LONG LIVE THE COW

USING MACHINE LEARNING TO ANALYZE WELFARE, LONGEVITY, PRODUCTIVITY, AND PROFITABILITY OF DAIRY COWS

BY

GABRIEL MACHADO DALLAGO

Department of Animal Science

McGill University, Montreal

December 2021

A thesis submitted to the McGill University in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

TABLE OF CONTENTS

ABSTRACT

Promoting the sustainability of dairy farming is imperative not only to meet the projected increase in demand for dairy, but also because it holds an important role in food security. Cow longevity and animal welfare are important factors associated with the sustainability of the dairy industry. Higher longevity, achieved by avoiding culling due to health issues such as mastitis, feet, and legs problems, as well as failure to reproduce (i.e., involuntary culling), could be associated with higher economic performance of farms, a lower environmental footprint of the milk industry, and a better welfare status of the animals. Yet, cow longevity has decreased in most high milk-producing countries over the last decades and its relationship with milk yield is not straightforward. While strategic culling decisions aiming at keeping cows more likely to avoid involuntary culling could help in increasing longevity among herds, the current metrics of longevity are limited to either the length of time a cow remains in the herd or if she is still alive at a given time, limiting such strategic decision-making. Early life indicators such as calving ease, birth size, and twinning could be used to help in the decision process since they are associated with a reduction in the longevity of the offspring. Other data such as the occurrence of health events during the pre-weaning period as well as failure of the passive immunity transfer do not seem to be considered by farmers to carry out culling decisions during the rearing period. Early life management practices regarding colostrum, feeding, and housing were also studied, and two distinct clusters of farms were identified: production-oriented farms used more modern management practices which were associated with increased productivity and profitability but reduced longevity, whereas resource-oriented farms used more traditional practices which in turn were associated with increased longevity but reduced productivity and profitability. Lastly, different animal welfare profiles were identified when investigating the link between herd outcome measures of welfare and both longevity and economic indicators, highlighting that management practices should be adjusted to improve animal outcomes in those with poorer welfare status. However, the set of solutions and innovations should be tailored according to the individual problems present on each farm. The pressing issues associated with climate change and competition for resources highlight the need to improve the efficiency of animal food production, but both farmers' and animals' interests should be aligned to meet the increasing demand while ensuring the sustainability of the sector.

RÉSUMÉ

Promouvoir la durabilité de la production laitière est un impératif, non seulement pour répondre à l'augmentation prévue de la demande pour les produits laitiers, mais également parce qu'elle joue un rôle important dans ce qui a trait à la sécurité alimentaire; à cet égard, la longévité et le bien-être des vaches laitières sont autant de facteurs importants associés à la durabilité de l'industrie laitière. L'atteinte d'une meilleure longévité via une diminution des réformes involontaires, souvent liées à des problèmes de santé tels que la mammite, les affections des pieds et membres et les trouble de reproduction, pourrait mener à une meilleure performance économique des fermes, une diminution de l'empreinte environnementale de la production, et un meilleur statut de bien-être des vaches. Pourtant, au cours des dernières décennies, la longévité des vaches a diminué dans la majorité des pays producteurs de lait, si bien que le lien unissant longévité et productivité ne semble pas aussi clair qu'il pourrait l'être. Si des décisions de réforme stratégiques visant à garder dans les troupeaux des vaches plus susceptibles d'éviter les troubles menant à une réforme involontaire pourraient aider à améliorer la longévité des troupeaux, les données actuelles liées à la longévité sont limitées à la durée de temps passée par la vache dans le troupeau, ou à la question de sa survie et de sa présence dans le troupeau au moment de l'évaluation, ce qui limite les possibilités pour la prise de décisions stratégiques. Puisqu'ils sont associés à une réduction de la longévité des vaches, des indicateurs notés tôt dans la vie (ex. : la facilité du vêlage, la taille à la naissance et la naissance gémellaire) pourraient servir pour la prise de décisions. D'autres données telles que la survenue de problèmes de santé au cours de la période pré-sevrage ou l'échec du transfert d'immunité passive ne sont que peu considérés par les producteurs lors de la prise de décisions de réforme au courant de la période d'élevage des femelles de remplacement. Lors de l'étude des pratiques de gestion en début de vie à la ferme (ex. : gestion du colostrum, logement des veaux), deux groupes de producteurs ont été identifiés : des producteurs dont la gestion est plus axée sur la productivité, qui utilisent des pratiques modernes de gestion des veaux, et qui obtiennent une plus grande productivité des vaches en lactation et une plus grande valeur de lait produit, mais une moins grande longévité, et des producteurs dont la gestion est plus axée sur les ressources, qui utilisent des pratiques plus traditionnelles de gestion des veaux, et qui ont une meilleure longévité du troupeau, mais des niveaux de production et une valeur de lait produit moins élevés. Différents profils en fait de bien-être animal ont été identifiés lors de l'investigation du lien entre les mesures de bien-être à la ferme et les mesures de longévité ainsi que les indicateurs

économiques; ces trouvailles démontrent que les pratiques de gestion de troupeau devraient être ajustées pour améliorer le bien-être des vaches chez les troupeaux où cet aspect se révèle plus problématique. Toutefois, il n'existe pas d'ensemble unique de solutions applicable à toutes les fermes; il serait donc important de prioriser une approche adaptée aux problèmes spécifiques identifiés sur chacune des fermes. Le contexte lié aux changements climatiques et à la compétition pour les ressources met en lumière un besoin d'améliorer davantage l'efficacité des productions animales incluant la production laitière, d'une manière qui respectera à la fois les intérêts des producteurs et les besoins des animaux, afin d'augmenter la productivité du secteur et ainsi répondre à la demande en augmentation, mais de manière durable.

ACKNOWLEDGMENTS

Alright, alright! Here we are! What a journey has it been. Though the acknowledgements section usually comes at the beginning of a thesis, it is probably something that is written last. It is for me at least. We do not usually stop to think about where we are and how we got there. So, I would like to take this opportunity to reflect on the road taken up to this point. At some point in our lives (hopefully) all the floating pieces fall into place, and everything suddenly makes sense. This is not what happened during my Ph.D. It might have been just quite the opposite. A "famous" quote probably describes it best: "*I knew exactly what to do, but in a much more real sense, I had no idea what to do*" (Michael Scott; The Office, S5E14 – Stress Relief Part 1, 9:11). We have all been there, haven't we? There have been ups and downs through out this journey, but I surely did not do it alone. Rather, I could not have done it alone. This thesis is the result of a collaborative effort, and this is my humble attempt to thank the people involved.

First and foremost, I would like to thank my supervisor, Dr. Elsa Vasseur. It may come across as a formality to thank your supervisor, but this could not be more sincere. Thank you, Elsa, for always believing in me over the past three years and for providing me with all the support, incentive, and wake-up calls I needed. You have always been a bright, kind, and optimistic light regardless of the situation.

I am grateful I had Drs. Kevin M. Wade, Roger I. Cue, and J T. McClure as my committee members. Thank you, Dr. Wade and Dr. Cue for always helping me with the precarious data-related troubles and for providing insightful suggestions. I would like to thank Dr. McClure for his supervision and support during my research time at the University of Prince Edward Island (UPEI). Thanks, as well to Drs. Gregory P. Keefe and Ibrahim Elsohaby for data provision and the work done with me not only while I was at the UPEI, but afterwards as well.

I truly appreciate all the funding support I received, which made it possible for me to carry out my Ph.D. research. I would like to thank the funding support provided by Novalait, Dairy Farmers of Canada, Lactanet, and Natural Sciences and Engineering Research Council (NSERC) as part of Industrial Research Chair in the Sustainable Life of Dairy Cattle and Agriculture and Agri-Food Canada and Dairy Farmers of Canada through AgriScience Clusters program, Growing Forward 3. My extended gratitude to the program CREATE in Milk Quality of NSERC, the Op+lait group of the Fonds de Recherche du Québec – Nature et Technologies (FRQNT), McGill

University, through the Pilarczyk Fellowship and Graduate Excellence Awards for providing stipend funding as well as McGill University for the GREAT Travel Award and the FRQNT for providing a scholarship.

I would like to acknowledge both Lactanet and Dairy Farmers of Canada for providing Dairy Herd Improvement (DHI) and animal welfare outcomes data, respectively. Most of the studies in my Ph.D. would not have been possible without these data. My sincere gratitude!

I would like to express my appreciation to everyone else with whom I shared my Ph.D. journey with. Thank you to all my fellow CowLife McGill lab mates for extended lunch breaks, coffee breaks, and breaks in general. You all made the process lighter and more pleasant when it seemed to have reached a *cul-de-sac* point. My sincere gratitude to all McGill University – Macdonald Campus Staff. Always working behind the curtains and providing the very needed backbone so we can do what we do (best)! I'm also thankful for the reminders throughout the entire process to always Live, Laugh, and Love! Lastly, but not least, I would like to thank my parents Natália M. de Almeida and Celso Dallago as well as my brother Bruno M. Dallago. Though far away, I have always received their unconditional support.

There might not be a point in life where all floating pieces fall into place because there is no such thing as right place. Being able to navigate through the floating pieces is a privilege that many wishes they had, and it takes us a lot of work to appreciate it. Specially during tough times where we are tempted to doubt ourselves. I am happy I did not do it alone and grateful for all the people whose paths were intertwined with my own. It is always the beginning of something new.

Hope you enjoy the reading :)

CONTRIBUTION TO KNOWLEDGE

CHAPTER 2

Raising concerns about the state of dairy cow longevity have been voiced by the industry, warranting a comprehensive review of the literature. This chapter reviewed the current longevity metrics available and provided both an original and standardized approach to evaluate cow longevity at the country level. By doing so, we were able to not only depict the current state of cow longevity, but also how it has changed over the last decades. We discussed potential influencing factors and stated specific factors that should be considered to improve the definition of longevity.

CHAPTER 3

This chapter provided an evaluation of birth conditions routinely collected on an animal-level by Dairy Herd Improvement agencies and their effect on offspring longevity. We found that factors commonly associated with a difficult calf delivery impaired offspring longevity, providing evidence of long-term effects of early life conditions. Such knowledge contributes to identifying animals less likely to achieve their potential and to carrying out strategic culling decisions at an early stage.

CHAPTER 4

The relationship between early life animal-level outcomes and lifetime indicators of longevity, productivity, and profitability were analyzed. By analyzing the data from a cohort of animals made available through a collaboration with the University of Prince Edward Island, we found that problems in the early life such as the occurrence of diseases during the pre-weaning period are not necessarily linked with a reduction in animal longevity, productivity, and profitability. Herd was also shown to be the most important variable associated with the lifetime indicators, highlighting the importance of management decisions towards cow longevity, productivity, and profitability.

CHAPTER 5

Chapter 5 furthers our understanding of the relationships between early life and outcomes of longevity, productivity, and profitability, this time at herd-level and based on a greater number of farms compared to Chapter 4. We discovered the existence of stable and well-separated groups of farms that adopted different sets of early life management practices. In turn, clear patterns were observed between these management practices and the evaluated herd outcomes. This knowledge provides the groundwork towards developing intervention strategies to improve cow longevity.

