

Article

Estimation of Dairy Cow Survival in the First Three Lactations for Different Culling Reasons Using the Kaplan–Meier Method

Wilhelm Grzesiak ¹, Krzysztof Adamczyk ², Daniel Zaborski ^{1,*} and Jerzy Wójcik ¹

¹ Department of Ruminants Science, West Pomeranian University of Technology in Szczecin, Klemensa Janickiego 29, 71-270 Szczecin, Poland; wilhelm.grzesiak@zut.edu.pl (W.G.); jerzy.wojcik@zut.edu.pl (J.W.)

² Department of Cattle Breeding, Institute of Animal Sciences, University of Agriculture in Krakow, Mickiewicza 24/28, 30-059 Kraków, Poland; rzadamcz@cyfronet.pl

* Correspondence: daniel.zaborski@zut.edu.pl

Simple Summary: From a breeding and production point of view, the length and quality of life of dairy cows are directly determined (more or less) by voluntary decisions made by the breeders and technical staff in human–animal–environment relationships. In this case, economic conditions are the key roles that greatly complicate the decision processes, especially when it concerns the whole herd (not solely single animals). On the other hand, increasing social pressures on the continuous improvements of animal welfare and pro-environmental agricultural practices, including high-producing dairy cows, can be seen. Therefore, the aim of the present study was to analyze survival curves for cows culled for different reasons over three successive lactations and to determine the effects of various factors on cow survival. The main culling categories were reproductive disorders—40%, udder diseases—13 to 15%, and locomotor system diseases—above 10%. The survival curves for cows from individual culling categories had similar shapes. The greatest influences on the relative culling risks were exerted by: age at first calving, lactation length, calving interval, production subindex, breeding value for longevity, temperament, and average daily milk yield. A more accurate method of determining culling reasons would be required.



Citation: Grzesiak, W.; Adamczyk, K.; Zaborski, D.; Wójcik, J. Estimation of Dairy Cow Survival in the First Three Lactations for Different Culling Reasons Using the Kaplan–Meier Method. *Animals* **2022**, *12*, 1942. <https://doi.org/10.3390/ani12151942>

Academic Editor: Victor E. Cabrera

Received: 3 July 2022

Accepted: 27 July 2022

Published: 30 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: The aims of the study were: (i) to compare survival curves for cows culled for different reasons over three successive lactations using the Kaplan–Meier estimator; (ii) to determine the effects of breeding documentation parameters on cow survival; (iii) to investigate the similarity between culling categories. The survival times for a subset of 347,939 Holstein-Friesian cows culled between 2017 and 2018 in Poland were expressed in months from calving to culling or the end of lactation. The survival tables were constructed for each culling category and lactation number. The survival curves were also compared. The main culling categories were reproductive disorders—40%, udder diseases—13 to 15%, and locomotor system diseases—above 10%. The survival curves for cows from individual culling categories had similar shapes. A low probability of survival curves for metabolic and digestive system diseases and respiratory diseases was observed in each of the three lactations. The contagious disease category was almost non-existent in the first lactation. The greatest influence on the relative culling risk was exerted by age at first calving, lactation length, calving interval, production subindex, breeding value for longevity, temperament, and average daily milk yield. A more accurate method of determining culling reasons would be required.

Keywords: cluster analysis; Cox proportional hazards model; culling; dairy cow; Kaplan–Meier curve; survival analysis; survival table

1. Introduction

From a breeding and production point of view, the length and quality of life of dairy cows are directly determined (more or less) by voluntary decisions made by the breeders and technical staff in human–animal–environment relationships [1]. In this case, the

key roles involve macroeconomic (e.g., changes in the global relationships between milk prices and animal feeding costs as well as breeding animal prices) and microeconomic (e.g., veterinary service expenditures) conditions, which greatly complicate the decision processes, especially when it concerns the whole herd (not solely single animals) [2]. On the other hand, it can be seen that social pressure on the continuous improvements of animal welfare (promoting grazing management systems and free-stall housing instead of tie-stall housing, social recommendations, and expectations about the humane treatment of animals at each stage of their handling, as well as respecting the five freedoms as minimum obligations to animals) and pro-environmental agricultural practices (especially the reduction of carbon footprints) involving high-producing dairy cows increases [3–7]. These factors result in the need for an increasingly detailed analysis of the length and quality of a cow's life, first utilizing routinely collected breeding data [8]. Among them, special attention should be paid to data on the longevity and survivability of cows and their culling reasons [9,10]. Culling decisions constitute a major challenge to dairy owners and managers. Culling has a huge impact on breeding progress but it also generates costs associated with replacement heifers that account for a substantial part of the dairy budget. The importance of culling decisions is reflected in the fact that owners of particularly large herds, who are not normally involved in the management decisions of individual cows, often directly participate in the decision-making process for selecting animals to cull [11,12]. However, the aim of the breeding work is not to completely eliminate culling, but to minimize the rate of cow culling unplanned by breeders (involuntary culling) at an increased level of voluntary culling [13,14]. Therefore, the knowledge of culling reasons and their dynamics (especially for the first three lactations of Holstein-Friesian cows) exert the greatest effects on breeding decisions as well as breeding and production results at the herd and population levels [15–19]. However, it should be emphasized that the outcome (animal survivability or longevity and culling) is not only associated with the single culling decision, but also with the life history of an animal, including the age of first calving and feeding intensity that affects the milk yield of cows [7,20]. A significant association between longevity and culling reasons with milk performance traits and the reproduction of cows has been indicated, taking into account management conditions at the same time (e.g., herd size, culling season). For instance, the need to consider factors, such as the percentage of Holstein-Friesian genes, the breeding values of cows and sires, age at first calving, the length of calving interval, and milk performance, has been confirmed in the evaluation of longevity and culling reasons in high-yielding animals [21–27]. Considering the above, the analysis of longevity and survivability of dairy cows in terms of breeding and production aspects has been a top research topic in recent years. Special attention has been paid to the selection of methods precisely determining the relationships among variables [28–31]. It appears that survival analysis is one of them. It is used for investigating the association between longevity, survivability, culling reasons, and routine herd data in terms of dairy cow culling.

The basis for survival analysis is the duration process, characterized by a certain relationship, i.e., the occurrence of two events (e.g., the initial and final ones, such as a cow's birth and her culling) [32–38]. The time interval between these events is a random variable [39,40]. Such an approach allows for the development of a survival model, which is the probability distribution of this variable, assuming positive values [34]. The Kaplan–Meier estimator is one of the most frequently used methods to analyze survival data and make comparisons between groups of individuals [36,41], such as cows culled for different reasons [14,42,43]. One advantage of the survival analysis, also known as the failure time or event time analysis, is the fact that the information from cows that are not culled by the time the data recording is complete can also be retained. Thus, records from culled (uncensored) and surviving (censored) cows can be treated jointly and included in the analysis, so that all the available information is properly used [17–19,37,38,44,45]. Cows still alive at the end of the study period cause problems when using other methods for the survival evaluation because their actual life spans are unknown. Therefore, the survival

analysis offers multiple advantages over traditional linear models in terms of better fitting the survival data and using time-dependent variables properly [14,19,42]. It is based on the so-called hazard rate, i.e., the probability of a cow being culled at time t , given that the cow has been alive prior to t [41]. The hazard rate is usually modeled as a product of a baseline hazard function (which relates to the natural aging process) and an exponential function of factors potentially affecting culling risk, such as herd–year–season, milk production level, or genetic causes (e.g., sire effect). The hazard rate can be calculated from both uncensored and censored records [17,38,44,45]. Historical data are usually used to construct survival models, although this method can also be the basis for prediction [46]. Cluster analysis, on the other hand, allows for the grouping of objects according to specific criteria, which in turn enables the creation of a useful classification system [47,48].

The main aims of the study were: (i) to describe and compare survival curves for cows culled for different reasons over three successive lactations using the Kaplan–Meier estimator; (ii) to determine the effects of breeding documentation parameters on the survival of cows, especially their influences on the relative risks in animals culled for various reasons; (iii) to investigate the similarities between individual culling categories based on survival curves and the cluster analysis.

2. Materials and Methods

A total of 347,939 milk-recorded Holstein-Friesian cows culled between 2017 and 2018 in Poland were included in the study. They were mostly kept under an outdoor/indoor system (i.e., grazed during the spring and summer and had access to paddocks throughout the year). A stratified sample of the original dataset was used for a detailed analysis. Table 1 presents the structure of cull cows in the original dataset and the selected subset. The cows were divided into nine culling categories according to the breeding documentation (the analysis did not include the “old age” reason), whereas the reproductive disorder category also comprised sterility (Table 1).

Table 1. Culling categories between 2017 and 2018 according to the breeding documentation in the original dataset and the selected subset.

Class	Lactation 1				Lactation 2				Lactation 3			
	Original Dataset		Subset		Original Dataset		Subset		Original Dataset		Subset	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
LMY ¹	2849	3.72	195	3.97	5416	3.38	427	3.40	7323	3.15	608	3.12
UD ²	9973	13.03	686	13.98	22,811	14.23	1862	14.80	35,205	15.12	2990	15.36
RD ³	33,619	43.92	1878	38.28	69,526	43.36	5090	40.47	97,116	41.71	7741	39.76
CD ⁴	121	0.16	2	0.04	251	0.16	13	0.10	352	0.15	30	0.15
MDSD ⁵	5654	7.39	435	8.87	12,621	7.87	1085	8.63	19,598	8.42	1746	8.97
RSD ⁶	533	0.70	41	0.84	1060	0.66	87	0.69	1485	0.64	126	0.65
LSD ⁷	8062	10.53	591	12.05	16,807	10.48	1438	11.43	25,347	10.89	2250	11.56
ACC ⁸	8177	10.68	577	11.76	16,455	10.26	1360	10.81	24,109	10.36	2109	10.83
OTH ⁹	7552	9.87	501	10.21	15,383	9.59	1215	9.66	22,274	9.57	1870	9.60

¹ low milk yield, ² udder diseases, ³ reproductive disorders, ⁴ contagious diseases, ⁵ metabolic and digestive system diseases, ⁶ respiratory system diseases, ⁷ locomotor system diseases, ⁸ accidents, ⁹ others (the “others” category included those culling reasons that were difficult to classify to the remaining categories or more than one culling reason existed).

The survival time is expressed in months from calving to culling or the end of lactation (the analysis included three successive lactations). A variable indicating censored observations was also created (with values of 0 and 1, 1—with censored observation, i.e., those that survived the observation period and their further fate was not followed at a given moment of analysis and 0—cows with full observation, i.e., those that were culled during the observation period) [16,18,32,33,37,39,45].

For the initial analysis, survival tables (presented in Supplementary Tables S1–S3), including individual indicators in class intervals were constructed for each culling category and lactation number [41] (the number of cases entering the interval alive; the number of deaths in the interval; the proportion of deaths, i.e., the ratio of the number of deaths to the number of cases at risk in the interval; the cumulative proportion of surviving cows until the time interval (Sp) equal to the product of the survival probabilities of all preceding intervals; a probability density function (f_i), i.e., the probability of culling in the time interval per unit time [39]:

$$f_i = \frac{P_i - P_{i+1}}{h_i} \quad (1)$$

and the hazard rate (γ_i), i.e., the probability (per unit time) that a cow that survived to the start of the interval would be culled in this interval:

$$\gamma_i = \frac{f_i}{0.5 \cdot (P_i + P_{i+1})} \quad (2)$$

where f_i is the probability density function, P_i is the cumulative proportion of surviving cases at the beginning of the interval and the beginning of the subsequent interval (P_{i+1}), h_i is the time interval width).

Survival curves were compared for the first three lactations of cows culled for different reasons, recorded in the breeding documentation [41,49]. The Kaplan–Meier method was applied for the survival curve analysis. It estimates the survival function directly from the continuous survival time (the estimate is independent of data grouping) [14,33,49,50]. The probability estimate was calculated as the product of successive conditional probabilities estimated separately (with censored observations) [34,41].

The Kaplan–Meier estimator (Kaplan i. Meier, 1958) was estimated according to the following formula [34,37,39]:

$$\hat{S}(t) = \prod_{t_j \leq t} \left(1 - \frac{d_j}{r_j} \right) \quad (3)$$

where r_j is the number of cows at risk at time t_j , d_j is the number of cows culled at time t_j , Π is the product of all cases lower than or equal to t .

