy herd, which he raises each year. Because of this, selection based on longevity will not give

longer realised herd life. Selection on lifespan only gives a longer herd life in case the dairy farmer really decides to raise less young stock and starts milking less new heifers (lower replacement rate). With a constant replacement rate and a genetic improvement of the herd, involuntary culling will be replaced by voluntary culling (on the population level).

The breeding value for longevity is a figure that indicates how satisfactory the daughters of a certain bull function. It is a measure to rank bulls for the ability of their daughters to resist culling. Until December 2017 breeding values for longevity were estimated based on survival analysis using the "Survival Kit" (Ducrocq and Sölkner, 1998). Since April 2018 breeding values are estimated with a random regression model to get more stable breeding values for bulls (Van Pelt, 2017). From April 2019 onwards the information of cows in Flanders is also included in the genetic evaluation for longevity.

▪ Direct and indirect information

The milk recording data provides information concerning when a cow is calving for the first time and when her last test milking occurred. The productive lifespan can then be calculated as the difference between these two dates. This information is the *direct* information that is analysed in the breeding value estimation. As the heritability of longevity is low (about 12%) and it may take quite some time before enough direct information is available concerning the lifespan of the daughters of a bull, information on correlated (predictive) traits is also used in the case of young bulls. This information is the *indirect* information. The final breeding value for longevity consists of two components: one based on the direct breeding value for longevity (direct information), and one based on the breeding values for other traits (indirect information). The amount of information from both sources may vary per bull. For this reason, direct and indirect information are weighted per bull.

When the breeding value of a bull is estimated for the first time, the indirect information results in an increase in the reliability of about 10%. As a bull grows older, more direct information from his daughter about lifespan becomes available and so the indirect information becomes less significant.

▪ Direct breeding value estimation

a. Data
The direct breeding value estimation is based on the direct information concerning the lifespan of the herd book daughters of a bull. The productive lifespan is calculated from the data from the milk recording system: the number of days between the first calving and the last test milking. If a cow is transferred to a different herd that participates in the milk recording system, then this has no effect whatsoever on the calculation of her lifespan. If she is transferred to a herd that is not participating in the milk recording system, then her definite lifespan will not be known, but her lifespan up until that moment, which is a minimum value of the definite lifespan, will be known.

The breeding value estimation uses data that meet the following requirements:
- The cow has to be herdbook registered;
- The production goal of the cow is 'milk';
- The cow has a test milking or her first calving after January 1, 1988 when residing in the Netherlands or January 1, 2006 when residing in Flanders;

- Only official lactations are used;
- The cow must be housed in a herd with milk recording;
- The cow has known locations;
- Age at first calving is at least 640 days;
- Data of a cow is used after a waiting period of 120 days after first calving;
- Culled heifers without a milk testing, mostly culled before the first milk test, present in a herd with milk recording, are included;
- Herd-year-months need to have a survival rate of at least 70%;
- Herd-year-months with 5 or more culled animals need to have a survival rate that is higher than the mean survival rate of the past 12 months minus three times the standard deviation of survival of the past 12 months.

Culled heifers without a milk testing are included in the breeding value estimation, because this is the most expensive culling. The farmer had all rearing costs, but hardly any milk revenues. The two requirements at herd-year-month level are set to use only records from 'normal' culling. For example, non-normal culling is a herd that stopped milk production and sold all animals at once or in batches, or the entire herd was culled at once because of an outbreak of a disease, or herds that shrinked because of legislation. Such extreme culling at the herd level does not give a representative comparison of the culled animals with their herd mates, because under normal conditions a smaller part of these cows would have been culled.

b. random regression
The direct lifespan data is analysed using random regression. This is the same methodology that is used for the breeding value estimation for milk production traits. The methodology of the random regression is characterised by the following:
1. The animals of which the definite lifespan is not (yet) known are also included in the analysis;
2. It allows for an exact adjustment for interfering effects, and these effects are modelled time dependent;
3. The survival per month (or the chance of culling) can be modelled per parity and within parity for lactation stage.
The random regression model does not examine the lifespan, but rather the survival per month. Lifespan and survival per month are closely connected.

As long as cows are still alive, their actual lifespan remains unknown. The data of such cows cannot be included in methodologies such as those used in the breeding value estimation for other traits. If the data were to be included, it would mean that the breeding value for the lifespan would be continually underestimated, especially in the case of young bulls. After all, a young bull has young daughters by definition. If the method only analyses the data of the daughters that have already been culled, then this data will always concern daughters that were replaced at a young age. This will mean a low breeding value, independent of how many daughters are still alive.
The random regression model also considers the data of the daughters that have not (yet) been replaced. This is possible because the analysis considers survival per month after first calving, rather than the lifespan itself. A cow that is still alive has not been replaced in the competition with her herd mates and this provides information on her lifespan compared to that of her herd mates. In this way, one can obtain a useful breeding value for longevity for a young bull as well.