CHAPTER 6

Dairy cows are the backbone of the dairy industry. Therefore, to map the sustainability of dairy farming, we used herd-level animal outcomes of welfare, production, and performance as indicators of social, economic, and environmental sustainability pillars respectively. We uncovered and characterized different welfare profiles, highlighting how animals respond to their surroundings, and the need for flexible solutions to improve the social acceptability of dairy farming. We also found a pressing need for better indicators of social, economic, and environmental profiles to fully assess the sustainability of dairy farming.

CONTRIBUTION OF AUTHORS

Five co-authored manuscripts are presented in this thesis.

Authors of manuscripts 1, 2, 3, 4 and 5 (Chapters 2, 3, 4, 5, and 6, respectively) are as follows:

Gabriel M. Dallago (primary author, manuscripts 1-5), Elsa Vasseur (manuscripts 1-5), René Lacroix (manuscripts 1, 2, 3, and 5), Kevin M. Wade (manuscripts 1, 2, and 5), Roger I. Cue (manuscripts 1, 2, and 5), J T. McClure (manuscripts 1 and 3), Ibrahim Elsohaby (manuscript 3), Doris Pellerin (manuscript 1), Daniel Warner (manuscript 5), and Abdoulaye B. Diallo (manuscript 5).

Gabriel M. Dallago conceptualized manuscript 2, co-conceptualized manuscripts 1, 3, 4, and 5, and designed and conducted the research studies which resulted in the findings presented in manuscripts 1-5. Additionally, Gabriel analyzed the data for manuscripts 1-5, conceptualized the presentation of results and was the primary author on all five manuscripts.

Elsa Vasseur supervised the primary author, co-conceptualized the manuscripts 1, 3, 4, and 5, and assisted in the elaboration of all study designs. Elsa also reviewed and co-authored all five manuscripts.

René Lacroix oversaw the extraction of longevity, productivity, and profitability data used on manuscript 3. René also reviewed and co-authored manuscripts 1, 2, 3, and 5.

Kevin M. Wade reviewed and co-authored manuscripts 1, 2, and 5.

Roger I. Cue provided the data used on manuscript 2, and reviewed and co-authored manuscripts 1, 2, and 5.

J T. McClure and Ibrahim Elsohaby provided the data on early life animal outcomes, reviewed, and co-authored manuscript 3. J T. McClure also reviewed and co-authored manuscript 1.

Doris Pellerin reviewed and co-authored manuscript 1.

Daniel Warner extracted the data, reviewed, and co-authored manuscript 5.

Abdoulaye B. Diallo co-conceptualized the data analysis, reviewed, and co-authored manuscript 5.

LIST OF FIGURES

LIST OF TABLES

LIST OF SUPPLEMENTARY MATERIAL

CHAPTER 1 – General introduction

In simple terms, dairy cow longevity refers to how long a cow lives. Though it may refer to a simple concept, its ramifications under current commercial conditions are far from simple. The life of a dairy cow can be split into non-productive and productive stages. The non-productive stage of a female dairy offspring starts when her dam gets pregnant. Once born, the female calf receives a liquid diet up until she is weaned off at about 2 months of age (Vasseur et al., 2010). The calf grows into a heifer, which is then artificially inseminated at about 18 months old and calves for the first time at 27 months old (Lactanet, 2020). Calving for the first time marks the end of the non-productive stage and the beginning of the productive stage. That is when the heifer, now a cow, starts to produce milk and enters the lactating herd. The cow will produce milk for about a year before she is dried prior to giving birth to another calf (Lactanet, 2020). The lactating-dry-lactating cycle repeats until she is culled from the herd.

Culling is the process of removing an animal from the herd due to death, salvage, sale, or slaughtering (Fetrow et al., 2006). It is a management decision commonly carried out by farmers and the most frequently reported reasons are fertility issues, mastitis, leg problems, and low milk yield (Heise et al., 2016, Compton et al., 2017, CDIC, 2021). Regardless of the reason, all cows are eventually culled. Though they can live for 20 years (De Vries and Marcondes, 2020), culling often happens at a younger age (Dallago et al., 2021) depending on farmers' production priorities (Rilanto et al., 2022). Instead of culling, farmers may decide to try different reproduction technologies on cows that have fertility issues and treat the sick ones. Looking at dairy farming as a business enterprise, such decision should be based on maximizing profitability. However, cumulative costs and profitability of a dairy herd are generally not known or underestimated (Vasseur et al., 2012; Duplessis et al., 2021), whereas the past occurrence of costly events such as health problems are seldom considered (Beaudeau et al., 2000).

Removing a cow too early from the herd is not without consequences. Depending on cow productive levels, she only becomes profitable from her third lactation onwards, mainly because of her rearing costs (Horn et al., 2012, Delgado et al., 2017, Habel et al., 2021). Cows with longer longevity are also more environmentally sustainable. Methane emission does not increase as cows get older (Grandl et al., 2016) and the longer a cow remains in the herd, the lower the methane emission per kg of milk produced (Grandl et al., 2019). Lastly, animal welfare is the primary issue mentioned by consumers thinking towards an ideal dairy farms (Cardoso et al., 2016) and most of the reported reasons for culling are associated with diseases, implying a poor animal welfare status. Therefore, increasing dairy cow longevity would be beneficial for the dairy industry not only from a consumer's perception, but also from the perspective of animal welfare, while making it more profitable for dairy farmers and more environmentally sustainable.

Reducing the occurrence of both reproduction failures and health issues could potentially increase dairy cow longevity as they are the most prevalent reported reasons for culling. In Canada, for example, 72.0% of the culling in 2020 were due to reasons other than low milk production (CDIC, 2021). One possible strategy to achieve that reduction would be by minimizing the risk factors for various undesirable health events. For one, concrete floors (Somers et al., 2003, Vanegas et al., 2006) and bed comfort (McPherson and Vasseur, 2020) are some of the risk factors associated with the occurrence of lameness, thus acting on those factors could contribute. Another strategy to employ would be increasing the cleaning frequency of dairy barns since it improves the hygiene scores for the udder (DeVries et al., 2012), which contributes to a decrease in somatic cell count (Schreiner and Ruegg, 2003, Reneau et al., 2005) and a reduction of clinical mastitis (Santman-Berends et al., 2016).

Another possibility is to identify and raise resilient animals, which exhibit high adaptability to challenges and good cumulative fertility and health, resulting in an increased longevity (Adriaens et al., 2020). As opposed to controlling risk factors, much less attention has been drawn to the early identification of resilient animals that are more likely to reach their potential. An example of this is the occurrence of diarrhea during rearing, which impacts the subsequent milk production (Svensson and Hultgren, 2008) and reproductive performance (Aghakeshmiri et al., 2017). Therefore, efforts to improve cow longevity should be made even before a dairy animal enters the lactating herd, maximizing the efficiency in which resources are used to produce milk with inherited sustainability.

Understanding the connections between early life conditions and animal longevity, productivity, and performance is fundamental to move towards improving the dairy industry sustainability. It provides new insights to develop recommendations for dairy producers to improve their ability to keep healthier and comfortable cows longer in their herds. Overall, such an approach would contribute to a gain in economic and environmental sustainability of dairy farms, while leading to improvements in animal welfare and creating a more positive public perception of the dairy industry.

1.1. General and specific objectives

We have hypothesized that early life aspects of dairy cattle influence their subsequent longevity, productivity, and profitability. The general objective of this thesis was to test this hypothesis from different perspectives. A final objective was to provide insights on the sustainability of dairy farming as it is currently practiced based on animal welfare outcome measures as well as on longevity, productivity, and profitability. Chapter 2 aimed to provide an integrated review of dairy cow longevity by reviewing the most common longevity metrics and describing the current state of longevity in high milk-producing countries. Chapters 3 to 5 delve deeper into analyzing different early life aspects and their association with longevity, productivity, and performance. Both Chapters 3 and 4 used animal-level responses, while Chapter 5 used herd-level response metrics. Chapter 3 aimed to evaluate the effect of calving ease, calf size, and twinning, which are birth conditions routinely collected by Dairy Herd Improvement agencies, on subsequent offspring longevity. The objective of Chapter 4 was to analyze the associations between early life animal outcomes of body weight, health events, and immune status of dairy calves and not only their subsequent longevity, but also their productivity and profitability. In Chapter 5, the objective was to characterize dairy herds based on a set of diverse early life management practices and to analyze the association between practices and herd longevity, productivity, and profitability. Lastly, the objective of the study presented in Chapter 6 was to map the sustainability of dairy farming using an animal centric approach based on herd-level indicators of animal welfare outcomes, longevity, productivity, and profitability.

CHAPTER 2 – Keeping dairy cows for longer: A critical literature review on dairy cow longevity in high milk-producing countries

Gabriel M. Dallago[1,*], Kevin M. Wade[1], Roger I. Cue[1], J T. McClure[2], René Lacroix[3], Doris Pellerin[4] and Elsa Vasseur[1]

[1] Department of Animal Science, McGill University, Sainte-Anne-de-Bellevue, Quebec, H9X 3V9, Canada; gabriel.dallago@mail.mcgill.ca (G.M.D.); kevin.wade@mcgill.ca (K.M.W.); roger.cue@mcgill.ca (R.I.C.); elsa.vasseur@mcgill.ca (E.V.)

[2] Department of Health Management, Atlantic Veterinary College, University of Prince Edward Island, Charlottetown, Prince Edward Island, C1A 4P3, Canada; jmcclure@upei.ca

[3] Lactanet, Sainte-Anne-de-Bellevue, Quebec, H9X 3R4, Canada; rlacroix@lactanet.ca

[4] Département des Sciences Animales, Université Laval, Québec, Québec, G1V 0A6, Canada; doris.pellerin@fsaa.ulaval.ca

* Corresponding author: gabriel.dallago@mail.mcgill.ca

Received: 15 January 2021; Accepted: 9 March 2021; Published: 13 Mach 2021

Animals 11(3): 808. https://doi.org/10.3390/ani11030808.

2.1. Abstract

The ability of dairy farmers to keep their cows for longer could positively enhance the economic performance of the farms, reduce the environmental footprint of the milk industry, and overall help in justifying a sustainable use of animals for food production. However, there is little published on the current status of cow longevity and we hypothesized that a reason may be a lack of standardization and an over narrow focus of the longevity measure itself. The objectives of this critical literature review were: (1) to review metrics used to measure dairy cow longevity; (2) to describe the status of longevity in high milk-producing countries. Current metrics are limited to either the length of time the animal remains in the herd or if it is alive at a given time. To overcome such a limitation, dairy cow longevity should be defined as an animal having an early age at first calving and a long productive life spent in profitable milk production. Combining age at first calving, length of productive life, and margin over all costs would provide a more comprehensive evaluation of longevity by covering both early life conditions and the length of time the animal

remains in the herd once it starts to contribute to the farm revenues, as well as the overall animal health and quality of life. This review confirms that dairy cow longevity has decreased in most high milk-producing countries over time and its relationship with milk yield is not straight forward. Increasing cow longevity by reducing involuntary culling would cut health costs, increase cow lifetime profitability, improve animal welfare, and could contribute towards a more sustainable dairy industry while optimizing dairy farmers' efficiency in the overall use of resources available.