The significances between the survival curves were determined using the Cox–Mantel statistic (C), which (assuming the null hypothesis is true) has a standard normal distribution [43]:

$$C = \frac{U}{\sqrt{I}} \quad (4)$$

$$U = r_2 - \sum_{i=1}^k m_i A(i) \quad (5)$$

$$I = \sum_{i=1}^k \frac{m_i(r_i - m_i)}{r_i - 1} \cdot A(i)(1 - A(i)) \quad (6)$$

where k is the number of time moments in which cows were culled, m_i is the number of cows culled at the i th time moment, r_i is the number of observations from both groups being compared, in which the survival times are at least as long as or longer than the i th time moment, $A(i)$ is the proportion of observations from Group 2, r_2 is the total number of culled cows from Group 2.

Based on the Kaplan–Meier survival curves, survival probabilities ($P1$) were calculated for cows at the beginning of the age interval ($w1$), in which cow culling began, as well as at the end of the decrease when the curve stabilized for a given culling category ($P2$ and $w2$). The rate of decrease in survival probability was also calculated as the difference, i.e., $DP = (P1 - P2) \cdot 100\%$ and expressed as a percentage (DP).

For additional descriptions of cow survivability, regression model parameters were estimated and the values of the hazard ratio (*HR*) were determined over three successive lactations to identify the predictors most strongly correlated with the survival times of cows in each culling category [16,34,35,37,44,45,51].

The following variables were included in the model: calving season (1—autumn, 2—winter, 3—spring, 4—summer), the percentage of Holstein-Friesian genes (%), age at first calving (months), age at first breeding (months), the average number of days in milk during a complete lactation, the length of calving interval (days), the components of the production index used for the evaluation of Holstein-Friesian cows in Poland (production subindex (PSI), conformation subindex (CSI), fertility subindex (FSI), breeding value for somatic cell count (BVSCC), breeding value for longevity (BVL), body frame subindex (BFSI), strength and milk yield subindex (SMYSI), leg and hoof subindex (LHSI), udder subindex (USI), heifer conception rate (HCR), cow conception rate (CCR), the interval from calving to first insemination (CFI), calving-to-conception interval (CCI)), as well as milk performance traits (average daily milk yield (kg), the average milk fat content (%), and the average milk protein content (%)). The manner of calculating subindices is presented in Supplementary File S1. The partial maximum likelihood method was used for parameter estimation (an algorithm based on the Newton–Raphson iterative scheme) [52]. The Cox proportional hazards model was applied, which expresses the risk at time *t* for the analyzed set of independent variables (Cox, 1972):

$$h(t : x_1, x_2, \dots, x_k) = h_0(t) e^{\sum_{i=1}^k a_i x_i} \quad (7)$$

where $h(t : x_1, x_2, \dots, x_k)$ is the resulting hazard given *k* predictors and the appropriate survival time, *e* is the natural logarithm base, $\sum_{i=1}^k a_i x_i$ is the linear combination of independent variables (x_1, x_2, \dots, x_k), and model parameters (a_1, a_2, \dots, a_k), $h_0(t)$ is the baseline hazard dependent on duration only.

The hazard ratio (HR_i) was determined as:

$$HR_i = e^{\sum_{i=1}^k a_i (\bar{x}_i - x_i)} \quad (8)$$

where $\sum_{i=1}^k a_i (\bar{x}_i - x_i)$ is the (summary) total effect of explanatory variables on the final risk (hazard) for a given cow, independent of duration, a_i is the estimated regression coefficient.

HR_i indicates the change in the risk of survival time shortening for one unit increase in the independent variable when adjusted for the remaining independent variables included in the model (it is assumed that they are constant when the independent variable increases by one unit). When $HR > 1$, the risk of survival time shortening increases, when $HR < 1$, the risk decreases. The values close to unity are interpreted as no risk. These values were calculated only for statistically significant regression coefficients.

In addition, a cluster analysis was carried out to show the similarities between cows culled for different reasons. The above-mentioned variables were used for the description of cull cows. The hierarchical Ward's method was applied in the first stage of the cluster analysis [53]. It aims to obtain clusters with cases (culling categories) that are as similar as possible to each other and as different as possible from cases (culling categories) belonging to other clusters [54,55]. This can be obtained by merging all possible cluster pairs and selecting, each time, the cluster with the minimum sum of squared deviations [56–59] using an approach based on the analysis of variance to determine the distance between clusters [55,60–63]. The measure of the distance between cases (culling categories) and the mean value of a given cluster was the error sum of squares (*ESS*), given by the following formula [64–66]:

$$ESS = \sum_{i=1}^k (x_i - \bar{x})^2 \quad (9)$$

where x_i is the value of the variable that is a clustering criterion for the i th case, k is the number of cases (culling categories) within the cluster, \bar{x} is the mean value of this variable within the cluster.

The clusters created using Ward's method were presented in the dendrogram showing the hierarchical structure of the set of culling categories [67] with decreasing similarities between them. Consequently, the analysis of the dendrogram revealed clusters formed by culling categories that were most similar to each other and represented by separate branches. The optimal number of clusters was finally determined through the pruning of the dendrogram. All the assumptions of Ward's method were also verified [62,63].

To confirm the number of clusters determined with Ward's method, the non-hierarchical k -means clustering algorithm was applied [47,48,68,69]. At first, all grouping variables were standardized [47,70] and the initial cluster centers were determined [68,71,72] by sorting all the distances between cases (culling categories), taking those at constant intervals [73]. Subsequently, the algorithm iteratively updated cluster centers [68,74] and the distances between these centers and individual cases (culling categories) to obtain the highest possible similarity between culling categories within clusters [71] and the highest possible degree of heterogeneity among clusters [60,69,75,76]. The Euclidean distance between culling categories and cluster centers was used [47,64,68,72]:

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^p (x_i - y_i)^2} \quad (10)$$

where $\mathbf{x} = (x_1, x_2, \dots, x_p)$, $\mathbf{y} = (y_1, y_2, \dots, y_p)$, p is the number of variables defining the p -dimensional space.

The number of clusters obtained with k -means clustering was the same as that determined with Ward's method. All the assumptions of k -means clustering (i.e., the lack of collinearity among the grouping variables and the lack of outliers) were also verified [70,74,77–79].

All the calculations were carried out using Statistica 13.3 software (Tibco, Inc., Tulsa, OK, USA).

3. Results

3.1. Survival Tables and the Kaplan–Meier Survival Curves

Cows in their first lactation were most frequently culled between 24 and 31 months of age and the main culling reasons were udder diseases (about 20%) and locomotor system diseases (18%). The hazard rates ranged from 0.058 for udder diseases to 0.083 for metabolic and digestive system diseases (Supplementary Table S1). There were no cows culled due to contagious diseases. The first culling already occurred at about 20 months of age on average, whereas culling in older cows (above 31 months of age) was not observed in the first lactation.

Based on the Kaplan–Meier functions, the survival curve decline could be observed at around 24–25 months of age (until about 31 months of age) for cows from individual culling categories (Figure 1). The smallest kink in the curve was noted only for cows culled due to reproductive problems (the difference in the cumulative proportions of surviving cases (DP) was about 7%). The sharpest kink could be seen for cows culled due to metabolic and digestive system diseases ($DP = 43\%$) (Table 2). The shapes of these two curves were statistically significantly different from the curve shapes of cows from the remaining culling categories (Table 3). The survival probability for cows culled for different reasons ranged from 0.6094 for those from the respiratory system disease category to 0.9276 for those from the reproductive disorder category (Supplementary Table S2).

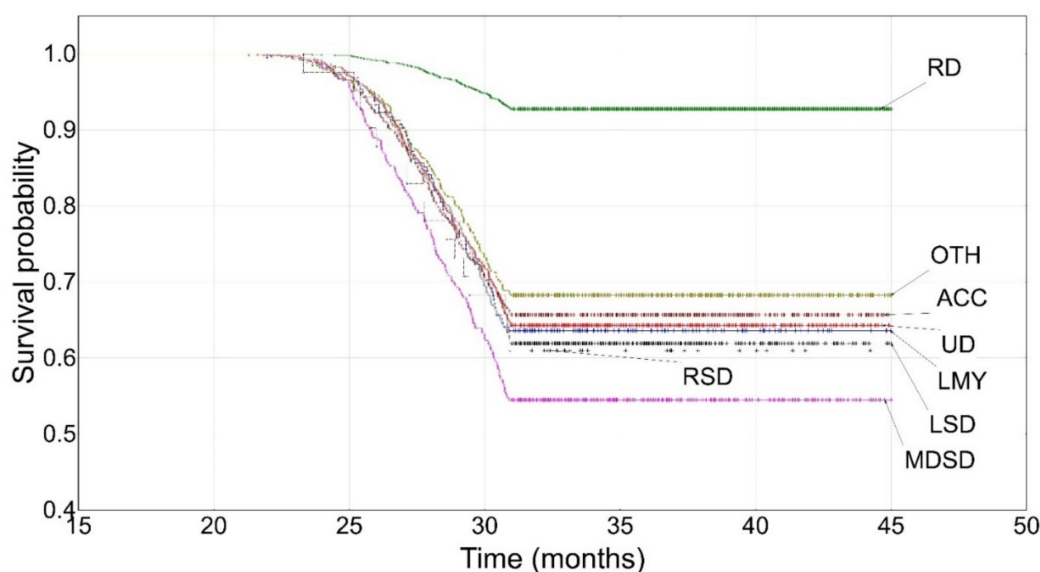


Figure 1. The Kaplan–Meier survival probability until the first lactation. UD—udder diseases, RD—reproductive disorders, MDSD—metabolic and digestive system diseases, RSD—respiratory system diseases, LSD—locomotor system diseases, ACC—accidents, OTH—others, LMY—low milk yield.

The highest culling number (until the second lactation, inclusive) was observed from 36 to 60 months of age and the most frequent culling categories were reproductive disorders (almost 33%) and udder diseases (17%). The hazard rates ranged from 0.08 to 0.12 (Supplementary Table S2). The lowest culling numbers were recorded for respiratory system diseases (0.8%) and contagious diseases (0.1%). The shallowest slope of the Kaplan–Meier survival curve was found for cows with reproductive problems ($DP = 68\%$) (Figure 2). Only at around 45 months of age did a noticeable decline in the curve start, which stabilized at about 60 months of age at a survival probability of 0.2229. The survival curve for cows culled due to contagious diseases dropped sharply at around 38 months of age and stabilized ten months later at a survival probability of 0.2778. This curve was quite similar to that for cows culled due to respiratory system diseases; however, the latter stabilized later (at about 58 months of age) at a survival probability of 0.0394. The survival curves for cows culled due to low milk yield and udder diseases were also characterized by similar shapes. The decline was observed at about 36 months of age and stabilized at 60 months of age at a survival probability of 0.07. The curve for cows culled due to low milk yield additionally differed from that for cows culled for other reasons (the “others” category included those culling reasons that were difficult to classify to the remaining categories or more than one culling reason existed), whereas the curve for cows culled due to udder diseases was different from that for cows culled due to metabolic and digestive system diseases. The curves for cows culled for the remaining reasons declined at about 36 months of age and stabilized at around 60 months of age at a survival probability of 0.0553–0.1153.

Table 2. Cumulative proportions of cows surviving (*P1* and *P2*) until the age interval (*w1* and *w2*) and the percentage decrease (*DP*).

Reason	Until the First Lactation					Until the Second Lactation					Until the Third Lactation				
	<i>P1</i>	<i>w1</i>	<i>P2</i>	<i>w2</i>	<i>DP</i>	<i>P1</i>	<i>w1</i>	<i>P2</i>	<i>w2</i>	<i>DP</i>	<i>P1</i>	<i>w1</i>	<i>P2</i>	<i>w2</i>	<i>DP</i>
LMY ¹	0.9708	24	0.6364	31	33.44	0.8000	36	0.0742	60	72.58	0.8325	40	0.0742	71	75.83
UD ²	0.9718	24	0.6429	31	32.89	0.8280	36	0.0742	60	75.38	0.8613	40	0.0742	71	78.71
RD ³	0.9973	25	0.9276	31	6.97	0.9001	40	0.2229	60	67.72	0.9012	40	0.2620	70	63.92
CD ⁴	-	-	-	-	-	0.9999	38	0.2778	47	72.21	0.9999	40	0.3771	70	62.28
MDSD ⁵	0.9711	24	0.5455	31	42.56	0.8000	36	0.0553	60	74.47	0.8207	40	0.0553	71	76.54
RSD ⁶	0.9705	24	0.6098	31	36.07	0.7772	36	0.0394	58	73.78	0.8206	40	0.0394	71	78.12
LSD ⁷	0.9722	24	0.6201	31	35.21	0.8280	36	0.0872	60	74.08	0.8613	40	0.0872	71	77.41
ACC ⁸	0.9721	24	0.6521	31	32.00	0.8532	36	0.1020	60	75.12	0.8613	40	0.1020	71	75.93
OTH ⁹	0.9705	24	0.6751	31	29.54	0.8557	36	0.1153	60	74.04	0.8676	40	0.1153	71	75.23

¹ low milk yield, ² udder diseases, ³ reproductive disorders, ⁴ contagious diseases, ⁵ metabolic and digestive system diseases, ⁶ respiratory system diseases, ⁷ locomotor system diseases, ⁸ accidents, ⁹ others (the "others" category included those culling reasons that were difficult to classify to the remaining categories or more than one culling reason existed), *P1*—the cumulative proportion of cows surviving until the interval in which culling started; *P2*—the cumulative proportion of cows surviving until the interval in which culling ended; *w1*—the month of life of a cow, in which culling started for a given category; *w2*—the month of life of a cow, in which culling ended for a given lactation; *DP*—the rate of decrease = $(P1 - P2) \cdot 100\%$; the values determined from Kaplan–Meier curves.