The second advantage of the random regression model is the precise modelling. It is assumed in traditional methodologies that the influence of any interfering effect is continuous. In the random

regression model, it is possible to incorporate a time component (the effect is time-dependent). An example: previously the effect of the herd-year-season-parity (HYSP) was incorporated in the breeding value estimation for milk production with a lactation model. Depending upon the operation and the date of calving, a HYS-effect is defined for each parity; one that applies to that particular parity, that is to say, during the entire lactation period. Where the lifespan is concerned, it is obvious that a single HYS-effect for the entire life of a cow may be too long. For example, the quota system will cause the culling policy (and therefore the lifespan) in the beginning of the calendar year to differ from the other seasons. Random regression offers the opportunity to alter the HYS per season. The HYS-effect also changes if a cow is transferred to another herd. The survival per month due to the HYS-effect varies with each season. Even if the cow is transferred from one herd to the other, the HYS-effect on survival per month will change. It is also possible to model other effects with a time component in the same way, so that the adjustment for these interfering influences is very precise.

The third advantage of random regression is that the survival per month can be modelled per parity and within parity for lactation stage. The survival at day 250 of a lactation is higher than the survival at day 350 of the lactation. When a cow is culled at day 250 of the lactation, this has a bigger impact on the breeding value of her sire than when the cow is culled at day 350 of the lactation.

The fourth advantage of random regression is that an animal model can be used, resulting in breeding values for both bulls and cows. With the survival analysis, the method used until December 2017, a limited number of animal effects could be estimated because of the complexity of the calculation method and the excessive computer capacity needed, and only sire and maternal grandsire were used to estimate the genetic effects. In the animal model, all family relationships are taken into account, meaning that the dam is included as well, instead of the maternal grandsire. With the animal model, family relations within a herd are handled better, allowing to distinguish between cow families even when the pedigree for sire and maternal grandsire are the same.

The fifth advantage of random regression is that all data is used for the calculation of the reliability. That means that both animals still alive and culled animals provide information for the reliability of a bull. For the survival analysis only culled animals provided information for the reliability of a bull. For a bull with only daughters in first lactation the reliability is higher with random regression than with survival analysis.

c. statistical model
The model for the direct breeding value estimation for longevity, where survival per month is analysed, is a linear random regression model, and is as follows:

$$Y_{ijklmno} = HYS\_LS_i + YSAM\_LS_j + HSC_k + het_l + rec_m + animal_n + rest_{ijklmno}$$

in which:

$Y_{ijklmno}$: Observation for survival in month *o* after first calving; month 1 to 72;

HYS_LS_i : Herd-year-season x lactation-stage *i*; year-season of first calving, lactation split in 1, 2, 3+, stage of lactation split in month 1-2, 3-9, 10+ and dry period;

$YSAM\_LS_j$: Year-season x age of first calving x within-herd production level x lactation-stage *j*; year-season of first calving, age of first calving in months 21, 22, …, 34, 35+, within-herd production level is defined per 3 years and is divided in 5

		classes of 20% each for predicted or realised age-corrected 305-day yield of kg fat and protein;
HSC_k	:	Herd size change k; HSC is calculated by comparing the number of cows present in a herd in a year with the number of cows in the same herd one year later. 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 (more than 90% shrinkage).
het_l	:	Heterosis l of animal n;
rec_m	:	Recombination m of animal n;
$animal_n$	:	Additive genetic effect (or breeding value) of animal n; the additive genetic effect is estimated with a random regression function and describes the effect on survival for every month between month 1 and 72. This results in a breeding value for every month;
$rest_{ijklmno}$	:	Residual of $Y_{ijklmno}$, anything that is not explained by the model.

The change in the size of a herd is incorporated because a herd that is expanding will adopt a replacement policy that is totally different from that of a herd that is cutting back. Age of first calving is shown to have an effect on the lifespan of dairy cows. Cows that are relatively old at the time of the first calving have a shorter lifespan than cows that calved relatively young.

The additive genetic effect is modelled with a random regression function with a fifth order Legendre polynomial. The regression function is mathematical function describing the additive genetic effect on survival per month for every month between month 1 and 72 after first calving. These effects do not have to be the same for every month. Therefore, it is possible that animals are genetically good in first lactation (higher survival), but bad in later life (lower survival). Or the other way around, an animal has genetically a higher chance of surviving third lactation than surviving first lactation. The random regression model yields 72 breeding values, one for every month between month 1 and 72. The breeding values per month are extrapolated up to 180 months to derive an overall breeding value for the total productive lifespan. The overall breeding value over 180 months is derived as follows: the breeding value of a certain month is added to the mean survival rate of that specific month, then this value is multiplied with the value of the previous month to get the cumulative survival rate. This results in a survival curve of a bull, and the overall breeding value is the difference between the surface of this curve and surface of the mean survival curve.

d. Parameters
The heritability of direct longevity is 0.12 and the genetic standard deviation is 243 days.