2.2. Introduction

Dairy cow longevity is the length of life of the animal, which in turn is determined by either culling decision made by the producer or death of the animal. The removal of cows from dairy herds due to old age is rare in the modern dairy industry and the economic interest associated with farm animals, which require them to achieve expected production levels, to reproduce regularly, and stay healthy (Essl, 1998, Fetrow et al., 2006), influences the farmer's decision regarding the optimum moment to cull a cow. It is a complex decision process and a myriad of factors are to be considered by the dairy farmer (Roche et al., 2020). Therefore, longevity is a compound feature reflecting a successful combination of many different aspects during the lifespan of a cow (Van Doormaal, 2009). Dairy cow longevity is linked to the economic performance of farms, the environmental footprint of the milk industry, and the welfare status of the animals (Essl, 1998, Benbrook et al., 2010, Brickell and Wathes, 2011, Pellerin et al., 2014, Boulton et al., 2017, Grandl et al., 2019), and short cow longevity limits the achievement of a sustainable dairy industry.

The genetic potential for longevity has increased over the years (De Vries, 2017, CRV, 2020, DairyNZ, 2020) reflecting the inclusion of functional traits in the calculation of estimated breeding values (Van Doormaal, 2009). Even though dairy cow have a life expectancy of around 20 years (De Vries and Marcondes, 2020), this is rarely observed under modern commercial conditions. In Canada for example, the average age that Holstein cows die due to natural causes is 9.1 years (Van Doormaal, 2009). This would represent a productive life (length of time between first calving and culling/death) of 6.8 years or about 6 lactations if an average age at first calving of 27 months is assumed (Van Doormaal, 2009). This has been appointed as a problem by the dairy industry and, contrary to milk recording, there is no standardized approach to measure longevity which results in different metrics being used by different countries (Mark, 2004). Increasing dairy cow longevity could be a strategy to improve the efficiency of using resources available to the dairy farmer and to produce milk with inherited sustainability.

The purpose of this critical literature review is to provide an integrated view of dairy cow longevity combined with the analysis of its status by focusing on phenotypical aspects of longevity rather than its genetic aspects. The objectives were to 1) review metrics commonly used to measure dairy cow longevity and, 2) use the most common metric to describe the status of longevity in high milk-producing countries. We hypothesized that limitations exist on current longevity metrics such as the lack of both a standard metric and reporting by DHI agencies or national databases. The significance of this critical review is to overcome these limitations by developing a standard methodology to estimate longevity metrics, which allow for a fair comparison between different countries, and by demonstrating that dairy cow longevity has decreased over the years in most high milk-producing countries. Addressing these two objectives will then lead us to answering the following questions: (i) should we improve dairy cow longevity?, and (ii) how can we improve dairy cow longevity?, and result in (iii) proposing a more comprehensive definition of cow longevity.

2.3. How can we measure longevity?

The longevity of dairy cows is influenced by culling decisions made by the dairy farmer since culling ultimately defines the total length of time a cow remains in the herd. Therefore, common longevity metrics reflect culling strategies as well as the different stages of the life of a dairy cow.

2.3.1. Culling

Culling is the process of removing an animal from the herd due to death, salvage, sale, or slaughtering (Fetrow et al., 2006). Apart from death, culling is a decision usually made by the dairy farmer and it is influenced by the economic interest associated with farm animals. Culling can be further classified as voluntary or involuntary based on the main reason underlying the culling decision. Voluntary culling occurs when a fertile and healthy animal is culled due to low milk production (Weigel et al., 2003, Fetrow et al., 2006). On the other hand, an involuntary culling happens if low milk production is not the culling reason (Weigel et al., 2003, Fetrow et al., 2006).

Involuntary culling accounts for most of the removal of dairy cows with known reasons. For example in Canada, the average involuntary culling was 73.6% (Standard deviation; SD = 0.65) between 2014 and 2019, while the averages of voluntary culling and culling with unknown reason were 7.18% (SD = 0.28) and 20.7% (SD = 2.3), respectively, in the same period (CDIC, 2020a). The high percentage of culling with unknown reason indicates the existence of limitations among producers to keep track of culling records within the farm and reporting it. Reproduction (16.8%; SD = 0.51), mastitis (10.6%; SD = 0.66), and feet and leg problems (6.88%; SD = 0.33) were the main reasons for involuntary culling during this period (CDIC, 2020a). Similarly, infertility (20.4%), udder health (14.7%), and leg disorders (12.2%) have been reported as the main reasons for involuntary culling in Germany between 2010 and 2013 (Heise et al., 2016). The prevalence of the main culling reasons remained stable over time. In a meta-analysis conducted by Compton et al. (2017) on 51 published papers regarding 54 studies conducted in 22 different countries between 1989 and 2014, the annual incidence risk of culling due to udder and reproduction issues did not change for almost two decades starting at the mid-1980s. At the same time, there was a decrease in culling due to low milk production (voluntary culling). A similar condition was observed in Canada (Figure 2.1), in which the percentage of involuntary culling remained stable between 1997 and 2019 (CDIC, 2020a) for reproduction, mastitis, and feet and leg problems, while the culling for low milk production decreased up to 2008 after which it was seen a slight upwards trend. The main reason for such reduction in the voluntary culling is likely due to the genetic selection for high milk-yielding cows, which reduces the relative risk of being culled due to low milk production and is likely to continue as an objective of dairy farms (Compton et al., 2017).

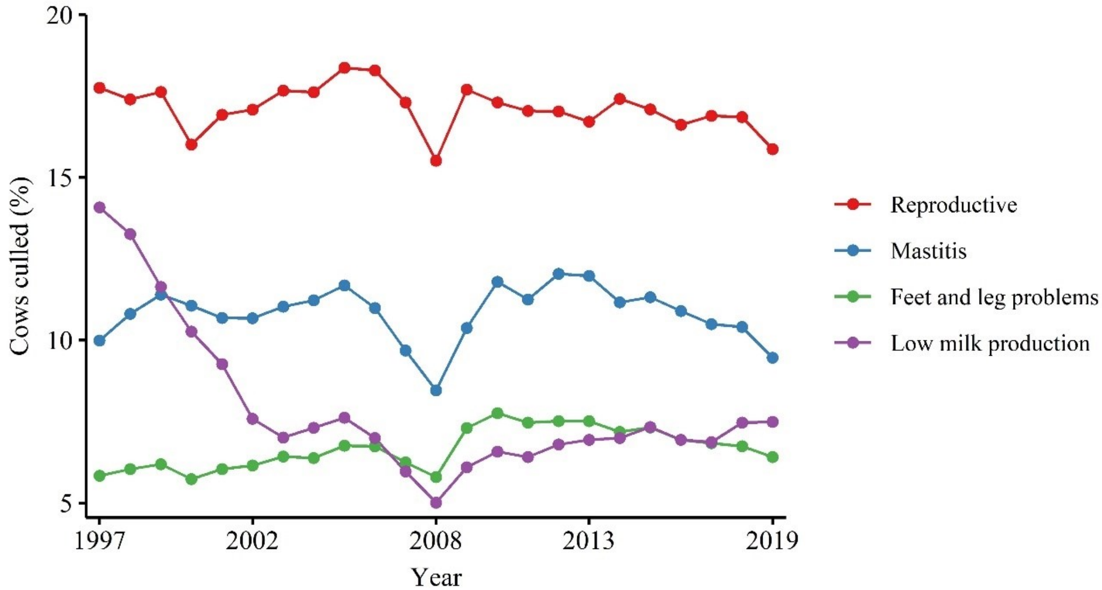


Figure 2.1. Change over time of the top four culling reasons based on the total number of cows culled with a known reported reason in Canada between 1997 and 2019 (CDIC, 2020a).

The risk of culling is not constant over the life of a cow. It depends on cow factors such as lactation number, stage of lactation, milk yield, and reproductive status as well as environmental factors such as season of calving and herd-production needs (Hultgren and Svensson, 2009, Pinedo et al., 2010, Pinedo et al., 2014, Heise et al., 2016, Haine et al., 2017). Death, as well as diseases and injury, are the main reasons for culling early after the onset of a new lactation (Pinedo et al., 2010, Pinedo et al., 2014). On the other hand, the risk of culling due to failure to reproduce and low milk production increases as the lactation progress and the highest risk is observed at later stages of the lactation (Pinedo et al., 2010, Pinedo et al., 2014, Heise et al., 2016). Milk yield and reproduction are protective factors against culling, in which pregnant (Hultgren and Svensson, 2009) and high yielding animals are less likely to be culled compared with its counterpart (Pinedo et al., 2010, Stojkov et al., 2020). Death is mostly associated with seasonal effects in which hot seasons are associated with a greater risk of dying (Pinedo et al., 2010). Cows are favoured to remain in the herd if they are healthy, reproduce regularly, have functional feet, legs, and udders, and produce enough milk (Essl, 1998, Fetrow et al., 2006, De Vries and Marcondes, 2020).

2.3.2. Longevity measures

Longevity can be categorized as true, functional, and residual longevity. True longevity indicates the ability of an animal to delay culling, but it is not adjusted for milk yield (Mark, 2004). Functional longevity indicates the ability of an animal to delay involuntary culling, and it is adjusted for milk yield within the herd (Mark, 2004). Lastly, residual longevity represents cow longevity after adjusting it for all other traits under consideration in the breeding program (Mark, 2004). Having culling or death as the endpoint and based on the different stages of the life of a dairy cow (Figure 2.2), different longevity metrics have been used (**Table 2.1**). These metrics can be obtained at the herd-level when they reflect the overall prevalence of animals that meet certain criteria such as the number of animals on third or greater lactation, or at the animal-level when each animal is individually evaluated.