Table 3. *p*-values for comparisons between survival curves for cows from individual culling categories.

Category	First Lactation							
	UD ¹	RD ²	CD ³	MDSD ⁴	RSD ⁵	LSD ⁶	ACC ⁷	OTH ⁸
LMY ⁹	0.8505	0.0000		0.0265	0.7102	0.7330	0.7049	0.2597
UD		0.0000		0.0005	0.6206	0.4351	0.7622	0.1799
RD				0.0000	0.0000	0.0000	0.0000	0.0000
MDSD					0.4333	0.0067	0.0003	0.0000
RSD						0.8199	0.5470	0.3106
LSD							0.3071	0.0440
ACC								0.3150
Until the Second Lactation								
LMY	0.7113	0.0000	0.2995	0.0001	0.0789	0.5368	0.1325	0.0236
UD		0.0000	0.3344	0.0000	0.0378	0.6205	0.0534	0.1591
RD			0.0037	0.0000	0.0000	0.0000	0.0000	0.0000
CD				0.1073	0.1750	0.3661	0.4543	0.5868
MDSD					0.8175	0.0000	0.0000	0.0000
RSD						0.0936	0.0762	0.0171
LSD							0.1800	0.0356
ACC								0.2776
Until the Third Lactation								
LMY	0.0110	0.0000	0.0026	0.0247	0.2958	0.0329	0.0402	0.0022
UD		0.0000	0.0098	0.0000	0.0136	0.3608	0.5817	0.2602
RD			0.8454	0.0000	0.0000	0.0000	0.0000	0.0000
CD				0.0003	0.0015	0.0087	0.0107	0.0205
MDSD					0.9918	0.0000	0.0000	0.0000
RSD						0.0237	0.0273	0.0062
LSD							0.9185	0.1517
ACC								0.1325

¹ udder diseases, ² reproductive disorders, ³ contagious diseases, ⁴ metabolic and digestive system diseases, ⁵ respiratory system diseases, ⁶ locomotor system diseases, ⁷ accidents, ⁸ others, ⁹ low milk yield.

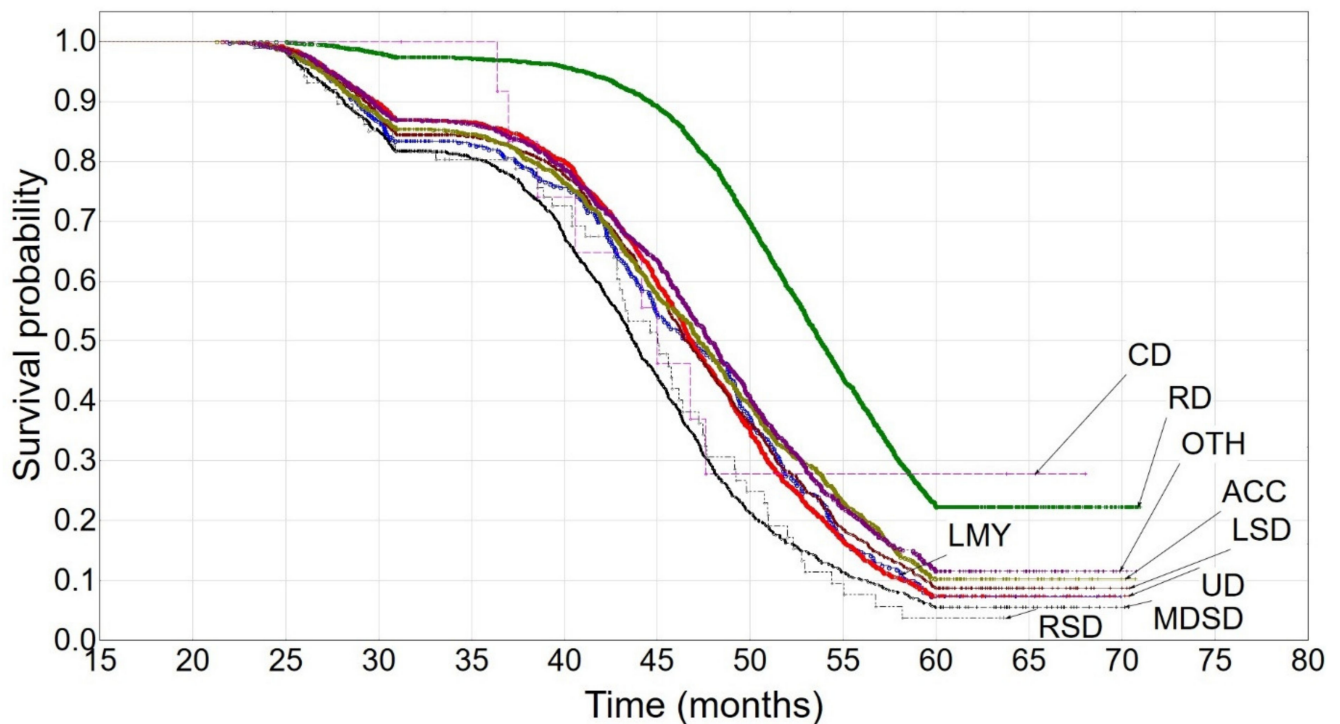


Figure 2. The Kaplan–Meier survival probability until the second lactation. UD—udder diseases, RD—reproductive disorders, CD—contagious diseases, MDSD—metabolic and digestive system diseases, RSD—respiratory system diseases, LSD—locomotor system diseases, ACC—accidents, OTH—others, LMY—low milk yield.

The highest culling number (until the third lactation, inclusive) was observed from 40 to 71 months of age. Cows culled due to reproductive disorders (32%) and udder diseases (almost 18%) predominated. The lowest culling number was recorded for contagious diseases (0.14%) and respiratory system diseases (0.74%). The hazard rates ranged from 0.041 to 0.073. The relatively shallow slope of the curve was characteristic of cows culled due to contagious diseases and reproductive disorders (Figure 3). The decline began at about 45 months of age and stabilized at around 70 months of age ($P_2 = 0.3771$ for contagious diseases and $P_2 = 0.2620$ for reproductive disorders; an approximately 63% decrease). The survival probability for the remaining curves ranged from 0.0394 (respiratory system diseases) to 0.1153 (others). The survival curve for cows with reproductive disorders (such as that for metabolic and digestive system diseases) was statistically significantly different from the remaining ones. It did not only differ from the curve for respiratory system diseases. The curves for cows with low milk yield were different from those for the “others” category and the curves for udder diseases differed from those for respiratory system diseases and accidents. The curve for locomotor system diseases was also different from the “others” category (Table 3).

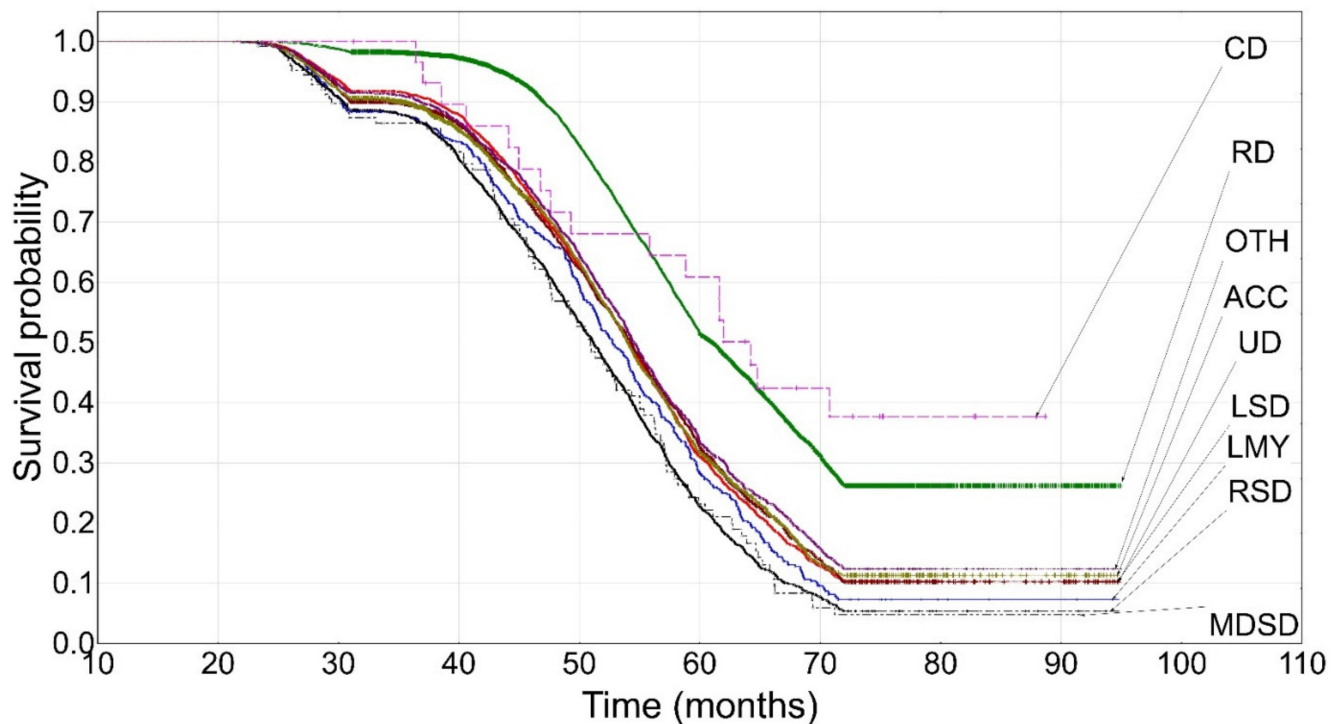


Figure 3. The Kaplan–Meier survival probability until the third lactation. UD—udder diseases, RD—reproductive disorders, CD—contagious diseases, MDSD—metabolic and digestive system diseases, RSD—respiratory system diseases, LSD—locomotor system diseases, ACC—accidents, OTH—others, LMY—low milk yield.

The estimators obtained with the Kaplan–Meier method are independent of data grouping. Therefore, they do not directly correspond to the results from survival tables based on the Weibull distribution (Supplementary Tables S1–S3).

3.2. The Cox Proportional Hazards Model Parameters Affecting Cow Survival and Hazard Ratio (HR) Coefficients

The probability values indicating the statistical significance of the parameter estimates for the Cox proportional hazards model determined with the Wald statistic are shown in Table 4, together with the hazard ratio(s) (*HR*) showing the risk of survival time shortening according to different culling categories and lactation numbers. For cows culled due to low milk yield, statistically significant *HR* values until the first lactation ranged between 0.10 for the age at first calving and 1.26 for the percentage of HF genes. The respective values until the second and third lactations ranged from 0.53 to 1.03 and from 0.47 to 1.03 for the same two factors (temperament and BVL, respectively). For the udder disease category, statistically significant *HR* values until the first lactation ranged between 0.05 for the age at first calving and 1.02 for PSI. The respective values until the second and third lactations ranged from 0.44 for milk protein content to 1.23 for milk fat content and from 0.71 for temperament to 1.21 for milk fat content. For the reproductive disorder category, statistically significant *HR* values until the first lactation ranged between 0.30 for milk protein content and 1.03 for BVL. The respective values until the second and third lactations ranged from 0.34 for milk protein content to 1.36 for milk fat content and from 0.53 for milk protein content to 1.09 for CSI. No statistically significant parameters were found for the contagious disease category, except for milk fat content until the second lactation at a very high value of *HR* (19.41). For the metabolic and digestive system disease category, statistically significant *HR* values until the first lactation ranged between 0.09 for the age at first calving and 1.46 for temperament. The respective values until the second and third lactations ranged from 0.31 to 1.27 and from 0.26 to 1.30 for the same two factors (milk

protein content and milk fat content, respectively). For the respiratory system disease category, statistically significant *HR* values until the first lactation ranged between 0.03 for the age at first calving and 0.88 for lactation length. The respective values until the second and third lactations ranged from 0.55 for the percentage of HF genes to 1.07 for BVL and from 0.86 for the age at first calving to 1.12 for BVL. For the locomotor system disease category, statistically significant *HR* values until the first lactation ranged between 0.22 for the age at first calving and 1.26 for USI. The respective values until the second and third lactations ranged from 0.33 for milk protein content to 1.08 for LHSI and from 0.59 for temperament to 1.20 for milk fat content. For the accident category, statistically significant *HR* values until the first lactation ranged between 0.32 for the age at first calving and 1.79 for FSI. The respective values until the second and third lactations ranged from 0.28 for milk protein content to 1.20 for FSI and from 0.26 for milk protein content to 1.20 for milk fat content. For the “others” category, statistically significant *HR* values until the first lactation ranged between 0.24 for the age at first calving and 1.11 for BFSI. The respective values until the second and third lactations ranged from 0.32 for milk protein content to 1.35 for milk fat content and from 0.37 for milk protein content to 1.22 for milk fat content and CSI.