▪ Indirect breeding value estimation

a. productive longevity
In the direct breeding value estimation breeding values for functional longevity are estimated, because survival per month is adjusted for within-herd production level in the statistical model. Bull owners and farmers are used to using the breeding value for productive longevity. With a selection index productive longevity is derived from functional longevity and the production traits kg milk, kg

fat, and kg protein. Table 1 shows the heritabilities, genetic correlations and genetic standard deviations that are used in the selection index to derive productive longevity.

Table 1. Heritabilities (bold on diagonal), genetic correlations and genetic standard deviations for functional survival, kg milk, kg fat and kg protein to derive productive longevity.

Trait	Func.lon	Kg milk	Kg fat	Kg protein	Prod.lon	Gen.sd[1]
Func.lon	**0.12**					243
Kg milk	0.02	**0.51**				775
Kg fat	-0.05	0.39	**0.52**			32
Kg protein	-0.02	0.87	0.61	**0.44**		22
Prod.lon	0.87	0.37	0.34	0.44	**0.14**	258

[1]genetic standard deviation for functional and productive longevity is expressed in days

b. Predictors

Essentially, the breeding value for longevity is estimated using data on the lifespan of the daughters of a certain bull. However, the reliability of the breeding value estimation of a bull depends upon the heritability that applies for the trait for which a breeding value is being estimated, and on the amount of information that is available. The heritability for longevity is not very high: 0.12. For comparison: the heritability for milk production is 0.51. This means that in order to obtain a certain degree of reliability in the breeding value estimation for longevity, one will require more information than in the case of milk production. Particularly when young bulls are concerned, little direct information is available. Therefore, indirect information is used next to the direct information for longevity.

For the predictor traits breeding values are estimated, as well as parent averages. The breeding values minus the parent average is used as predictor. The advantage of this approach is that information is not counted double. After all, the information of the predictors from the parents is already included in the parent average for longevity for the bull. This method is used for young bulls and cows to increase the reliability of the breeding value for longevity. In the case of older proven bulls, the indirect information does not contribute anything new to the breeding value for lifespan.

There are a number of traits that show a correlation with longevity. Here, the most important are those that, combined, provide the best prediction. The traits of which the breeding values, in addition to the direct information concerning longevity, are incorporated in the estimation of the breeding value longevity are subclinical mastitis, claw health and locomotion. Table 1 shows the heritabilities, genetic correlations and genetic standard deviations that are used in the selection index to add information from predictors to get a better estimate for productive longevity.

Tabel 2. Heritabilities (bold on diagonal), genetic correlations and genetic standard deviations to add information from predictors to productive longevity.

Kenmerk	Subcl.mast	Claw health	Locomotion	Prod.lon	Gen.sd[1]
Subcl.mast	**0.056**				4.5
Claw health	0.09	**0.175**			4.5
Locomotion	0.05	0.77	**0.14**		4.5
Prod.lon	0.62	0.52	0.27	**0.14**	258

[1]genetic standard deviation for functional and productive longevity is expressed in days, other traits are expressed on a relative scale with a genetic standard deviation of 4.5 points

▪ Publication

The breeding value for longevity is presented as a complete breeding value, expressed in days. The standard deviation of the breeding value longevity is 258 days. The breeding value for longevity is published for an AI sire if the reliability is at least 25%.

▪ Base definitions

Breeding values for longevity are published based on the 2015-base. Cows born in 2010 determine the base of 2015. There are four different bases: Milk goal Black, Milk goal Red, Dual purpose and Belgian Blue. The definitions of these bases are as follows:

Milk goal Black (Z)
Herdbook-registered cows born in 2010 with at least 87.5% HF-blood and up to 12.5% FH-blood and hair colour black pied, with at least one observation in the genetic evaluation.

Milk goal Red (R)
Herdbook-registered cows born in 2010 with at least 87.5% HF-blood and up to 12.5% MRY-blood and hair colour red pied, with at least one observation in the genetic evaluation.

Dual purpose (D)
Herdbook-registered cows born in 2010 with at least 75% MRIJ-blood and 25% or less HF blood, with at least one observation in the genetic evaluation.

Belgian Blue (B)
The Belgian Blue base is determined by the animals that determine the Dual purpose base.

Every 5 years, in a year divisible by 5, the reference year for the base is moved 5 years. The base differences are shown in Table 3.

Table 3. Base differences for longevity.

	$Z \rightarrow R$	$Z \rightarrow D$	$R \rightarrow D$
Longevity	50	280	230

▪ Literature

Ducrocq, V.P. en J. Sölkner, 1998. "The Survival Kit" – a package for large analyses of survival data. Proc. 6th World Congress on Genetics Applied to Livestock Production, vol. 27, pag. 447.

Van Pelt, M.L., De Jong, G., Veerkamp, R.F., 2017. Improving the genetic evaluation for longevity in the Netherlands. Interbull bulletin 51: 33-37, Tallinn, Estonia, August 25-28.

Van Pelt, M.L., 2017. Genetic improvement of longevity in dairy cows. PhD-thesis, Wageningen University.