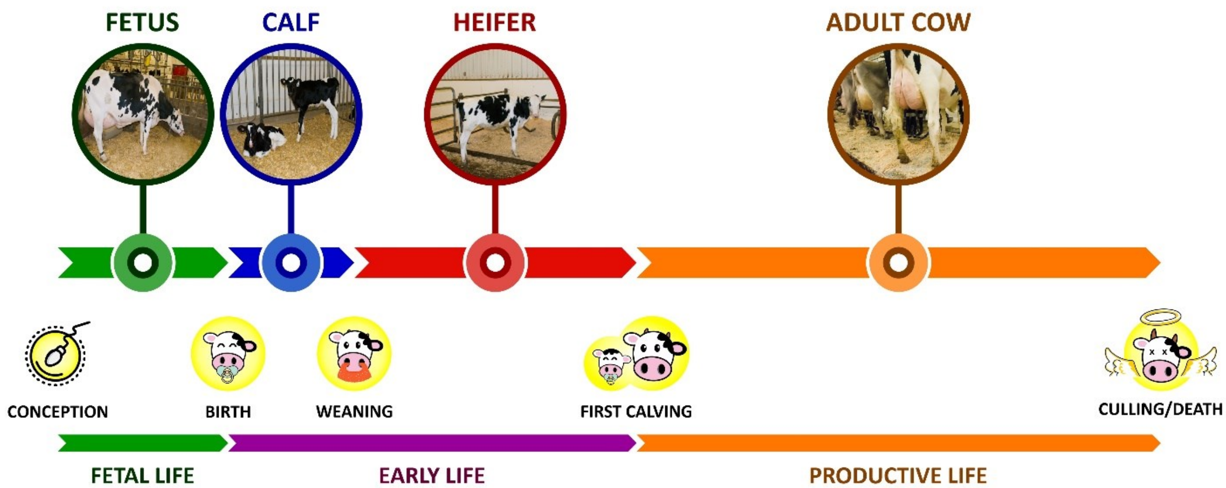


Figure 2.2. Schematic representation of the life of a dairy cow according to a chronological sequence of key events (conception, birth, weaning, first calving, and culling/death) that prompt a change to the different life status (fetus, calf, heifer, and adult cow) and respective life stages (fetal, early, and productive life). The length of the arrows is proportional to the duration of each status and stage as seen in the province of Quebec, Canada (Valacta, 2019).

Table 2.1. Dairy cow longevity metrics commonly used.

Measure	Unit	Time frame	Description	Reference
Lactation	Count	First calving to culling/death	Cumulative number of lactations	Essl (1998)
3+ lactation	Herd prevalence	The number at a given point	Percentage of cows on the third or greater lactation	Villettaz Robichaud et al. (2018), Villettaz Robichaud et al. (2019a), Villettaz Robichaud et al. (2019b)
Culling rate	Herd prevalence	The number at a given point	Percentage of culling	Villettaz Robichaud et al. (2018), Villettaz Robichaud et al. (2019a), Villettaz Robichaud et al. (2019b)
Length of life	Year	Birth to culling/death	Length of time between birth and culling	Haworth et al. (2008)
Length of productive life	Year	First calving to culling/death	Length of time between first calving and culling	Ducrocq (1994), Schneider et al. (2007)
Functional longevity	Rank	First calving to culling/death	Length of productive life adjusted for within-herd milk production level	Sewalem et al. (2008)
Longevity index	%	Birth to culling/death	Lifetime days in milk divided by length of life	Haworth et al. (2008), Brickell and Wathes (2011)

The different longevity metrics can be classified as stayability metrics or lifetime metrics. Stayability metrics have a binary nature and indicate if a dairy cow is alive at a given moment in time (van Pelt, 2017) and can be updated as the animal grows. An example would be if the cows reach the third or greater lactation (Villettaz Robichaud et al., 2018, Villettaz Robichaud et al., 2019a, Villettaz Robichaud et al., 2019b). Even though such metrics do not provide a complete picture of cow longevity, one of their advantages is that they can be measured at any time (van Pelt, 2017). On the other hand, lifetime metrics take into account the completed life stages of the animals (van Pelt, 2017). For example, the life of a dairy cow can be split into early life (non-productive) and productive stages (Figure 2) from a production perspective. Based on that, longevity can be measured as the length of the productive life of a dairy cow (Ducrocq, 1994, Schneider et al., 2007). Since lifetime metrics take into account the entire stage of life, they can only be calculated when such a stage is completed, which is one of the main limitations of such metrics.

Most lifetime metrics of dairy cow longevity do not specifically account for the early life stage (Figure 2), since they typically have first calving as the starting point. The longevity index (Table 1) is a proposed metric that overcomes such limitation by taking into account both the length of life of an animal and the length of time spent on producing milk (Haworth et al., 2008); therefore, accounting for the entire non-productive period of life (early life stage) and days dry of a dairy animal.

2.3.3. Limitation of common longevity measures

There is no standard metric to measure dairy cow longevity and even though each different metric reflects an aspect of dairy cow longevity, they are not comparable since they do not have the same meaning (Van Doormaal, 2009). Mark (2004) estimated the correlation between longevity metrics used by different Interbull (an international network that carries on genetic analysis of livestock animals) member countries. The author reported a low correlation coefficient (0.59) and high variability (range = 0.96) among all countries evaluated, regardless of how longevity was measured/defined in each country. However, the correlation increased (0.71) and the variability decreased (range = 0.51) while analyzing only countries that used comparable longevity traits. Differences in how longevity was measured between countries could be partially the reason for low correlation and great variability in both cases, but a correlation lower than unity among countries that used comparable metrics could also be due to differences in culling reasons and trait definitions (Mark, 2004), which indicates a lack of standardization on measuring longevity. At the same time, the differences in the environment could be a reason for slight differences in traits among different countries as well.

2.4. What is the current status of dairy cow longevity and milk yield?

The average length of productive life, which is one of the most common longevity metrics, can be estimated based on the culling rate (De Vries, 2013, 2020, De Vries and Marcondes, 2020). Since information at country level regarding culling is not available for most countries, a proxy can be estimated based on slaughtering data at country level, even though this approach would not take into account animals that died in the farm and would assume the accuracy of slaughter records reported by each country. Once this information is obtained, it can be used to evaluate the trend over time in the status of dairy cow longevity along with milk yield per animal. The following methodology was used to identify high milk-producing countries and estimate dairy cow longevity at the country level.

2.4.1. Sourcing the information

Countries were first ranked based on total milk production to identify the ones with the highest production. As a starting point, the average total whole fresh cow milk was calculated based on 2016, 2017, and 2018 information provided by the FAO (2020). The top 21st high milk-producing countries were kept. Next, we searched for yearly official statistics publications for each of these countries regarding the total number of dairy cows, total milk production (kg), average milk yield per animal (kg), and the number of slaughtered cows. No date limit was imposed at this stage and all information available was gathered and aggregated into a single data file. For countries that reported milk production in liters it was converted to kg using the 1.03 conversion factor. For countries that did not officially report average milk yield, it was estimated by dividing the reported total milk production over the number of dairy cows (Supplementary Table 2.1). References and official sources are presented in Supplementary Table 2.1.

The length of productive life was estimated based on the culling rate. A proxy of the average culling rate was estimated at the country level by dividing the number of dairy cows slaughtered per year by the total number of dairy cows in each year for the countries that we were able to find both information. For countries that did not specify the number of dairy cows slaughtered, we used the number of cows slaughtered (Supplementary Table 2.1). The inverse of the culling rate was then used as an estimation of the length of productive life (De Vries, 2013, 2020, De Vries and Marcondes, 2020).

Once the data was gathered and calculations were completed, two criteria were used in data cleaning to define its sufficiency and reliability, respectively. First, only countries that we were able to find information for at least two consecutive decades were kept for further steps. Next, information from countries in which cows had a length of productive life lower than 1.5 years (Argentina, Australia, and Mexico) in earlier decades or greater than 7 years (Turkey and United Kingdom) in recent decades were considered unreliable and excluded from further steps in this review. After cleaning, information from 10 countries remained (Figure 2.3; Supplementary Table 2.1).

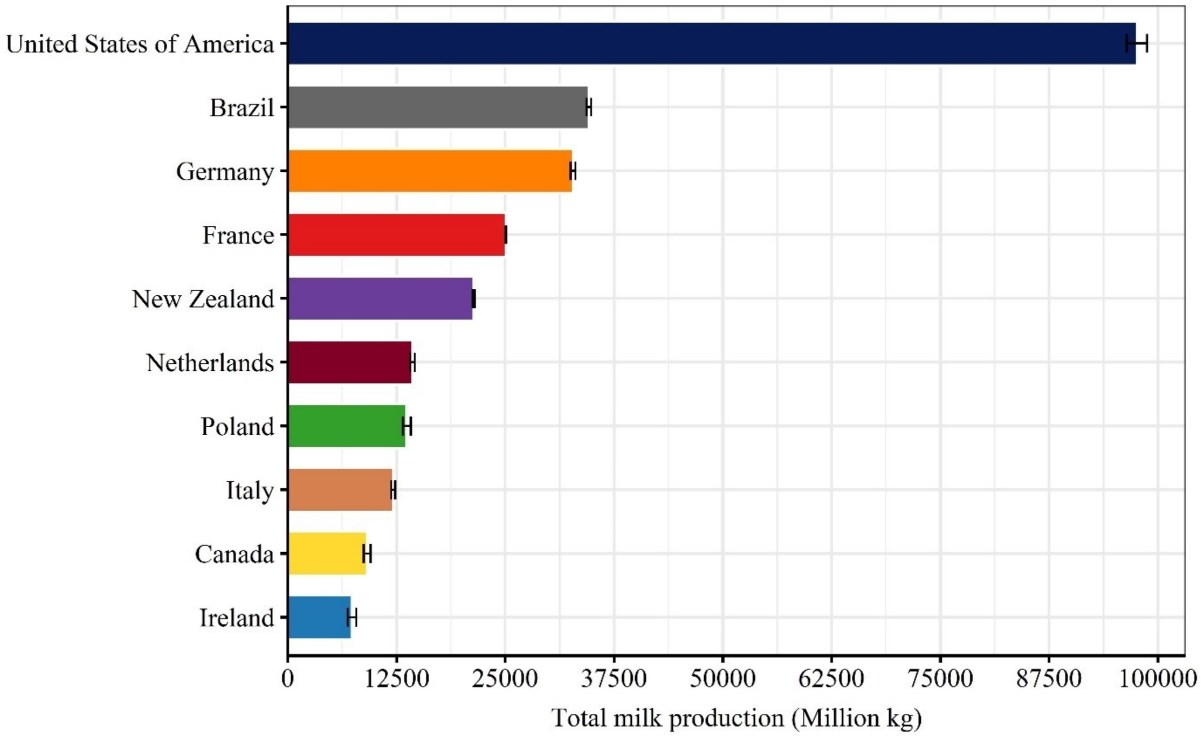


Figure 2.3. Top 10 high milk-producing countries based on total milk production averaged over the years 2016 to 2018. Columns represent the averages followed by the standard deviation (error bars). The list of countries is limited to those for which we were able to provide sufficient and reliable data on the length of productive life. Data sources are provided in Supplementary Table 2.1.

Linear regression was used to describe the trend over time in both milk yield and length of productive life. For milk yield, we reduced the number of observations to standardize the time window interval for all countries. Therefore, we considered only the information ranging from 1961 to 2018, which represented 96.9% of the data available after cleaning. For the length of productive life, it was not possible to establish a standard time window. For some countries, we were able to find reliable information only from more recent years while for others the collection was more extensive. The following polynomial regression model was used to describe both trends:

$$Y_j = \beta_0 + \beta_1 Year_j + \varepsilon_j, \tag{1}$$

in which Y_j represented the milk yield per animal (kg) or length of productive life (year), β_0 was the intercept, β_1 was the linear regression coefficient, $Year_j$ was the value observed in the jth year and ε_j was the residual error ~ N (0, σ^2). Statistical significance level was set at $\alpha < 0.05$.