Table 4. HR values for the individual parameters of the Cox proportional hazards model according to culling categories and lactation groups (1—until the first lactation, 2—until the second lactation, 3—until the third lactation; HR values for statistically significant parameters are marked in bold).

Category	LMY ¹			UD ²			RD ³			CD ⁴			MDS ⁵			RSD ⁶			LSD ⁷			ACC ⁸			OTH ⁹		
	1	2	3	1	2	3	1	2	3	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	
CS ¹⁰	1.04	1.07	0.98	0.95	1.08	0.98	1.15	0.99	0.99	1.52	1.22	0.89	1.03	1.00	0.45	0.84	1.04	0.99	1.01	1.04	0.87	0.95	1.00	0.73	1.04	1.00	
HF ¹¹	1.26	1.00	0.99	1.00	0.99	0.98	0.95	0.99	0.98			0.98	0.99	1.00	0.16	0.55	0.77	1.03	0.99	0.99	0.94	0.97	0.97	0.98	0.97	0.99	
AFC ¹²	0.10	0.82	0.89	0.05	0.78	0.85	0.40	0.82	0.86	0.87	0.96	0.09	0.77	0.88	0.03	0.59	0.86	0.22	0.77	0.85	0.32	0.77	0.84	0.24	0.80	0.86	
AFB ¹³	0.87	0.99	0.98	1.01	1.03	1.02	0.97	1.00	0.99	1.78	0.98	0.99	0.99	0.99	0.98	0.95	0.94	0.92	1.03	1.03	0.98	1.04	1.02	0.87	1.01	1.02	
LL ¹⁴	0.92	0.99	1.00	0.91	0.99	1.00	0.96	1.00	1.00	0.96	1.00	0.90	0.99	1.00	0.88	0.99	1.00	0.94	0.99	1.00	0.96	0.99	1.00	0.94	0.99	1.00	
CI ¹⁵		0.99	1.00		0.99	1.00		0.99	0.99	1.01	0.99		0.99	1.00		0.99	0.99		0.99	1.00		1.00	1.00		0.99	1.00	
PSI ¹⁶	0.98	1.02	1.02	1.02	0.99	1.02	1.01	0.98	1.00	0.98	1.16	1.00	1.02	1.02	1.17	0.99	1.00	1.02	1.01	1.01	1.00	1.02	1.02	0.97	1.00	1.02	
CSI ¹⁷	0.96	1.01	1.01	0.87	1.11	1.03	0.67	0.97	1.09	0.85	0.01	1.22	1.05	1.13	1.02	1.11	0.71	0.71	0.87	1.13	1.59	0.88	1.05	0.95	1.19	1.22	
BFSI ¹⁸	0.94	1.00	1.00	1.05	0.99	1.01	1.01	1.01	0.99	1.05	2.63	0.98	1.00	0.99	0.98	0.95	1.06	1.11	1.05	0.99	0.97	1.02	1.00	1.11	0.98	0.98	
SMYSI ¹⁹	1.10	1.00	1.00	1.02	0.99	1.00	1.08	1.01	1.00	0.88	2.07	0.97	1.01	1.00	1.06	1.01	1.11	1.05	1.01	1.01	0.95	1.03	1.02	0.91	0.98	0.99	
LHSI ²⁰	0.97	0.98	0.99	1.09	0.96	0.99	1.17	1.01	0.97	1.55	11.30	0.94	0.97	0.95	1.05	1.00	1.14	1.13	1.08	0.96	0.85	1.06	0.98	1.07	0.95	0.93	
USI ²¹	0.91	1.00	1.00	1.05	0.92	0.98	1.37	1.03	0.96	0.88	25.26	0.87	0.98	0.92	1.04	0.91	1.30	1.26	1.10	0.92	0.73	1.11	0.98	1.09	0.91	0.88	
FSI ²²	1.08	1.25	0.97	1.08	0.99	1.01	1.04	0.99	1.01	0.11	0.14	0.85	0.96	0.98	1.02	1.13	1.18	1.13	1.01	1.02	1.79	1.20	1.09	0.95	1.13	1.11	
HCR ²³	0.91	0.81	1.03	0.93	1.00	0.98	1.04	1.00	0.99	7.68	6.13	1.19	1.03	1.01	1.06	0.89	0.85	0.89	0.97	0.97	0.62	0.83	0.91	1.03	0.86	0.89	
CCR ²⁴	0.95	0.98	0.99	0.97	1.00	1.00	0.95	1.00	0.99	1.14	1.21	0.98	1.00	1.02	0.98	0.97	1.02	0.97	1.02	1.01	0.89	0.98	1.00	1.03	1.01	0.99	
CFI ²⁵	0.96	0.99	0.96	0.98	0.97	0.98	1.01	1.01	1.00	1.18	1.93	1.05	1.00	1.02	1.05	0.98	1.03	0.97	1.00	1.01	0.86	0.98	1.01	1.02	0.97	0.98	
CCI ²⁶		0.95	1.06		1.02	1.02		0.99	1.01	1.10	0.96		1.02	0.99		1.03	0.91		1.00	0.99		0.98	0.98		1.01	1.00	
BVSCC ²⁷	0.95	1.01	1.01	1.00	0.99	1.01	1.00	1.00	0.99	0.82	1.21	1.00	1.01	1.00	0.99	0.96	1.00	1.05	1.01	1.00	0.95	0.98	0.99	0.97	1.00	1.00	
BVL ²⁸	1.08	1.03	1.03	1.00	1.03	1.03	1.03	1.02	1.02	1.11	0.87	1.02	1.02	1.03	1.05	1.07	1.12	1.00	1.01	1.03	1.05	1.03	1.03	0.97	1.01	1.02	
Temp ²⁹	0.83	0.53	0.47	0.79	0.63	0.71	0.79	0.81	0.94	2.72	0.42	1.46	0.47	0.49	0.38	0.88	0.62	1.13	0.51	0.59	0.81	0.61	0.58	0.81	0.70	0.67	
MY ³⁰	0.98	0.96	0.96	0.96	1.01	0.98	1.02	1.02	0.99	1.17	0.77	1.02	0.97	0.97	1.05	1.00	0.97	1.00	0.99	0.98	1.00	0.98	0.98	1.04	1.00	0.99	
Fat ³¹	0.79	1.12	0.92	1.00	1.23	1.21	0.89	1.36	1.05	19.41	2.10	1.01	1.27	1.30	2.05	1.64	1.17	0.90	1.12	1.20	1.08	1.18	1.20	0.84	1.35	1.22	
Prot ³²	0.90	0.72	0.76	0.90	0.44	0.17	0.30	0.34	0.53	0.00	0.00	0.93	0.31	0.26	0.22	0.37	0.46	0.66	0.33	0.36	0.85	0.28	0.26	1.08	0.32	0.37	

¹ low milk yield, ² udder diseases, ³ reproductive disorders, ⁴ contagious diseases, ⁵ metabolic and digestive system diseases, ⁶ respiratory system diseases, ⁷ locomotor system diseases, ⁸ accidents, ⁹ others, ¹⁰ calving season, ¹¹ percentage of HF genes, ¹² age at first calving, ¹³ age at first breeding, ¹⁴ lactation length, ¹⁵ calving interval, ¹⁶ production subindex, ¹⁷ conformation subindex, ¹⁸ body frame subindex, ¹⁹ strength and milk yield subindex, ²⁰ leg and hoof subindex, ²¹ udder subindex, ²² fertility subindex, ²³ heifer conception rate, ²⁴ cow conception rate, ²⁵ the interval from calving to first insemination, ²⁶ calving-to-conception interval, ²⁷ breeding value for somatic cell count, ²⁸ breeding value for longevity, ²⁹ temperament, ³⁰ milk yield, ³¹ fat percentage, ³² protein percentage.

3.3. Cluster Analysis

The results of the cluster analysis using Ward's method are presented in Figures 4–6. As can be seen from these figures, three clusters could be distinguished for the first lactation and the first two lactations, whereas four clusters were formed for the first three lactations. The reproductive disorder category clearly differed from the remaining ones and constituted a separate cluster in each of the three lactation groups. On the other hand, udder disease and low milk yield categories were similar and grouped in the same cluster in all dendrograms. Metabolic, digestive, and respiratory system diseases also occurred in one cluster for all lactation groups. The remaining categories rarely constituted separate clusters. In the third lactation group, locomotor system diseases and accidents formed a new separate cluster.

The *k*-means clustering confirmed the number of clusters determined with Ward's method (Table 5), which means that the cluster members were exactly the same as those visible in the dendrogram. Three clusters existed until the first lactation. The first cluster included only reproductive disorders, whereas the second and third ones consisted of udder disease, low milk yield, accidents, and the "others" categories, as well as metabolic and digestive system diseases, locomotor system diseases, and respiratory system diseases, respectively. Three clusters could also be distinguished until the second lactation. The first one was the same as previously. The second cluster additionally included locomotor system diseases. The third cluster contained contagious diseases instead of locomotor system diseases. Four clusters existed until the third lactation. The first one was identical as previously. The second cluster was similar to those of the first two lactation groups; however, it included contagious diseases instead of accidents (such as that for the first lactation group) or accidents and locomotor system diseases (such as that for the second lactation group). Finally, the last two clusters consisted of accidents and locomotor system diseases as well as metabolic and digestive system diseases and respiratory system diseases, respectively.

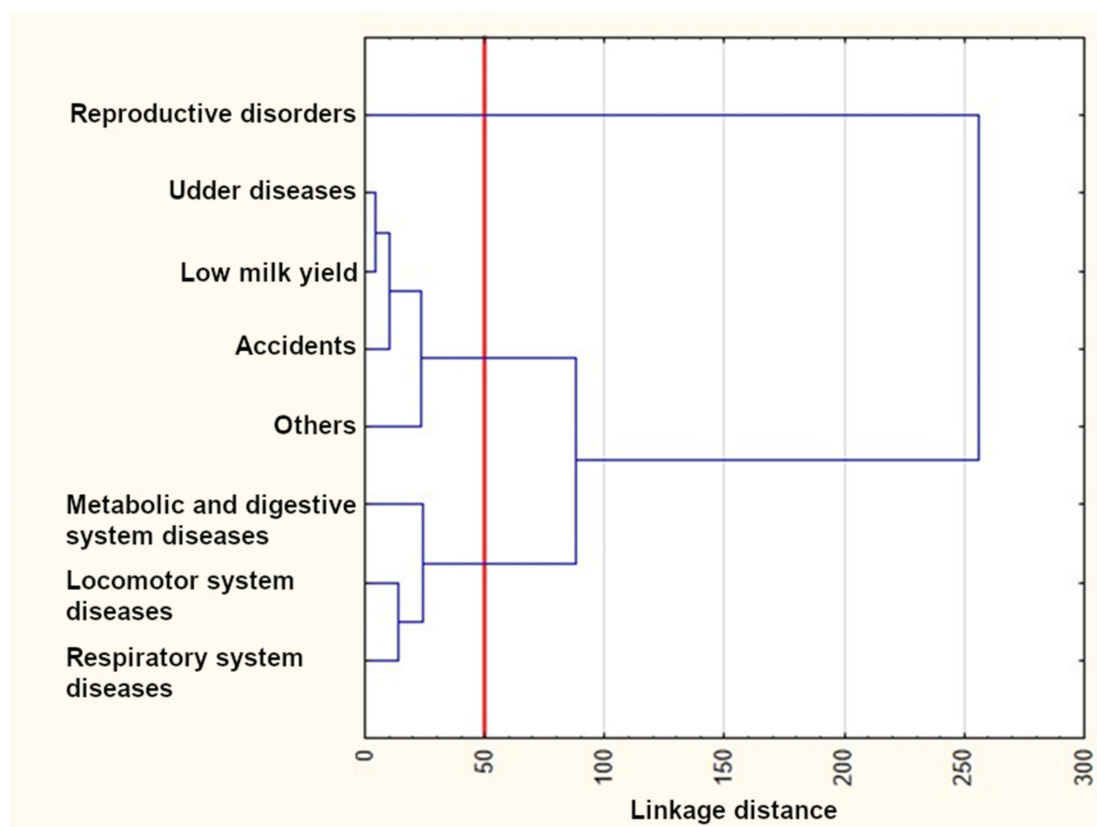


Figure 4. The results of the cluster analysis using Ward's method for the first lactation.

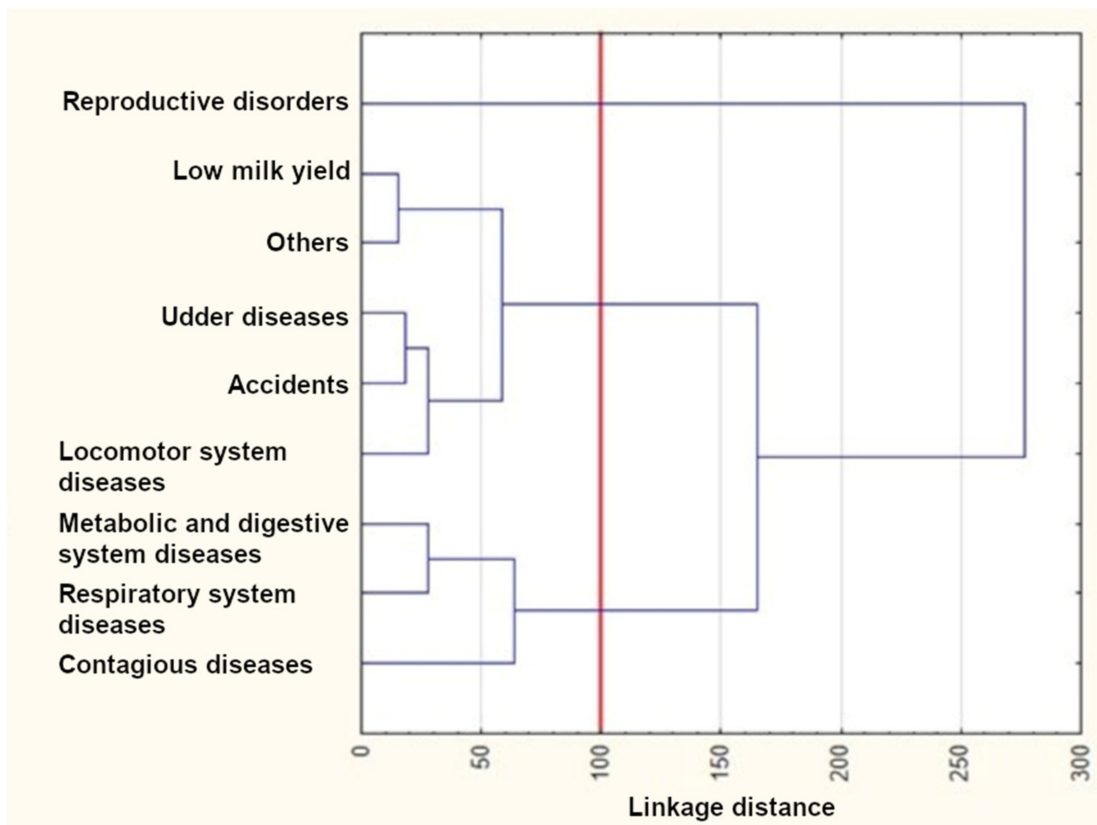


Figure 5. The results of the cluster analysis using Ward’s method for the first and second lactations.

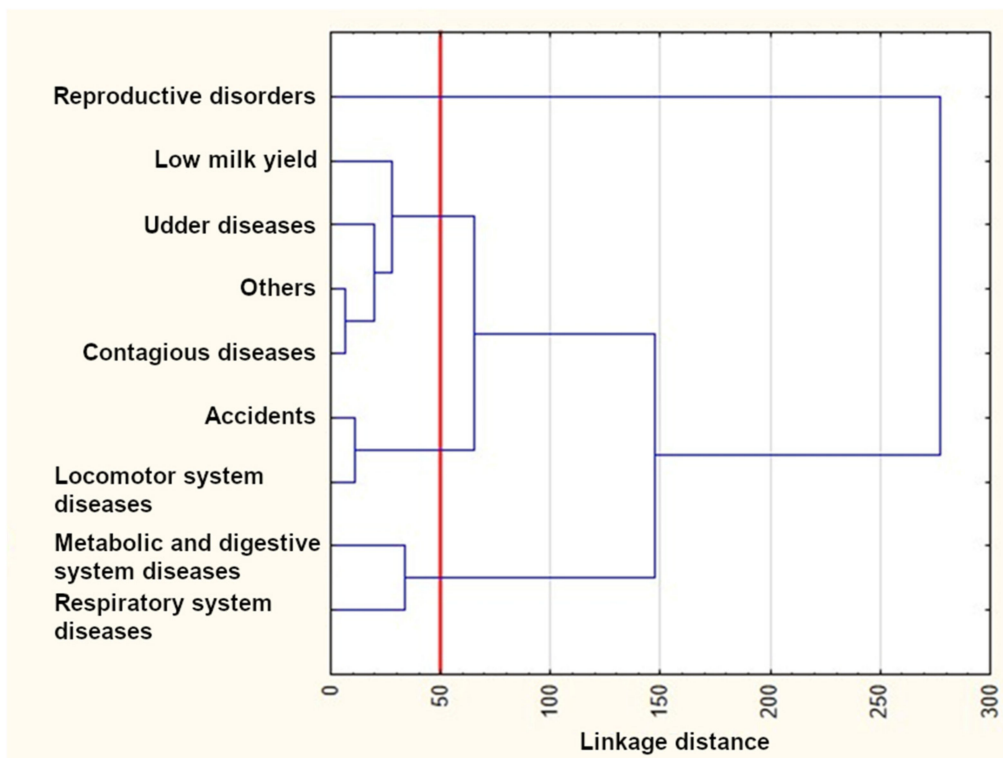


Figure 6. The results of the cluster analysis using Ward’s method for the first, second, and third lactations.

Table 5. Cluster elements for the *k*-means clustering method.

Cluster Number	Culling Category
First lactation	
1	Reproductive disorders
2	Udder diseases Others Low milk yield Accidents
3	Metabolic and digestive system diseases Locomotor system diseases Respiratory system diseases
First and second lactation	
1	Reproductive disorders
2	Udder diseases Others Low milk yield Accidents Locomotor system diseases
3	Metabolic and digestive system diseases Respiratory system diseases Contagious diseases
First, second and third lactation	
1	Reproductive disorders
2	Udder diseases Others Low milk yield Contagious diseases
3	Accidents Locomotor system diseases
4	Metabolic and digestive system diseases Respiratory system diseases

4. Discussion

As already mentioned, an important feature of survival analysis is the fact that information from cows that are not culled by the time the data recording is complete can also be used. Cows still alive at the end of the study period may cause difficulties in survival evaluation using other statistical methods since the actual life spans of the animals are not known. Therefore, the survival analysis offers numerous advantages over traditional linear models and allows for better fitting the survival data and using time-dependent variables appropriately [14,19,42]. The cluster analysis, on the other hand, creates homogenous groups of culling reasons. Cows have increased chances of remaining in the herd if they are healthy, reproduce regularly, have functional feet, legs, and udders, and produce enough milk [15]. Thus, knowledge of the association between various indicators monitored at the individual animal and herd levels with cow lifespans and culling reasons is essential to predict cow longevity and support optimal decisions in herd management [31]. The survival and cluster analyses performed in the present study showed some similarities between individual culling categories, which may indicate their direct connection. For instance, the survival curve for contagious diseases was similar to that for respiratory system diseases in the second lactation group (both categories grouped in the same cluster), which may suggest that pathogens causing respiratory system infections can finally lead to cow culling. The survival curve for low milk yield was also similar to that for udder diseases in the same lactation group (both included in the same cluster), which may imply that the milk yield of

cows with udder diseases (such as mastitis) decreases, ultimately resulting in their culling. Finally, the reproductive disorder category clearly differed from the remaining ones and formed a separate cluster. A more detailed description of the similarities and differences between individual survival curves and clusters and their consequences is given below.

The main culling categories analyzed in the present study (reproductive disorders—40%, udder diseases—13 to 15%, and locomotor system diseases—above 10%) corresponded to the frequency of culling reasons reported by others [15,24,80–84], although Reimus et al. [85] identified leg problems as the main culling category (above 26%). Similarly, Sewalem et al. [86] and Zavadilová et al. [87] indicated udder and leg diseases and all the resulting consequences as the most important culling reasons. The effect of lameness on culling is not unequivocal. Milian-Suazo et al. [88] and Barkema et al. [89] argued that lameness affected culling to a small extent. According to Gröhn et al. [90], the decision on culling due to lameness was mainly associated with calving age, milk yield, and fertility. Therefore, numerous variables should be included in the analysis of the influence of lameness on culling risk. Kugonza et al. [91] indicated fertility, poor productivity, health performance, as well as old age as the main factors affecting the culling rate. The remaining reasons (identified with the questionnaires on the long-horned Ankole cattle) included inbreeding, production of males only, bad temperament, high calf mortality, slow growth, unfavorable color, inappropriate conformation, and body condition. Ahlman et al. [92] analyzed six longevity traits in Swedish Holstein and Swedish Red cows (the length of productive life; survival through the first, second, and third lactations; fertility- and udder health-determined survival) and divided culling reasons into eight groups: udder health, low fertility, low production, leg problems, metabolic diseases, other diseases, other specified causes, and unspecified causes. Finally, economic factors affecting the culling rate include input and production prices and their seasonal variations [93,94].

The survival curves for cows from the reproductive disorder category clearly differed from the remaining ones in each of the three lactation groups (Figures 1–3 and Table 3) and constituted a separate cluster (Figures 4–6). Moreover, the curves for udder diseases and low milk yield were very similar (they differed significantly only for the third lactation) and were grouped in the same cluster. Furthermore, the survival curves for cows from the metabolic and digestive system disease category were different from most survival curves for the remaining culling categories (except for respiratory system diseases). The cluster analysis showed that digestive and respiratory system diseases occurred in one cluster for all lactations. In addition, the reproductive disorder category constituted a separate cluster for each of the three lactations. It should be emphasized that reproductive factors have low heritability, which means that their contribution to the variability of this trait is not great. It may suggest that not all cows culled due to reproductive problems were infertile [95]. The cows from the remaining categories rarely constituted separate clusters, which was also confirmed by the analysis of the differences among survival curves (Table 3). In the first lactation, the low milk yield and udder disease clusters could also be distinguished. It can be noticed that the three main clusters of cull cows were formed until the second lactation, whereas four clusters existed until the third lactation. It is also noteworthy that locomotor system diseases and accidents constituted a new separate cluster. This connection may suggest that the locomotor system disease and accident categories may have a lot in common; the health status of cows that are culled due to locomotor system diseases and accidents should be verified more precisely since accidents can implicate leg diseases or vice versa. Booth et al. [96] indicated that the culling risk due to lameness may be underestimated unless the temporal relationship between lameness and culling is included in the model. It results from the fact that such a model assumes that the culling risk of a lame cow is constant in time, even before the occurrence of lameness. On the other hand, Ahlman et al. [92] reported that low fertility was the main culling reason in the first lactation, regardless of breed or production system, whereas the change in the main culling cause from fertility (most common in the first lactation, irrespective of the production system) to udder health occurred at a lower age in organic herds.

Table 4 presents the relative hazard ratio(s) (*HR*) of culling a cow from individual culling categories, taking into account different factors. Fodor et al. [97] reported that the high age at first calving in heifers (above 30 months) increased culling probability by more than five times compared to that of less than 22 months. In the present study, the *HR* for the age at first calving was below unity practically for each culling category, which would suggest an increased survival chance with a one-month increase in the age at first calving. This result corresponds to the values close to 0.99 for a one-day increase in the age at first calving reported by Whitaker et al. [98]. However, it should be noticed that the age at first calving is characterized by a certain optimum range (22 to 25 months) and the values below or above this threshold contribute to a higher culling probability [99].

Calving season was practically insignificant for an increased culling risk in each culling category, except for udder diseases in the second lactation, in which this risk was slightly higher in the spring–summer season ($HR > 1$) and the “others” category, where it decreased in the same season ($HR < 1$). Bradley and Green [100] suggested that culling risk due to udder diseases increased over the autumn–winter period. Some authors indicated that the summer calving season reduced cow survival [95,98].