2.4.2. Milk yield and longevity over time

The average milk yield per animal per year increased in all the countries considered in this review (Figure 2.4). However, the magnitude of the increase was not the same across countries. The estimated increase ranged from 18.5 kg (Standard error; SE = 1.49) per animal per year in Brazil to 129.7 kg (SE = 1.20) kg in the United States both from 1961 to 2018 (Table 2.2).

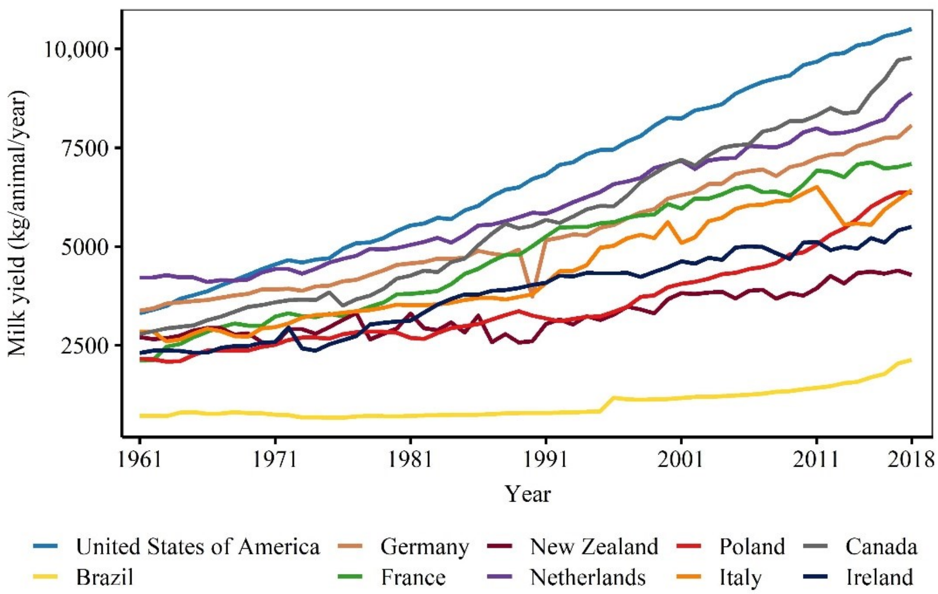


Figure 2.4. The average milk yield (kg) per animal from the top 10 high milk-producing countries over the years. The list of countries is limited to the world's top high milk-producing countries for which we were able to provide sufficient and reliable data on the length of productive life. Data sources are provided in Supplementary Table 2.1.

Table 2.2. The linear trend of milk yield (kg) per animal per year for each country between 1961 and 2018. The list of countries is limited to the world's top high milk-producing countries for which we were able to provide sufficient and reliable data on the length of productive life.

Country	**Model** [1]		**R^2** [3]	**RSE** [4]	***p*-value** [5]
	Intercept [2]	**Year** [2]			
United States of America	2941.6*** (40.6)	129.7*** (1.20)	0.99	152.5	< 0.001
Brazil	451.2*** (50.4)	18.5*** (1.49)	0.73	189.5	< 0.001
Germany	2904.6*** (83.4)	81.4*** (2.46)	0.95	313.4	< 0.001
France	2103.8*** (56.8)	91.9*** (1.68)	0.98	213.6	< 0.001
New Zealand	2419.1*** (70.9)	29.0*** (2.09)	0.77	266.7	< 0.001
Netherlands	3485.8*** (61.2)	84.3*** (1.80)	0.97	230.0	< 0.001
Poland	1603.6*** (107.4)	65.3*** (3.17)	0.88	403.8	< 0.001
Italy	2200.5*** (91.3)	72.5*** (2.69)	0.93	343.1	< 0.001
Canada	2081.3*** (84.1)	120.7*** (2.48)	0.98	316.3	< 0.001
Ireland	2035.0*** (53.3)	59.9*** (1.57)	0.96	200.5	< 0.001

[1] *** = *p*-value < 0.01; [2] Estimate (Standard error); [3] R2 = Coefficient of determination; [4] RSE = Residual standard error; [5] Model significance.

Improvements in nutrition, genetics, animal health, and management of environmental factors contributed to the increase in milk yield (Collier et al., 2006, Shook, 2006). However, the relative weight of such factors is likely not the same across countries. For instance, the tropical climate in Brazil limits the raising of high yielding animals such as Holstein cows, which are particularly susceptible to heat stress (Polsky and von Keyserlingk, 2017) and had their susceptibility highlighted due to intensive selection for milk production (Collier et al., 2006). Climatic conditions are not as limiting in countries under a similar low input pasture-based production system than Brazil but located in a cooler climate zone, such as New Zealand, where climatic conditions are adverse towards production for only up to 20% of days in a year (Bryant et al., 2007). On the other hand, milk yield increase in typical indoor-housing high input systems such as in the Netherlands, United States, and Canada was achieved by intense selection of animals based on milk production instead of increasing their resistance to climatic stressors and focused on improving nutritional management and developing artificial thermal conditioning systems (Collier et al., 2006).

Three different status of dairy cow longevity were observed in top high milk-producing countries (Figure 2.5). In most countries (6 out of 10), the length of productive life significantly decreased over the years, with a total estimated decrease ranging from 0.90 year in Ireland to 3.04 year in Poland (Table 2.3). New Zealand was the only country in which the length of productive life increased over time, with a total estimated increase of 1.85 years (Table 2.3). The length of productive life did not change in the United States, Germany, and the Netherlands (Table 2.3) with an average of 3.25 (SE = 0.09), 3.24 (SE = 0.07), and 3.14 (SE = 0.17) years, respectively.

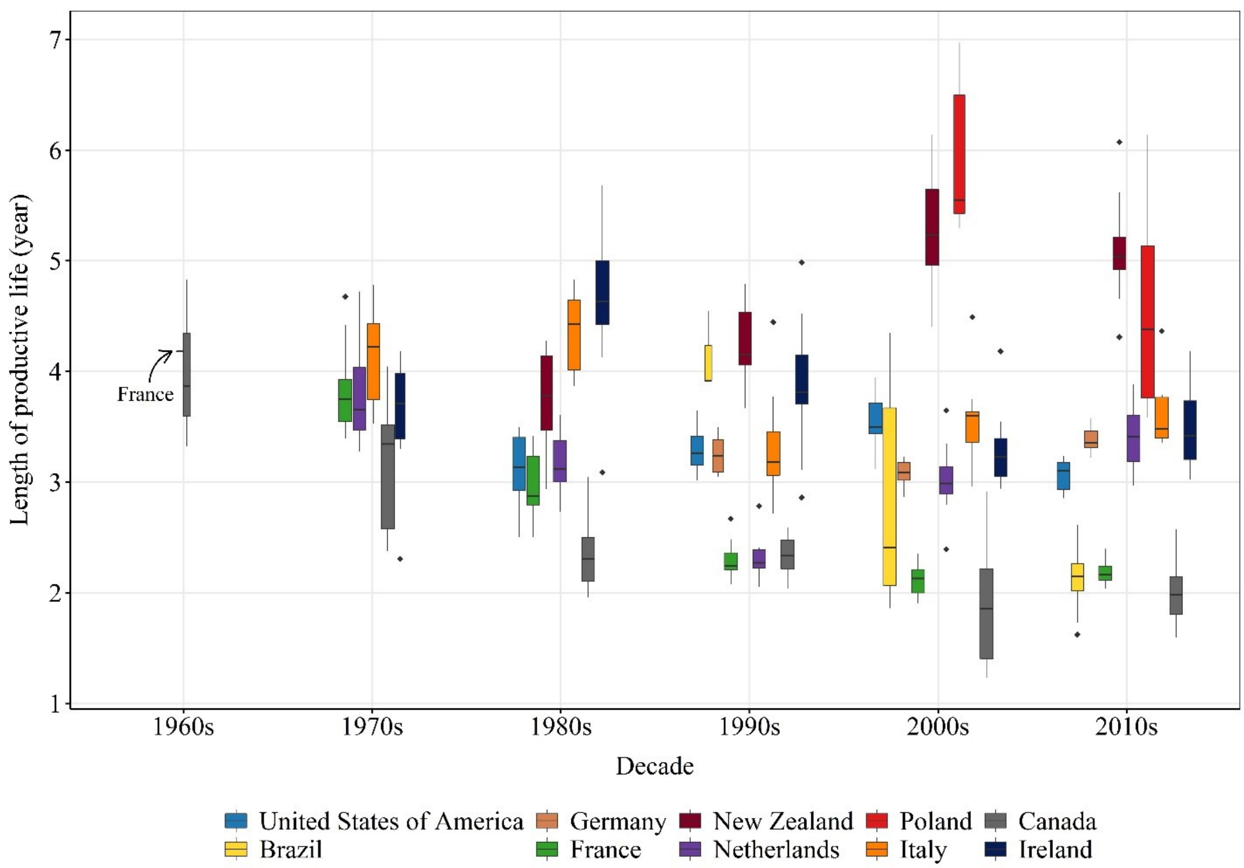


Figure 2.5. The length of productive life (year) of dairy cows from the top 10 high milk-producing countries on different decades. The relative width of each box per country within decades represents the number of observations available to generate it. The wider the box, the more observations were available. The list of countries is limited to the world's top high milk-producing countries for which we were able to provide sufficient and reliable data on the length of productive life. Full circles (•) represent values above or bellow the interquartile range. Data sources are provided in Supplementary Table 2.1.

Table 2.3. The linear trend of the length of productive life (year) in each country. The list of countries is limited to the world's top high milk-producing countries for which we were able to provide sufficient and reliable data on the length of productive life.

Country	Year	Model [1]		R^2 [3]	RSE [4]	*p*-value [5]
		Intercept [2]	Year [2]			
United States of America	1980 – 2019	3.25^{***} (0.10)	0.0004^{NS} (0.004)	0.0003	0.30	0.92
Brazil	1997 – 2018	4.06^{***} (0.24)	-0.12^{***} (0.02)	0.67	0.55	< 0.001
Germany	1993 – 2019	3.11^{***} (0.07)	0.01^{NS} (0.004)	0.13	0.18	0.06
France	1968 – 2019	3.89^{***} (0.10)	-0.04^{***} (0.003)	0.76	0.37	< 0.001
New Zealand	1982 – 2019	3.69^{***} (0.19)	0.05^{***} (0.01)	0.48	0.56	< 0.001
Netherlands	1970 – 2019	3.38^{***} (0.17)	-0.01^{NS} (0.01)	0.05	0.59	0.12
Poland	2003 – 2019	6.81^{***} (0.25)	-0.19^{***} (0.02)	0.79	0.50	< 0.001
Italy	1970 – 2019	4.26^{***} (0.14)	-0.02^{***} (0.005)	0.22	0.49	< 0.001
Canada	1967 – 2019	3.38^{***} (0.13)	-0.03^{***} (0.004)	0.57	0.47	< 0.001
Ireland	1974 – 2019	4.28^{***} (0.21)	-0.02^{*} (0.01)	0.14	0.70	0.01

[1] NS = Not significant, * = *p*-value <0.10, *** = *p*-value < 0.01; [2] Estimate (Standard error); [3] R2 = Coefficient of determination; [4] RSE = Residual standard error; [5] Model significance.

In order to look at the relationship between milk yield and longevity, the differences in production systems need to be considered since not every country uses the same system. For instance, most herds in New Zealand are under a low input pasture-based system while in Canada and the Netherlands cows are typically housed indoors. The average milk yield per animal in New Zealand in 2018 was 2.3 and 2.1 times lower than in Canada and the Netherlands, respectively (Figure 2.4), which was expected since milk yield in a pasture-based system is usually lower compared to indoor-housed systems. The opposite was observed for longevity between these countries. In the 2010s decade, the average length of productive life in New Zealand was 2.5 and 1.5 times higher compared to Canada and the Netherlands, respectively (Figure 2.5).

The highest incidence of involuntary culling due to fertility issues and health problems such as mastitis and lameness is one of the main factors responsible for a reduction in dairy cow longevity (Rushen and Passillé, 2013). Involuntary culling reduces the ability of dairy farmers to select animals based on production once they reach the productive life stage (Berry et al., 2005, Pritchard et al., 2013), forcing farmers to cull an animal that would otherwise be kept in the herd. However, such high incidence is not a reality in all farms within countries, indicating differences among farmers in their ability to keep animals healthy and comfortable for longer in the herd based on adopting management and housing practices that in turn prevent the occurrence of health problems associated with involuntary culling (Rushen and Passillé, 2013).

Differences in production systems could be associated with the longevity status of the animals in different countries. Indoor housing and high input milk production system are two of the main characteristics shared by most of the high milk-producing countries in this review in which the length of productive life decreased over time. In turn, these are also two of the main differences compared to the production system in New Zealand, where the length of productive life increased. Even though a comparison between systems regarding their effect on the main involuntary culling reasons (reproduction, mastitis, and feet and leg problems) would be inevitably confounded by milk production and animal characteristics between countries even within the same breed, it could be a starting point in exploring the reasons underlying such differences in longevity between countries.

2.4.3. Longevity and involuntary culling

Information on culling and culling reasons is not available at the country level for most of the high milk-producing countries covered in this review. Therefore, we rely on herd prevalence reported by epidemiological studies, which are usually conducted on a limited number of animals and farms.

2.4.3.1. Reproduction

Failure to reproduce is the most frequent reason for involuntary culling worldwide (Pinedo et al., 2010, Heise et al., 2016, CDIC, 2020a) and the incidence of uterine diseases have a negative effect on animal reproduction, which could lead to a shortened longevity. Endometritis is the most prevalent uterine disease in dairy cows. Its prevalence was 27.1 or 25.1%, depending on the diagnostic method (degree of purulent vaginal discharge or cytology of the endometrium, respectively) in New Zealand (McDougall et al., 2020). In the United States, the prevalence of clinical endometritis was 15.0% (Ribeiro et al., 2013) while the prevalence of subclinical endometritis ranged from 13.4% (Ribeiro et al., 2013) to 53% (Gilbert et al., 2005). Uterine diseases have a negative effect on animal reproduction by increasing the number of artificial inseminations per pregnancy, delaying the restart of estrous cyclicity (Ribeiro et al., 2013), and reducing the overall pregnancy rate (Gilbert et al., 2005, Dubuc et al., 2011, Giuliodori et al., 2013). Therefore, cows with high longevity are likely to have a better reproductive performance, such as shorter calving interval, require a lower number of inseminations to become pregnant, and reduced number of days to first service (Pritchard et al., 2013). However, having had uterine diseases do not put the cow at a greater risk of being culled if she gets pregnant (Dubuc et al., 2011), which demonstrate the protective effect of a positive reproduction status (being pregnant) against involuntary culling.

The reproductive performance of cows under different production systems has not been extensively studied. The reproductive health (calving difficulty, puerperal metritis, and endometritis) of seasonally bred dairy cows in a rotational grazing system tended to be better compared to cubicle housed cows in Ireland (Olmos et al., 2009b). A multi-year experimental study conducted by Washburn et al. (2002) at the North Carolina State University – the United States between 1995 and 1998 compared the reproductive performance of seasonally bred Jersey and Holstein cows kept under pasture or housed in a free-stall barn. Reproductive performance was measured as the percentage of pregnant animals in 75 days after the beginning of the breeding season and no difference was observed ($P > 0.05$) between systems or between breeds. However, such results need to be interpreted carefully, especially in places with climatic conditions different from those observed in these studies since animals on pasture are more susceptible to the climatic environment, which in turn can negatively affect reproduction. During summer months, the conception rate of Holstein cows kept in paddocks with little or no shade decreased by 18% in a study conducted in Florida, US (Cavestany et al., 1985). In addition, oocyte quality and the development of fertilized oocytes are negatively affected by the increase in temperature observed during the summer in Holsteins cows under pasture in Louisiana, US (Rocha et al., 1998). Such negative effects are likely to be intensified in the future, given the expected changes in climate conditions.

2.4.3.2. Mastitis

Mastitis is the most common disease in dairy cows and its occurrence varies between countries as well as within countries. The average incidence rate of clinical mastitis in Canada between November 2003 and July 2005 was 23.0%, but it ranged from 0.7 to 97.4% (Olde Riekerink et al., 2008), which indicates great variability between farms. In the Netherlands, the incidence of clinical mastitis was 33.8% (95% CI = 31.7 – 36.1) (van den Borne et al., 2010). A much lower average incidence rate of 12.7% as well as a narrowed range from 1.9 to 35.8% was observed in New Zealand between July 2004 and June 2005 (McDougall et al., 2007). In Brazil, where most of the dairy animals are on a pasture-based system similar to New Zealand, the average prevalence of clinical mastitis was 46.4%, but it ranged from 1.45 to 100% (Busanello et al., 2017) while in Northern Ireland, where pasture is also largely used, the incidence was 29% between 2010 and 2015 (Bell and Wilson, 2018).

Pasture-based systems are often associated with a lower occurrence of mastitis compared to indoor-housed cows. For instance, Jersey and Holstein cows housed in a free-stall barn had 1.8 times more cases of clinical mastitis compared to cows on pasture ($P < 0.05$), which resulted in free-stall cows having a culling rate due to mastitis eight times higher than cows on pasture in the United States (Washburn et al., 2002). Regular access to pasture was reported to be a protective factor against mastitis since it decreased the odds ratio of veterinary treated mastitis (Odds ratio; OR = 0.73, $P < 0.05$) in Austria (Firth et al., 2019). Indoor housing was also associated with a 4.86 OR of developing subclinical mastitis during the first 41 days of lactation in Germany (Krömker et al., 2012).

The cleanliness of the animals, which indicates the level of exposure to environmental pathogens, seems to be one of the reasons for such a protective factor of pasture. The cleanliness of stalls in a free-stall barn was positively correlated with the hygiene scores for udder (DeVries et al., 2012), which in turn was associated with increased somatic cell count (Schreiner and Ruegg, 2003, Reneau et al., 2005). Cows that had access to pasture were 3.75 (SE = 1.89; $P < 0.05$) times less likely to be dirty compared to cows that did not in Danish dairy farms (Nielsen et al., 2011). However, the cleanliness of cows on pasture or at outdoor paddocks is directly influenced by climatic conditions, which in the rainy season is associated with dirtier cows while the opposite is observed during the dry season (Sant'anna and Paranhos da Costa, 2011). In addition, the occurrence of mastitis is associated with hygiene practices and improving those are a low-cost solution that improves animal performance (Langford and Stott, 2012) and the incidence of mastitis. In indoor housing, increasing the frequency of cleaning the barns could be a strategy to reduce the level of exposure to pathogens, since that cleaning the floors more than 4 times per day was associated with a reduction in clinical mastitis incidence (OR = 0.77; 95% CI = 0.62 – 0.96; $P < 0.05$) in the Netherlands (Santman-Berends et al., 2016).

2.4.3.3. Feet and Leg

The occurrence of lameness is lower in cows on pasture compared to indoor-housed cows. The prevalence in New Zealand is 8.1% (Fabian et al., 2014) compared to 22.2% and 24.6% in Canada (Higginson Cutler et al., 2017) and the United States (Espejo et al., 2006), respectively. Such difference between systems is even present within the same countries. In the UK, Haskell et al. (2006) reported that zero-grazing farms had 2.6 more lame cows compared to grazing farms while in the United States, Adams et al. (2017) reported that in farms where cows were primarily housed in free-stall barns had an 6.9 (SE = 0.60) greater incidence density ratio of severely lame cows than in farms where cows were kept mainly in pasture. Concrete floor is a risk factor in increasing the incidence of claw lesions (Somers et al., 2003) and lameness (Somers et al., 2003, Vanegas et al., 2006). To that end, access to pasture could be beneficial because it has been associated with improving hoof health, healing of lesions, and decreasing the incidence of lameness (Somers et al., 2003, Hernandez-Mendo et al., 2007, Olmos et al., 2009a). The low incidence of lameness in New Zealand could result in a reduction of involuntary culling due to feet and leg problems and potentially increase the longevity of dairy cows in this country compared to Canada and United States where pasturing cows is seldom practised. However, in addition to information on culling not being available at the country level for most countries, failure in detecting lameness by farmers is a limiting factor in using herd prevalence as a proxy for culling reason. The prevalence of lameness is 3 to 4 times higher than that estimated by farmers (Espejo et al., 2006, Fabian et al., 2014, Beggs et al., 2019). In addition, lame cows are not necessarily culled since they can be treated if the producer chooses to do so. Lameness is also associated with negative reproductive performance and milk production (Huxley, 2013), which in turn might be the reason reported for culling by the farmer.

By itself, the pasture-based system is not responsible for reducing the prevalence of lameness in dairy cows. An overall prevalence of 39% was reported by Thompson et al. (2019) while evaluating 252 dairy cows from six pasture-based herds in the southern region of Brazil, which is higher than that reported in indoor housed animals in Canada (22%) (Higginson Cutler et al., 2017) and US (24.6%) (Espejo et al., 2006). Environmental conditions and management practices such as the amount of rainfall, condition of tracks to pasture, poor hygiene, and human-animal relationship are important factors associated with lameness in pasture-based farms (Ranjbar et al., 2016, Moreira et al., 2019). In addition, most of the time, farmers can only report one reason for culling, and feet and leg issues may not be the primary reason for culling due to their failure to detect lame animals (Espejo et al., 2006, Fabian et al., 2014, Beggs et al., 2019).

2.5. Should we improve dairy cow longevity?

Short longevity poses a threat to the sustainability of the dairy industry since it is associated with financial losses on farms, increased environmental footprint of milk production, and welfare issues for the animals, which in turn is a growing social concern among consumers (Essl, 1998, Benbrook et al., 2010, Brickell and Wathes, 2011, Pellerin et al., 2014, Boulton et al., 2017, Grandl et al., 2019). Therefore, improving dairy cow longevity would contribute to achieving a more sustainable industry, since it would have a positive effect towards the three pillars of sustainable agriculture: economic profit, environmental impact, and social concerns.