Gröhn et al. [90] reported that a higher milk yield decreased the culling rate due to udder diseases, reproductive problems, and metabolic diseases, which corresponds to the results of the present study to some extent. The *HR* values for milk yield were close to unity or slightly lower for these culling reasons. This coefficient was similar to that for SMYSI in the above-mentioned culling categories.

A significant association between locomotor system disease category and many variables was found. Increased age at first calving decreased the culling risk in each of the three lactations. Moreover, increased lactation length, temperament (in the second and third lactations), and the percentage of milk protein had the same effect. On the other hand, BFSI caused a higher culling risk in the first and second lactations ($HR = 1.11$ and $HR = 1.05$, respectively). Similarly, CSI, as a whole, increased this risk, especially in heifers ($HR = 1.55$). The same situation was observed for milk fat content in the third lactation ($HR = 1.2$) and USI (including traits such as: udder placement, fore udder attachment, rear udder attachment, medial suspensory ligament, udder width, rear teat placement, fore teat placement, and teat length) in the first lactation ($HR = 1.26$). Other traits had low *HR* values or their effects on culling risks were statistically non-significant.

The influence of exterior traits on the length of productive life in Slovak Simmental cows was analyzed by Čanji et al. [101] and Strapák et al. [102]. The udder had a great impact on this length among the main exterior traits. From the partial exterior traits, the largest effect on the length of productive life was observed for udder depth, teat length, thickness, and placement. Similar results were obtained by Strapák et al. [103], who found that udder depth, teat length, rear udder attachment, and the rear udder significantly influenced the length of productive life. In the study by Strapáková et al. [104] on Slovak Simmental cows, whose body conformation traits were evaluated according to the Fleckscore system, animals with deeper udders, stronger and tighter fore udder attachments, and centrally placed teats, achieved longer productive lives. The effect of udder traits on the culling rate was also observed in Slovak Spotted cattle by Strapák et al. [105]. The relationship between conformation traits and longevity indicators in Polish Holstein-Friesian Black-and-White cows was analyzed by Sawa et al. [106], who found that lifespan, the length of productive life, and the number of lactations depended primarily on udder conformation (udder as a general trait and conformation score) ($r = 0.11$) and, among detailed traits, on udder placement ($r = 0.14$) and fore udder attachment ($r = 0.10$). Vacek et al. [107], investigating the associations between conformation traits and longevity in Czech Holstein cattle, reported that rump angle, rear leg set, udder depth, and teat length were negatively correlated with the herd or productive life. Most body traits were weakly (but positively) associated with herd life, indicating that larger cows lived longer. However, body depth and chest width were not linearly related to longevity. The longest productive lives were observed in cows undersized in chest width and body depth and a

similar dependence was found for rump width. The ideal rear leg set for longevity was average or below average. Cows with well-attached fore udders, high-attached rear udders, strong central ligaments, close front teat placement, and moderately long teats showed the longest functional productive lives.

Studies by Čanji et al. [101], Strapák et al. [102], and Strapák et al. [103] also highlighted the effect of body frame on the culling rate in Slovak Simmental cattle. Among body conformation traits, rump angle, croup height, and body depth were the main factors affecting this rate. Strapáková et al. [104] found that smaller, longer, and well-muscled cows with deeper bodies had a lower risk of early culling, whereas hock angularity and pastern and hoof height with low or high scores increased this risk. Finally, the influence of body conformation traits on longevity in Slovak Spotted cows was investigated by Strapák et al. [105], who reported that body frame significantly affected the length of productive life in this breed.

According to Booth et al. [96], foot rot diagnosed during the second or third months of lactation decreased survival during the same time period (hazard ratio = 5.1; 95% confidence interval = 1.6 to 16.2). Sole ulcers occurring in the first four months of lactation reduced survival in several subsequent periods, in which the strongest association was observed between diagnosis in the third and fourth months of lactation and exit from the herd during the same period (hazard ratio = 2.7; 95% confidence interval = 1.3 to 6.0). In the study by Strapák et al. [102], feet and legs belonged to the main exterior traits affecting the culling rate. A more detailed analysis of the partial exterior traits showed that heel joint expression and feet significantly influenced the culling risk in Slovak Simmental cows. Similar results were obtained by Čanji et al. [101] and Strapák et al. [103], who found that fetlock, feet, and rear legs were the most important traits affecting the length of productive life. In the study by Sawa et al. [106], the conformation of legs and feet correlated with lifespan, the length of productive life, and the number of lactations ($r = 0.11$). Cramer et al. [108] included heifers, infectious lesions, white line lesions, hemorrhages, sole ulcers, other lesions, and free-stall housing as covariates in the final multivariate Cox proportional hazards model. Infectious hoof lesions were not significantly associated with culling probability. Hazard ratios for white line lesions, ulcers, and hemorrhages were 1.72, 1.26, and 1.36, respectively. The strength of the association between the grouped variables (other lesions) and culling risk decreased with time.

In the present study, the *HR* value for the percentage of milk protein was below unity for almost all culling categories, which suggested a decreased culling risk. A similar effect of increased milk protein content on culling rate was observed by Morek-Kopeć et al. [109], indicating that a lower protein content associated with the lactation stage and difficult calving may contribute to the culling of a cow. An increased percentage of milk fat was related to the higher culling risk for most culling categories, especially in the second and third lactations. A similar advantageous effect of protein content and a more neutral influence of fat content on the culling rate was found by Sasaki et al. [110], Chirinos et al. [111], and Morek-Kopeć et al. [109].

5. Conclusions

The identification of early predictors of a cow's productive life is of great significance to improve dairy herd profitability. The supply of quality replacement heifers, mitigation of transition cow disease risk, and good culling decisions have great impacts on the productivity and profitability of the herd. Culling decisions are made based on cows and external factors that are frequently specific to the herd. The survival analysis provides information about the potential of cows characterized by certain traits for breeders. It supports more rational management of reproduction and performance and the application of preventive measures that allow for expanding production and eliminating disadvantageous phenomena. The survival analysis may also provide some information about the size of the future population, enabling a better understanding of survival patterns according to population age, and determining the effects of different factors on these patterns, which was

presented and confirmed in the present study. The survival curves for cows from individual culling categories had similar shapes, although those for the reproductive disorder category (especially in the first lactation) and the contagious disease category (in the second and third lactations) had shallower slopes. The low level of probability of survival curves for cows from the metabolic and digestive system disease category and the respiratory system disease category could also be observed in each of the three lactations. The contagious disease category was almost non-existent in the first lactation.

The culling reason was not always precisely determined for the distinguished culling categories, which was confirmed by some results, e.g., the accident and locomotor system disease categories. A more accurate method of determining culling reasons would be required, especially for these two categories.

Among the indices from the breeding documentation included in the analysis, the greatest effect on the relative culling risk in individual lactations was observed for age at first calving, lactation length, calving interval, PSI, BVL, temperament, and average daily milk yield, whereas the smallest one was found for calving season, BFSI, SMYSI, CCR, CFI, and CCI.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ani12151942/s1>, Table S1: Survival tables for cows until the first lactation according to culling categories; Table S2: Survival tables for cows until the second lactation according to culling categories; Table S3: Survival tables for cows until the third lactation according to culling categories; Supplementary File S1: The PF (Productivity and Functionality) selection index.

Author Contributions: Conceptualization, W.G. and K.A.; methodology, W.G. and D.Z.; validation, W.G. and K.A.; formal analysis, J.W.; investigation, W.G. and D.Z.; resources, K.A. and J.W.; data curation, D.Z.; writing—original draft preparation, W.G. and D.Z.; writing—review and editing, K.A. and J.W.; visualization, W.G. and D.Z.; supervision, K.A. and J.W.; project administration, K.A. and J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data are available upon reasonable request from the corresponding author. The data set was provided by the Polish Federation of Cattle Breeders and Dairy Farmers in Warsaw (PFCBDF); Contract No. 47/2013, 36/2016 and PN/147/2019 between PFCBDF and the University of Agriculture in Krakow; the purchase of the data was financed from the funds for statutory activity, by the Ministry of Science and Higher Education in Poland.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Adamczyk, K. Dairy Cattle Welfare as a Result of Human-Animal Relationship—A Review. *Ann. Anim. Sci.* **2018**, *18*, 601–622. [[CrossRef](#)]
2. De Vries, A. Cow Longevity Economics: The Cost Benefit of Keeping the Cow in the Herd. In Proceedings of the Conference “Cow Longevity Conference”, Tumba, Sweden, 28–29 August 2013; pp. 22–52.
3. Sumner, C.L.; von Keyserlingk, M.A.; Weary, D.M. Perspectives of Farmers and Veterinarians Concerning Dairy Cattle Welfare. *Anim. Front.* **2018**, *8*, 8–13. [[CrossRef](#)]
4. Lovarelli, D.; Bacenetti, J.; Guarino, M. A Review on Dairy Cattle Farming: Is Precision Livestock Farming the Compromise for an Environmental, Economic and Social Sustainable Production? *J. Clean. Prod.* **2020**, *262*, 121409. [[CrossRef](#)]
5. Miele, M.; Veissier, I.; Evans, A.; Botreau, R. Animal Welfare: Establishing a Dialogue between Science and Society. *Anim. Welf.* **2011**, *20*, 103.
6. Rotz, C.A. Modeling Greenhouse Gas Emissions from Dairy Farms. *J. Dairy Sci.* **2018**, *101*, 6675–6690. [[CrossRef](#)] [[PubMed](#)]
7. Adamczyk, K.; Jagusiak, W.; Makulska, J. Analysis of Lifetime Performance and Culling Reasons in Black-and-White Holstein-Friesian Cows Compared with Crossbreds. *Ann. Anim. Sci.* **2018**, *18*, 1061–1079. [[CrossRef](#)]
8. Brouwer, H.; Stegeman, J.A.; Straatsma, J.W.; Hooijer, G.A.; van Schaik, G. The Validity of a Monitoring System Based on Routinely Collected Dairy Cattle Health Data Relative to a Standardized Herd Check. *Prev. Vet. Med.* **2015**, *122*, 76–82. [[CrossRef](#)] [[PubMed](#)]