2.5.1. Economic profit

For a dairy farm to be profitable, dairy cows need to be able to reproduce regularly, maintain high milk production, and do not fall ill for many years (Mulder and Jansen, 2001). Therefore, increasing the length of productive life is a potential option to improve the profitability of the dairy activity (Grandl et al., 2019). In fact, it is the second most economically important trait in dairy cows, while milk yield is the first most important trait (Komlósi et al., 2010). Short longevity indicates that animals are not expressing their maximum potential for productivity and profitability, since dairy cows become profitable at their third lactation due to high costs associated with the early life non-productive stage (Pellerin et al., 2014, Boulton et al., 2017). In addition, more first and second lactation cows are culled as culling rate increases (Dhuyvetter et al., 2007), which decreases animal longevity and reduces the profitability of the system. Overall, the most common reason for culling of first and second lactation cows is reproduction issues while death is the most common reason for third and greater lactation cows (Pinedo et al., 2010). During the initial third of the lactation, first calving cows are more likely to be culled due to low milk production and milkability while second lactation animals are culled due to the incidence of metabolic and other diseases (Heise et al., 2016). However, higher risk of culling due to failure to reproduce is observed in the final third of the lactation for both first and second lactation cows (Heise et al., 2016). With increasing longevity by decreasing the culling of animals in the beginning of their productive life, there will be a high number of cows on more profitable lactations in the herd and the replacement cost per day will be relatively reduced since it would be split into more lactations (Essl, 1998).

Having a greater proportion of mature cows because of increased longevity would reduce the number of replacement heifers required to achieve the same milk production since mature cows have a relatively higher milk yield compared to young animals. This is particularly relevant under a supply-management system such as the one present in Canada (Van Doormaal, 2009), where profitability is associated with increased efficiency of using the resources available and reducing input costs rather than increasing milk production. However, it would allow for the commercialization of extra heifers (Brickell and Wathes, 2011, Pritchard et al., 2013, De Vries, 2017) and potentially increasing this additional source of income. In addition, increasing longevity by reducing involuntary culling would improve lifetime profit (Essl, 1998) especially given the negative economic impact of factors underlying health problems associated with involuntary culling (Langford and Stott, 2012).

Even though longer longevity alone does not assure an increase in profitability, a farm with short longevity due to a high involuntary culling and its associated diseases is not likely to be profitable either (De Vries, 2020). The adoption of management practices and technologies to improve cow health and longevity is essential to achieve a profitable dairy industry in the future, which is a key factor in achieving sustainability (Walter et al., 2017, Britt et al., 2018).

2.5.2. Environmental impact

Increasing longevity would reduce the environmental toll of the dairy industry. Longer longevity would reduce the required number the replacement heifers needed on a farm, which contribute with 21 to 26% of the total enteric emission of methane in a herd (Wall et al., 2012). At the same time, it would reduce the proportional emission from replacement heifers. Assuming an age at first calving of 28 months, Knapp et al. (2014) estimated that increasing the length of productive life from 2.5 years (40% culling rate) to 4.0 years (25% culling rate) would reduce by 9.5% the enteric emission contribution of replacement heifers. In addition, methane emission per animal does not increase as the animal gets older (Grandl et al., 2016). In fact, an increase in the length of productive life was associated with a decrease in methane emission per kg of milk corrected for fat and protein (Grandl et al., 2019), which contributes to decreasing the footprint associated with milk production (Benbrook et al., 2010) and supports the argument that increasing dairy cow longevity would decrease the environmental burden of the dairy industry.

2.5.3. Social concerns

Early age at culling is a growing concern among consumers (Berry, 2015), especially because cow longevity is a global indicator of animal welfare since higher cow longevity indicates that the animal biological functions and health are not impairing the length of its life (Bruijnis et al., 2013). In addition, the health issues associated with the most common reasons for involuntary culling reported by dairy farmers bring into question the welfare conditions and ethical concerns towards dairy farming (De Vries, 2020).

The high incidence of involuntary culling due to reproduction problems (Pinedo et al., 2010, Heise et al., 2016, CDIC, 2020a) might hide underlying health problems. For instance, the occurrence of reproductive diseases (Gilbert et al., 2005, Dubuc et al., 2011, Giuliodori et al., 2013, Ribeiro et al., 2013) as well as lameness (Huxley, 2013) and mastitis (Kumar et al., 2017) have a negative effect on the ability of an animal to get pregnant and might result in animals being culled with failure to reproduce as the reported reason. However, increased longevity is not always associated with improved cow welfare. The incidence of health problems is directly associated with a poor cow welfare status and older animals are more likely to develop health problems such as lameness (Pötzsch et al., 2003) and mastitis (Firth et al., 2019) as well as body injuries (Bouffard et al., 2017). Therefore, the increase in cow longevity should be the result of improving the ability of dairy farmers to keep animals healthy and comfortable, which in turn improves the overall animal welfare status.

The main reported reasons for involuntary culling imply a lower status of animal welfare, which was the primary issue raised by consumers towards an ideal dairy farm (Cardoso et al., 2016). Leg problems such as lameness or foot disorders are considered the most detrimental condition on animal welfare (von Keyserlingk et al., 2009, Bruijnis et al., 2013), while peripartum problems such as dystocia and retained placenta, which are associated with decreased reproductive performance, can be life-threatening or occur because of chronic stressful conditions (Burnett et al., 2015).

Animal welfare becomes economically important to consumers once they attach importance to animal suffering (Molento, 2005). Consumers from Europe (European Commission,

2016) and the United States (Wolf and Tonsor, 2017) indicated a willingness to pay more for animal-based products obtained from farms with high welfare status. European consumers also stated that products imported from other countries should be subject to the same level of welfare standards that are imposed on farmers in the European Union (European Commission, 2016), which indicates that animal welfare could become a commercial barrier between countries (Molento, 2005, Bond et al., 2012). Even though willingness to pay does not always translate into action, it would be prudent to expect that future demand for higher welfare status of dairy cattle among consumers will remain, including a demand for longer longevity (De Vries, 2020).

2.6. How can we improve dairy cow longevity?

Dairy cow longevity is the outcome of decisions made by dairy farmers throughout the life of the animal, which dictates the moment and the reason a cow is culled. It is a dynamic process where multiple factors and their interactions are to be considered by the farmer (Roche et al., 2020). Therefore, all aspects of a cow's life need to be considered to reduce the rate of involuntary culling and increase longevity (Essl, 1998). In addition, most lifetime metrics of longevity only become available once the animal is culled. To overcome such limitation, a currently rich area for research is the identification of metrics available earlier in the life of the animal that, in turn, are correlated with lifetime longevity metrics available later in life.

2.6.1. Early Life Indicators

2.6.1.1. Age at first calving and its association with longevity metrics

Age at first calving (AFC) is a relatively early life metric, which has been extensively studied. The average AFC between high milk-producing countries is presented in Table 2.4, which ranged from 24.6 in the Netherland to 32.6 in Brazil. Age at first calving is associated with the ability of cows to remain in the herd and avoid culling, since animals that calved for the first time at a young age are less likely to be culled early during the productive life. Based on information from 437 herds across the United Kingdom, Sherwin et al. (2016) reported that cows with an AFC greater than 30 months were 1.71 times more likely ($P < 0.05$) of being culled compared to animals with an AFC of 23 – 24 months. In another study conducted on 7,768 Holstein heifers born between 2004 and 2006 in Spain, Bach (2011) reported that heifers which finished their first lactation had an average AFC of 23.8 months compared with an average AFC of 24.2 months of animals that did not.

Table 2.4. Age at first calving (month) and length of life (year) of dairy cows in high milk-producing countries. The list of countries is limited to the world's top high milk-producing countries for which we were able to provide sufficient and reliable data on the length of productive life and recent age at first calving from official milk recording agencies.

Country	Year	Recorded herds	Recorded cows	Percentage of recorded cows [1]	Breed	Recorded 1st lactations	Age at first calving	Reference	Length of life [2]
United State of America	2019	2,140	---	---	7 different dairy breeds	---	25.5 [4]	AgSource (2020)	4.98
Brazil	2018	334	15,459 [3]	0.09	Girolando (Holstein/Gir crossbreed)	12,384	32.6	GIROLANDO (2020)	4.34
Germany	2018	---	---	---	15 different dairy breeds	967,996	27.7	BRS (2019)	5.67
France	2018	35,253	2,437,250	69.0	20 different dairy breeds	776,679	30.0 [5]	idele (2019)	4.59
Italy	2019	15,316	1,351,442	72.7	30 different dairy breeds	321,298	27.3 [5]	AIA (2020)	5.69
Poland	2019	20,644	820,653	37.1	12 different dairy breeds	250,159	26.7	PFHBIPM (2020)	6.23
Netherlands	2019	14,367	1,459,287	91.9	Black-and-white dairy breeds, Red-and-white dairy breeds, and others	---	24.6	CRV (2020)	5.88
Ireland	April 2020	---	1,599,498	---	---	---	26.5	ICBF - Irish Cattle Breeding Federation (2020)	6.39 [7]
Canada	2019	7,063	658,311	68.0	---	---	25.0 [6]	CDIC (2020b), Lactanet (2020a, 2020b)	3.89

[1] Relative to the total number of cows in the country; [2] Age at first calving plus the length of productive life from each country (Figure 2.5); [3] Number of lactations recorded; [4] Average of averages weighted over the number of herds by breed, since the number of recorded 1st lactations was not available; [5] Average of averages weighted over the number of recorded 1st lactations; [6] Median; [7] Estimated using the length of productive life of 2018.

Age at first calving is also associated with the length of productive life. Swedish dairy cows that had an AFC of 27 – 28 months were 1.1 times more likely ($P < 0.05$) to have a shorter length of productive life compared to animals with an AFC younger than 25 months (Hultgren and Svensson, 2009). Similar results were reported in a study carried out by Nilforooshan and Edriss (2004) using production and pedigree data from Iranian Holstein cows collected between 1991 and 2011 from 45 herds, in which the length of productive life decreased as AFC increased ($P < 0.05$). The opposite was reported in a study using records from a single Australian farm from 1992 to 2005, in which animals with an AFC greater than 36 months had a longer length of life ($P < 0.05$) compared to animals calving for the first time between less than 24 to 36 months (Haworth et al., 2008). However, the opposite was observed for the longevity index in the same study ($P < 0.05$), indicating that animals with an older AFC had a longer length of life, possibly because the animals were inseminated older for the first time since the number of parities per lifetime did not differ ($P = 0.28$) between animals (Haworth et al., 2008).