9. Pritchard, T.; Coffey, M.; Mrode, R.; Wall, E. Understanding the Genetics of Survival in Dairy Cows. *J. Dairy Sci.* **2013**, *96*, 3296–3309. [[CrossRef](#)] [[PubMed](#)]
10. Van Pelt, M. *Genetic Improvement of Longevity in Dairy Cows*; Wageningen University and Research: Wageningen, The Netherlands, 2017; ISBN 10.18174/404192.
11. Lehenbauer, T.W.; Oltjen, J.W. Dairy Cow Culling Strategies: Making Economical Culling Decisions. *J. Dairy Sci.* **1998**, *81*, 264–271. [[CrossRef](#)]
12. Dziekiewicz-Mrugasiewicz, M.; Wierzbicka, M. The reasons of culling of cattle in dairy cows herd—A Review. *Pol. J. Nat. Sci.* **2021**, *36*, 197–209.
13. Weigel, K.A.; Palmer, R.W.; Caraviello, D.Z. Investigation of Factors Affecting Voluntary and Involuntary Culling in Expanding Dairy Herds in Wisconsin Using Survival Analysis. *J. Dairy Sci.* **2003**, *86*, 1482–1486. [[CrossRef](#)]
14. Pascottini, O.B.; Probo, M.; LeBlanc, S.J.; Opsomer, G.; Hostens, M. Assessment of Associations between Transition Diseases and Reproductive Performance of Dairy Cows Using Survival Analysis and Decision Tree Algorithms. *Prev. Vet. Med.* **2020**, *176*, 104908. [[CrossRef](#)]
15. Dallago, G.M.; Wade, K.M.; Cue, R.I.; McClure, J.T.; Lacroix, R.; Pellerin, D.; Vasseur, E. Keeping Dairy Cows for Longer: A Critical Literature Review on Dairy Cow Longevity in High Milk-Producing Countries. *Animals* **2021**, *11*, 808. [[CrossRef](#)]
16. Ojango, J.M.K.; Ducrocq, V.; Pollott, G.E. Survival Analysis of Factors Affecting Culling Early in the Productive Life of Holstein-Friesian Cattle in Kenya. *Livest. Prod. Sci.* **2005**, *92*, 317–322. [[CrossRef](#)]
17. Forabosco, F.; Bozzi, R.; Filippini, F.; Boettcher, P.; Van Arendonk, J.A.M.; Bijma, P. Linear Model vs. Survival Analysis for Genetic Evaluation of Sires for Longevity in Chianina Beef Cattle. *Livest. Sci.* **2006**, *101*, 191–198. [[CrossRef](#)]
18. Neerhof, H.J.; Madsen, P.; Ducrocq, V.P.; Vollema, A.R.; Jensen, J.; Korsgaard, I.R. Relationships between Mastitis and Functional Longevity in Danish Black and White Dairy Cattle Estimated Using Survival Analysis. *J. Dairy Sci.* **2000**, *83*, 1064–1071. [[CrossRef](#)]
19. Jovanovac, S.; Raguž, N. Analysis of the Relationships between Type Traits and Longevity in Croatian Simmental Cattle Using Survival Analysis. *Agric. Conspec. Sci.* **2011**, *76*, 249–253.
20. Heise, J.; Liu, Z.; Stock, K.F.; Rensing, S.; Reinhardt, F.; Simianer, H. The Genetic Structure of Longevity in Dairy Cows. *J. Dairy Sci.* **2016**, *99*, 1253–1265. [[CrossRef](#)]
21. Jankowska, M.; Sawa, A.; Kujawska, J. Effect of Certain Factors on the Longevity and Culling of Cows. *Acta Sci. Pol. Zootech.* **2017**, *13*, 19–30.
22. Zijlstra, J.; Ma, J.; Cao, Z.; van der Fels, J.B. Longevity and Culling Rate: How to Improve? In Proceedings of the SDDDC Progress Report, SDDDC, Beijing, China, 1 June 2016; pp. 1–12.
23. Adamczyk, K.; Makulska, J.; Jagusiak, W.; Węglarz, A. Associations between Strain, Herd Size, Age at First Calving, Culling Reason and Lifetime Performance Characteristics in Holstein-Friesian Cows. *Animal* **2017**, *11*, 327–334. [[CrossRef](#)]
24. De Vries, A.; Marcondes, M.I. Review: Overview of Factors Affecting Productive Lifespan of Dairy Cows. *Animal* **2020**, *14* (Suppl. S1), s155–s164. [[CrossRef](#)] [[PubMed](#)]
25. Schuster, J.C.; Barkema, H.W.; De Vries, A.; Kelton, D.F.; Orsel, K. Invited Review: Academic and Applied Approach to Evaluating Longevity in Dairy Cows. *J. Dairy Sci.* **2020**, *103*, 11008–11024. [[CrossRef](#)] [[PubMed](#)]
26. Shabalina, T.; Yin, T.; König, S. Influence of Common Health Disorders on the Length of Productive Life and Stayability in German Holstein Cows. *J. Dairy Sci.* **2020**, *103*, 583–596. [[CrossRef](#)] [[PubMed](#)]
27. Adamczyk, K.; Jagusiak, W.; Węglarz, A. Associations between the Breeding Values of Holstein-Friesian Bulls and Longevity and Culling Reasons of Their Daughters. *Animal* **2021**, *15*, 100204. [[CrossRef](#)]
28. Vergara, C.F.; Döpfer, D.; Cook, N.B.; Nordlund, K.V.; McArt, J.A.A.; Nydam, D.V.; Oetzel, G.R. Risk Factors for Postpartum Problems in Dairy Cows: Explanatory and Predictive Modeling. *J. Dairy Sci.* **2014**, *97*, 4127–4140. [[CrossRef](#)] [[PubMed](#)]
29. Phuong, H.N.; Blavy, P.; Martin, O.; Schmidely, P.; Friggens, N.C. Modelling Impacts of Performance on the Probability of Reproducing, and Thereby on Productive Lifespan, Allow Prediction of Lifetime Efficiency in Dairy Cows. *Animal* **2016**, *10*, 106–116. [[CrossRef](#)]
30. Van der Heide, E.M.M.; Veerkamp, R.F.; van Pelt, M.L.; Kamphuis, C.; Ducro, B.J. Predicting Survival in Dairy Cattle by Combining Genomic Breeding Values and Phenotypic Information. *J. Dairy Sci.* **2020**, *103*, 556–571. [[CrossRef](#)]
31. Adamczyk, K.; Grzesiak, W.; Zaborski, D. The Use of Artificial Neural Networks and a General Discriminant Analysis for Predicting Culling Reasons in Holstein-Friesian Cows Based on First-Lactation Performance Records. *Animals* **2021**, *11*, 721. [[CrossRef](#)] [[PubMed](#)]
32. Flynn, R. Survival Analysis. *J. Clin. Nurs.* **2012**, *21*, 2789–2797. [[CrossRef](#)]
33. Goel, M.K.; Khanna, P.; Kishore, J. Understanding Survival Analysis: Kaplan-Meier Estimate. *Int. J. Ayurveda Res.* **2010**, *1*, 274. [[CrossRef](#)]
34. Clark, T.G.; Bradburn, M.J.; Love, S.B.; Altman, D.G. Survival Analysis Part I: Basic Concepts and First Analyses. *Br. J. Cancer* **2003**, *89*, 232–238. [[CrossRef](#)]
35. Bradburn, M.J.; Clark, T.G.; Love, S.B.; Altman, D.G. Survival Analysis Part II: Multivariate Data Analysis—an Introduction to Concepts and Methods. *Br. J. Cancer* **2003**, *89*, 431–436. [[CrossRef](#)]
36. Etikan, İ.; Abubakar, S.; Alkassim, R. The Kaplan-Meier Estimate in Survival Analysis. *Biom. Biostat. Int. J.* **2017**, *5*, 00128. [[CrossRef](#)]

37. Schneider, M.D.P.; Strandberg, E.; Ducrocq, V.; Roth, A. Survival Analysis Applied to Genetic Evaluation for Female Fertility in Dairy Cattle. *J. Dairy Sci.* **2005**, *88*, 2253–2259. [[CrossRef](#)]
38. Carlén, E.; Emanuelson, U.; Strandberg, E. Genetic Evaluation of Mastitis in Dairy Cattle Using Linear Models, Threshold Models, and Survival Analysis: A Simulation Study. *J. Dairy Sci.* **2006**, *89*, 4049–4057. [[CrossRef](#)]
39. Kartsonaki, C. Survival Analysis. *Diagn. Histopathol.* **2016**, *22*, 263–270. [[CrossRef](#)]
40. Pohar, M.; Stare, J. Relative Survival Analysis in R. *Comput. Methods Programs Biomed.* **2006**, *81*, 272–278. [[CrossRef](#)] [[PubMed](#)]
41. Stel, V.S.; Dekker, F.W.; Tripepi, G.; Zoccali, C.; Jager, K.J. Survival Analysis I: The Kaplan-Meier Method. *Nephron Clin. Pract.* **2011**, *119*, c83–c88. [[CrossRef](#)] [[PubMed](#)]
42. Carlén, E.; Schneider, M.D.P.; Strandberg, E. Comparison between Linear Models and Survival Analysis for Genetic Evaluation of Clinical Mastitis in Dairy Cattle. *J. Dairy Sci.* **2005**, *88*, 797–803. [[CrossRef](#)]
43. Gilbert, R.O.; Shin, S.T.; Guard, C.L.; Erb, H.N.; Frajblat, M. Prevalence of Endometritis and Its Effects on Reproductive Performance of Dairy Cows. *Theriogenology* **2005**, *64*, 1879–1888. [[CrossRef](#)] [[PubMed](#)]
44. Vukasinovic, N.; Moll, J.; Casanova, L. Implementation of a Routine Genetic Evaluation for Longevity Based on Survival Analysis Techniques in Dairy Cattle Populations in Switzerland. *J. Dairy Sci.* **2001**, *84*, 2073–2080. [[CrossRef](#)]
45. Caraviello, D.Z.; Weigel, K.A.; Shook, G.E.; Ruegg, P.L. Assessment of the Impact of Somatic Cell Count on Functional Longevity in Holstein and Jersey Cattle Using Survival Analysis Methodology. *J. Dairy Sci.* **2005**, *88*, 804–811. [[CrossRef](#)]
46. Clark, T.G.; Bradburn, M.J.; Love, S.B.; Altman, D.G. Survival Analysis Part IV: Further Concepts and Methods in Survival Analysis. *Br. J. Cancer* **2003**, *89*, 781–786. [[CrossRef](#)] [[PubMed](#)]
47. Park, J.H.; Kang, Y.J. Evaluation Index for Sporty Engine Sound Reflecting Evaluators' Tastes, Developed Using k-Means Cluster Analysis. *Int. J. Automot. Technol.* **2020**, *21*, 1379–1389. [[CrossRef](#)]
48. Sinaga, K.P.; Yang, M.-S. Unsupervised K-Means Clustering Algorithm. *IEEE Access* **2020**, *8*, 80716–80727. [[CrossRef](#)]
49. Stel, V.S.; Dekker, F.W.; Tripepi, G.; Zoccali, C.; Jager, K.J. Survival Analysis II: Cox Regression. *Nephron Clin. Pract.* **2011**, *119*, c255–c260. [[CrossRef](#)] [[PubMed](#)]
50. Zhao, L.; Claggett, B.; Tian, L.; Uno, H.; Pfeffer, M.A.; Solomon, S.D.; Trippa, L.; Wei, L.J. On the Restricted Mean Survival Time Curve in Survival Analysis. *Biometrics* **2016**, *72*, 215–221. [[CrossRef](#)] [[PubMed](#)]
51. George, B.; Seals, S.; Aban, I. Survival Analysis and Regression Models. *J. Nucl. Cardiol.* **2014**, *21*, 686–694. [[CrossRef](#)] [[PubMed](#)]
52. Breslow, N. Covariance Analysis of Censored Survival Data. *Biometrics* **1974**, *30*, 89–99. [[CrossRef](#)]
53. Ward, J.H. Hierarchical Grouping to Optimize an Objective Function. *J. Am. Stat. Assoc.* **1963**, *58*, 236–244. [[CrossRef](#)]
54. Majerova, I.; Nevima, J. The Measurement of Human Development Using the Ward Method of Cluster Analysis. *J. Int. Stud.* **2017**, *10*, 239–257. [[CrossRef](#)]
55. Muthahharah, I.; Juhari, A. A Cluster Analysis with Complete Linkage and Ward's Method for Health Service Data in Makassar City. *J. Varian* **2021**, *4*, 109–116. [[CrossRef](#)]
56. Azizi, E.; Kharrati-Shishavan, H.; Mohammadi-Ivatloo, B.; Shotorbani, A.M. Wind Speed Clustering Using Linkage-Ward Method: A Case Study of Khaaf, Iran. *Gazi Univ. J. Sci.* **2019**, *32*, 945–954. [[CrossRef](#)]
57. Chengar, O.V.; Shevchenko, V.I.; Maschenko, E.N.; Moiseev, D.V.; Soina, A.S. Strategy for Primary Processing of Social Networks Data Using Hierarchy Analysis Method. *J. Phys. Conf. Ser.* **2020**, *1679*, 022082. [[CrossRef](#)]
58. Huh, D.; Kim, J.-H.; Jo, I.-H. A Novel Method to Monitoring Changes in Cognitive Load in Video-Based Learning. *J. Comput. Assist. Learn.* **2019**, *35*, 721–730. [[CrossRef](#)]
59. Roszkowska, E.; Filipowicz-Chomko, M. Measuring Sustainable Development Using an Extended Hellwig Method: A Case Study of Education. *Soc. Indic. Res.* **2021**, *153*, 299–322. [[CrossRef](#)]
60. Bártoová, B.; Bína, V. Early Defect Detection Using Clustering Algorithms. *Acta Oecon. Prag.* **2019**, *2019*, 3–20. [[CrossRef](#)]
61. Bienvenido-Huertas, D.; Bermúdez, J.; Moyano, J.J.; Marín, D. Influence of ICHTC Correlations on the Thermal Characterization of Façades Using the Quantitative Internal Infrared Thermography Method. *Build. Environ.* **2019**, *149*, 512–525. [[CrossRef](#)]
62. Eszergár-Kiss, D.; Caesar, B. Definition of User Groups Applying Ward's Method. *Transp. Res. Procedia* **2017**, *22*, 25–34. [[CrossRef](#)]
63. Hossen, M.B.; Siraj-Ud-Doula, M. Identification of Robust Clustering Methods in Gene Expression Data Analysis. *Curr. Bioinform.* **2017**, *12*, 558–562. [[CrossRef](#)]
64. Martha, A.W.; Zain, I. The Clustering of Households in Madura Based on Factors Affecting Their Ingestion of Clean Water Using Similarity Weight and Filter Method. *Inferensi* **2019**, *2*, 45–51. [[CrossRef](#)]
65. Ogasawara, Y.; Kon, M. Two Clustering Methods Based on the Ward's Method and Dendrograms with Interval-Valued Dissimilarities for Interval-Valued Data. *Int. J. Approx. Reason.* **2021**, *129*, 103–121. [[CrossRef](#)]
66. Ridwan, M.R.; Retnawati, H. Application of Cluster Analysis Using Agglomerative Method. *Numer. J. Mat. Dan Pendidik. Mat.* **2021**, *5*, 33–48. [[CrossRef](#)]
67. Nisa, K.; Sari, R.F.; Cipta, H.; Husein, I. Cluster Analysis Using the Hierarki Method for Grouping Sub-Districts in the District Steps Based on Health Indicators. *ZERO: J. Sains Mat. Dan Terap.* **2020**, *4*, 28–35. [[CrossRef](#)]
68. Inan, O.; Uzer, M.S. A Method of Classification Performance Improvement via a Strategy of Clustering-Based Data Elimination Integrated with k-Fold Cross-Validation. *Arab. J. Sci. Eng.* **2021**, *46*, 1199–1212. [[CrossRef](#)]
69. Yang, M.-S.; Sinaga, K.P. A Feature-Reduction Multi-View k-Means Clustering Algorithm. *IEEE Access* **2019**, *7*, 114472–114486. [[CrossRef](#)]

70. Siwi, C.P.; Nurfirdaus, Y. K-Means Cluster Analysis of Sub-Districts in Sidoarjo Based on Long-Term Contraceptive Method. *J. Biom. Dan Kependud.* **2020**, *9*, 161–170. [[CrossRef](#)]
71. Syakur, M.A.; Khotimah, B.K.; Rochman, E.M.S.; Satoto, B.D. Integration K-Means Clustering Method and Elbow Method for Identification of the Best Customer Profile Cluster. *IOP Conf. Ser. Mater. Sci. Eng.* **2018**, *336*, 012017. [[CrossRef](#)]
72. Yuan, C.; Yang, H. Research on K-Value Selection Method of K-Means Clustering Algorithm. *Comput. Sci.* **2019**, *2*, 226–235. [[CrossRef](#)]
73. Jain, A.K. Data Clustering: 50 Years beyond K-Means. *Pattern Recognit. Lett.* **2010**, *31*, 651–666. [[CrossRef](#)]
74. Wang, Q.; Wang, Y.; Niu, R.; Peng, L. Integration of Information Theory, K-Means Cluster Analysis and the Logistic Regression Model for Landslide Susceptibility Mapping in the Three Gorges Area, China. *Remote Sens.* **2017**, *9*, 938. [[CrossRef](#)]
75. Abdulhafedh, A. Incorporating K-Means, Hierarchical Clustering and PCA in Customer Segmentation. *J. City Dev.* **2021**, *3*, 12–30. [[CrossRef](#)]
76. Xie, H.; Zhang, L.; Lim, C.P.; Yu, Y.; Liu, C.; Liu, H.; Walters, J. Improving K-Means Clustering with Enhanced Firefly Algorithms. *Appl. Soft Comput.* **2019**, *84*, 105763. [[CrossRef](#)]
77. Ardiyanti, A.Y.; Notobroto, H.B.; Indawati, R. Risk Factors Clustering of Malnutrition Case Based on the Utilization of Toddler Health Service at Lamongan Regency in 2016. *Int. J. Public Health Clin. Sci.* **2018**, *5*, 231–238.
78. Dini, S.K.; Fauzan, A. Clustering Provinces in Indonesia Based on Community Welfare Indicators. *EKSAKTA J. Sci. Data Anal.* **2020**, *20*, 56–63. [[CrossRef](#)]
79. Ohana-Levi, N.; Bahat, I.; Peeters, A.; Shtein, A.; Netzer, Y.; Cohen, Y.; Ben-Gal, A. A Weighted Multivariate Spatial Clustering Model to Determine Irrigation Management Zones. *Comput. Electron. Agric.* **2019**, *162*, 719–731. [[CrossRef](#)]
80. Pedersen, J.; Jensen, J. Evaluation of Female Fertility of Danish Dairy Sires. *Interbull Bull.* **1996**, *12*, 72–77.
81. Seegers, H.; Beaudeau, F.; Fourichon, C.; Bareille, N. Reason for Culling in French Holstein Cows. *Prev. Vet. Med.* **1998**, *36*, 257–271. [[CrossRef](#)]
82. Bascom, S.S.; Young, A.J. A Summary of the Reasons Why Farmers Cull Cows. *J. Dairy Sci.* **1998**, *81*, 2299–2305. [[CrossRef](#)]
83. Krausslich, H. Grenzen Der Leistungs Steigerung-Hat Des Zweintutzungs System Eine Zuknuft. *Zucht. Und Besam.* **2003**, *149*, 26–29.
84. Fetrow, J.; Nordlund, K.V.; Norman, H.D. Invited Review: Culling: Nomenclature, Definitions, and Recommendations. *J. Dairy Sci.* **2006**, *89*, 1896–1905. [[CrossRef](#)]
85. Reimus, K.; Alvåsen, K.; Emanuelson, U.; Viltrop, A.; Mötus, K. Herd-Level Risk Factors for Cow and Calf on-Farm Mortality in Estonian Dairy Herds. *Acta Vet. Scand.* **2020**, *62*, 15. [[CrossRef](#)] [[PubMed](#)]
86. Sewalem, A.; Kistemaker, G.J.; Miglior, F.; Van Doormaal, B.J. Analysis of the Relationship between Type Traits and Functional Survival in Canadian Holsteins Using a Weibull Proportional Hazards Model. *J. Dairy Sci.* **2004**, *87*, 3938–3946. [[CrossRef](#)]
87. Zavadilová, L.; Němcová, E.; Štípková, M. Effect of Type Traits on Functional Longevity of Czech Holstein Cows Estimated from a Cox Proportional Hazards Model. *J. Dairy Sci.* **2011**, *94*, 4090–4099. [[CrossRef](#)] [[PubMed](#)]
88. Milian-Suazo, F.; Erb, H.N.; Smith, R.D. Descriptive Epidemiology of Culling in Dairy Cows from 34 Herds in New York State. *Prev. Vet. Med.* **1988**, *6*, 243–251. [[CrossRef](#)]
89. Barkema, H.W.; Westrik, J.D.; Van Keulen, K.A.S.; Schukken, Y.H.; Brand, A. The Effects of Lameness on Reproductive Performance, Milk Production and Culling in Dutch Dairy Farms. *Prev. Vet. Med.* **1994**, *20*, 249–259. [[CrossRef](#)]
90. Gröhn, Y.T.; Eicker, S.W.; Ducrocq, V.; Hertl, J.A. Effect of Diseases on the Culling of Holstein Dairy Cows in New York State. *J. Dairy Sci.* **1998**, *81*, 966–978. [[CrossRef](#)]
91. Kugonza, D.R.; Nabasirye, M.; Hanotte, O.; Mpairwe, D.; Okeyo, A.M. Pastoralists' Indigenous Selection Criteria and Other Breeding Practices of the Long-Horned Ankole Cattle in Uganda. *Trop. Anim. Health Prod.* **2012**, *44*, 557–565. [[CrossRef](#)]
92. Ahlman, T.; Berglund, B.; Rydhmer, L.; Strandberg, E. Culling Reasons in Organic and Conventional Dairy Herds and Genotype by Environment Interaction for Longevity. *J. Dairy Sci.* **2011**, *94*, 1568–1575. [[CrossRef](#)] [[PubMed](#)]
93. Van Arendonk, J.A.M. Studies on the Replacement Policies in Dairy Cattle. II. Optimum Policy and Influence of Changes in Production and Prices. *Livest. Prod. Sci.* **1985**, *13*, 101–121. [[CrossRef](#)]
94. Van Arendonk, J.A.M.; Dijkhuizen, A.A. Studies on the Replacement Policies in Dairy Cattle. III. Influence of Variation in Reproduction and Production. *Livest. Prod. Sci.* **1985**, *13*, 333–349. [[CrossRef](#)]
95. Ansari-Lari, M.; Mohebbi-Fani, M.; Rowshan-Ghasrodashti, A. Causes of Culling in Dairy Cows and Its Relation to Age at Culling and Interval from Calving in Shiraz, Southern Iran. *Vet. Res. Forum* **2012**, *3*, 233–237.
96. Booth, C.J.; Warnick, L.D.; Gröhn, Y.T.; Maizon, D.O.; Guard, C.L.; Janssen, D. Effect of Lameness on Culling in Dairy Cows. *J. Dairy Sci.* **2004**, *87*, 4115–4122. [[CrossRef](#)]
97. Fodor, I.; Lang, Z.; Ózsvári, L. Relationship of Dairy Heifer Reproduction with Survival to First Calving, Milk Yield and Culling Risk in the First Lactation. *Asian-Australas. J. Anim. Sci.* **2020**, *33*, 1360. [[CrossRef](#)] [[PubMed](#)]
98. Whitaker, D.A.; Kelly, J.M.; Smith, S. Disposal and Disease Rates in 340 British Dairy Herds. *Vet. Rec.* **2000**, *146*, 363–367. [[CrossRef](#)] [[PubMed](#)]
99. Heinrichs, A.J.; Zanton, G.I.; Lascano, G.J.; Jones, C.M. A 100-Year Review: A Century of Dairy Heifer Research. *J. Dairy Sci.* **2017**, *100*, 10173–10188. [[CrossRef](#)] [[PubMed](#)]
100. Bradley, A.J.; Green, M.J. Aetiology of Clinical Mastitis in Six Somerset Dairy Herds. *Vet. Rec.* **2001**, *148*, 683–686. [[CrossRef](#)]

101. Čanji, V.; Strapák, P.; Strapáková, E.; Juhás, P. Effect of Conformation Traits on Longevity of Cows of Slovak Simmental Breed. *Slovak J. Anim. Sci.* **2008**, *41*, 83–90.
102. Strapák, P.; Juhás, P.; Strapáková, E. The Relationship between the Length of Productive Life and the Body Conformation Traits in Cows. *J. Cent. Eur. Agric.* **2011**, *12*, 239–254. [[CrossRef](#)]
103. Strapák, P.; Juhás, P.; Strapáková, E.; Halo, M. Relation of the Length of Productive Life and the Body Conformation Traits in Slovak Simmental Breed. *Arch. Anim. Breed.* **2010**, *53*, 393–402. [[CrossRef](#)]
104. Strapáková, E.; Strapák, P.; Candrák, J.; Pavlík, I.; Dočkalová, K. Fleckscore System of Exterior Evaluation as a More Accurate Indirect Predictor of Longevity in Slovak Simmental Dairy Cows. *Czech J. Anim. Sci.* **2021**, *66*, 487–494. [[CrossRef](#)]
105. Strapák, P.; Juhás, P.; Čanji, V. The Relationship between the Body Conformation to Longevity in the Slovak Spotted Cows. *Biotechnol. Anim. Husb.* **2005**, *21*, 53–59. [[CrossRef](#)]
106. Sawa, A.; Bogucki, M.; Krežel-Czopek, S.; Neja, W. Relationship between Conformation Traits and Lifetime Production Efficiency of Cows. *Int. Sch. Res. Not.* **2013**, *2013*, 124690. [[CrossRef](#)]
107. Vacek, M.; Stipkova, M.; Nemcová, E.; Bouska, J. Relationships between Conformation Traits and Longevity of Holstein Cows in the Czech Republic. *Czech J. Anim. Sci.* **2006**, *51*, 327. [[CrossRef](#)]
108. Cramer, G.; Lissemore, K.D.; Guard, C.L.; Leslie, K.E.; Kelton, D.F. The Association between Foot Lesions and Culling Risk in Ontario Holstein Cows. *J. Dairy Sci.* **2009**, *92*, 2572–2579. [[CrossRef](#)] [[PubMed](#)]
109. Morek-Kopec, M.; Zarnecki, A.; Ptak, E.; Otwinowska-Mindur, A. Effect of Calving Difficulties and Calf Mortality on Functional Longevity in Polish Holstein-Friesian Cows. *Animals* **2021**, *11*, 2792. [[CrossRef](#)]
110. Sasaki, Y.; Uematsu, M.; Kitahara, G.; Osawa, T.; Sueyoshi, M. Effects of Stillbirth and Dystocia on Subsequent Reproductive Performance in Japanese Black Cattle. *Vet. J.* **2014**, *200*, 462–463. [[CrossRef](#)] [[PubMed](#)]
111. Chirinos, Z.; Carabaño, M.J.; Hernández, D. Genetic Evaluation of Length of Productive Life in the Spanish Holstein-Friesian Population. Model Validation and Genetic Parameters Estimation. *Livest. Sci.* **2007**, *106*, 120–131. [[CrossRef](#)]