2.6.1.2. Other early life indicators and their association with longevity metrics

Looking at longevity with AFC as the starting point overlooks early life (Figure 2.2) management practices and decisions made by the dairy farmer and their effect on the productive life of dairy cows. Even though it has received much less attention in the literature, there is an increasing interest in the subject. Housing calves from 3 to 7 months in litter pens with ≤ 12 calves resulted in a median increase ($P < 0.05$) of 18.2 months in the survival time compared to calves housed in slatted pens with > 7 calves (Hultgren and Svensson, 2009). The age in which the animal first consumed 0.91 kg/d of grain (dry matter basis) was positively associated with the age when removed from the herd (Heinrichs and Heinrichs, 2011). Also, housing automatically fed calves in small groups (6-9 calves) was associated with a higher growth rate (0.022 cm/day, about 40 g/day, $P < 0.05$) compared to calves housed in larger groups (12-18 calves; Svensson and Liberg (2006)). In turn, the higher the average daily gain (ADG) of weight in different ages before the first calving, the younger ($P < 0.05$) the AFC (Vacek et al., 2015). In pasture-based dairy herds, Chuck et al. (2018) reported a positive association ($P < 0.05$) between ADG from 1 month of age to first breeding on cumulative milk, fat, and protein yield at 100 and 250 days in milk in primiparous. Average daily gain from birth to weaning was also negatively associated ($P < 0.05$) with the occurrence of veterinary treated cases of mastitis from 7 to 30 days post-partum in primiparous cows (Svensson et al., 2006).

Health events in early life, the season of birth, and inbreeding are associated with cow longevity. The occurrence of severe calfhood respiratory disease was associated with a 12% increase ($P < 0.05$) in the calving interval of Swedish Red dairy cows (Hultgren and Svensson, 2010). Fall- and winter-born calves had a higher 8-week calf starter intake (48.3 kg vs 42.75 kg), ADG (0.66 vs 0.625 kg/d), and body weight (77.5 vs 75.0 kg) compared to spring- and summer-born calves ($P < 0.05$) (Chester-Jones et al., 2017). While using Dairy Herd Improvement (DHI) data between 1980 and 2004 from Canadian Holstein cows (n = 1977311), Sewalem et al. (2006) reported that animals with an inbreeding coefficient of 6.25 to 12.5% were 1.14 times more likely to have a shorter length of productive life, with the likelihood increasing as high as 1.51 times with inbreeding coefficient ≥ 25.0%.

2.6.1.3. Fetal life and its association with longevity metrics

Birth conditions are other less explored factors. The effect of complications during calving on dam longevity is well described in the literature. For instance, Holstein cows that require a hard pull and surgery during calving were 1.27 and 1.92 times more likely of being culled ($P < 0.05$), respectively compared to animals with unassisted calving (Sewalem et al., 2008). However, the effect on offspring longevity has been less studied. A study conducted by Heinrichs and Heinrichs (2011) on 21 dairy farms located in Pennsylvania, US reported that delivery scores indicating unassisted, easy pull, hard pull, mechanical extraction, or cesarean section were not associated ($P = 0.11$) with age when the offspring were removed from the herd. However, more studies are needed.

Fetal programming (the effect of dam conditions during conception and gestation on offspring performance) has been more extensively studied in beef (Du et al., 2010, Du et al., 2017) compared with dairy animals, but a few studies demonstrated associations between the dam's conditions on outcomes observed later in the life of the offspring. For instance, dam's intrauterine conditions associated with milk production seem to have an effect on offspring performance and survival (Berry et al., 2008), even though metabolic stress due to milk production might have a stronger effect than dam milk production alone (Bach, 2012). In addition, high milk urea nitrogen is associated with decreased fertility in dairy cows (Butler et al., 1996, Rajala-Schultz et al., 2001) and have a negative effect on early stages of oocyte development (De Wit et al., 2001, Kowsar et al., 2018). The longevity of calves originated from oocytes of cows with high milk urea nitrogen before ovulation has not been studied and accounting for the dam condition during pregnancy might be a possibility to improve offspring longevity (Opsomer et al., 2017).

2.6.2. Lack of space and quota constraints

Dairy cow longevity is not only influenced by intrinsic cow factors, but also by extrinsic factors such as availability of space in the farm as well as market characteristics, which could also influence the involuntary culling. In a situation where there is a surplus of heifers, farmers would decide to cull animals to make space for heifers that just calved (De Vries and Marcondes, 2020). In places under a supply management system such as in Canada (Van Doormaal, 2009), it would be difficult to accommodate the increase in milk production as a result of having more cows in more productive lactations (Brickell and Wathes, 2011) as a result of increased longevity or when producers have contracts with dairy processors that limit the amount of milk they can deliver. Both conditions would influence decreasing the rate of involuntary culling and increasing longevity.

A possible alternative would be to combine the use of sexed semen with extending the lactation of high-yielding cows by increasing the voluntary waiting period. Such a strategy would reduce the frequency that cows undergo the beginning of the lactation, which is associated with greater risk for involuntary culling due to death and diseases (Pinedo et al., 2010, Pinedo et al., 2014, Heise et al., 2016). At the same time, the use of sexed semen would reduce the negative effect of extended lactation on the genetic return (Clasen et al., 2019). Extending the duration of lactation of high yielding animals was also shown to have no negative effect on the gain of body condition score, udder health, milk production, and culling (Niozas et al., 2019a) while improving their reproductive performance (Niozas et al., 2019b).

2.7. Proposing a more comprehensive definition of cow longevity

Contrarily to milk and milk components, dairy cow longevity is neither routinely measured nor reported. This could be partly justified by the lack of a sound definition of the term and, as a result, the nonexistence of a standard metric designed to cover all aspects outlined in the definition. The definition of longevity should take into account the health, reproductive performance, and milk production of any given animal during its entire lifespan, which in turn are key factors associated with culling (Heise et al., 2016, Compton et al., 2017, CDIC, 2020a) and the profitability of the dairy industry (Essl, 1998, Mulder and Jansen, 2001, Fetrow et al., 2006, De Vries and Marcondes, 2020). As much as possible, the definition should allow for the use of metrics that are already routinely collected from either farms or DHI agencies, making it easy to be implemented and to increase the chances of being widely adopted. To that end, dairy cow longevity could be defined as an animal having an early age at first calving and a long productive life spent under profitable levels of milk production.

This definition covers both early life conditions and the stayability of the animal once it reaches the lactating herd as well as its overall health and quality of life. Assuming that an animal would be inseminated for the first time as soon as it is ready, early age at first calving would indicate that the animal was raised under healthy and favourable early life conditions. Next, a long and profitable productive life would imply that the animal produced enough milk to justify keeping it under milking, reproduced regularly avoiding a potential extension of the lactation to unprofitable levels or unnecessarily long dry periods, and maintained good health since the incidence of health issues are directly linked with reproduction failures and reduction in milk production. Age at first calving, length of productive life, and margin over all costs are metrics that could be used as indicators of early life conditions, length of life, and profitability, respectively. Combined, they would provide a more comprehensive approach to measure dairy cow longevity (Figure 2.6).

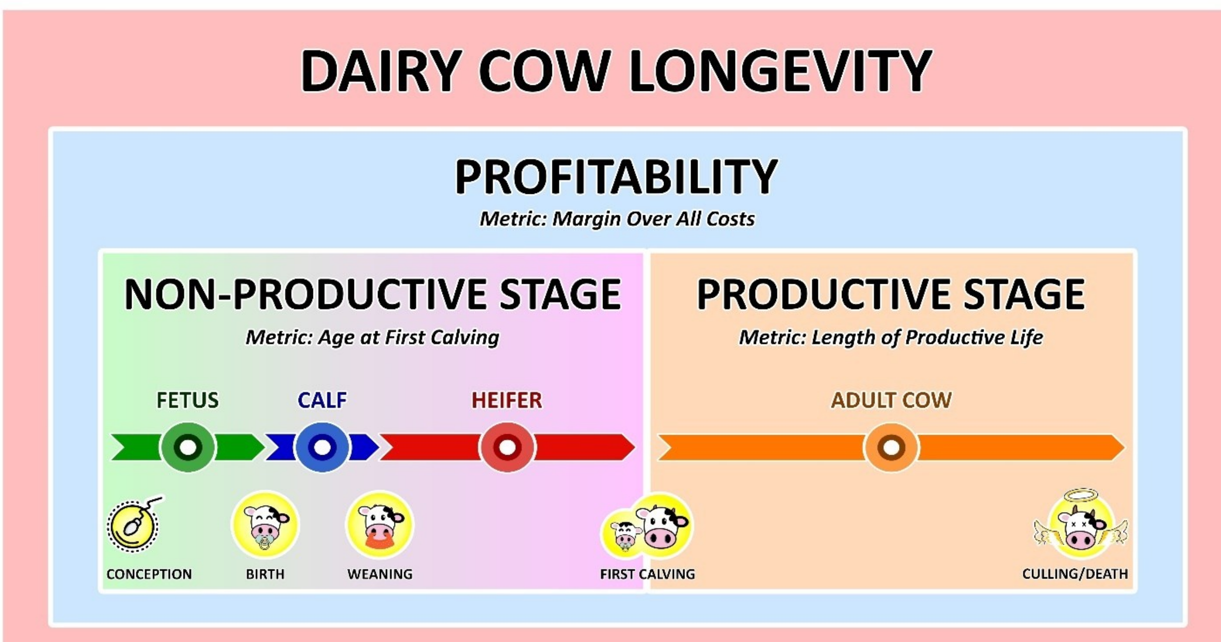


Figure 2.6. Relationship between the concepts of profitability, non-productive and productive life stages of dairy cows for a more comprehensive definition of cow longevity along with proposed metrics representing each respective concept.

2.8. Conclusions

The current metrics available to measure longevity often starts at the first lactation, overlooking early life management practices and decisions made by the dairy farmer before that point. To overcome such limitation, first, we propose that dairy cow longevity should be defined as an animal having an early age at first calving and a long productive life spent under profitable levels of milk production. Next, a combination of the metrics age at first calving, length of productive life, and margin over all (available) costs would provide a more comprehensive evaluation of longevity and cover all aspects of the definition.

By using a standard methodology, this critical literature review confirms the concerns raised by the dairy industry and other stakeholders that dairy cow longevity has decreased in most high milk-producing countries. Early life indicators are needed to support farmers in the early selection of animals that are more likely to reach their maximum potential. Increasing cow longevity due to a reduction in involuntary culling would reduce health costs, increase cow lifetime profitability, improve animal welfare and quality of life, and contribute towards a more sustainable dairy industry by producing milk with inherited sustainability while optimizing dairy farmers' efficiency in the use of resources.

2.9. References

Adams, A. E., J. E. Lombard, C. P. Fossler, I. N. Román-Muñiz, and C. A. Kopral. 2017. Associations between housing and management practices and the prevalence of lameness, hock lesions, and thin cows on US dairy operations. Journal of Dairy Science 100(3):2119-2136. https://doi.org/10.3168/jds.2016-11517.

AgSource – AgSource Dairy Cooperative Services. 2020. Industry benchmarks: AgSource Breed averages. Accessed on 21 April 2020. https://dairy.agsource.com/wp-content/uploads/2020/03/2020_cow_breed_new_avgs.xls.

AIA – Associazione Italiana Allevatori. 2020. Bollettino dei controlli della produttività del latte. LVIII ed., Roma, Italy. Accessed on 19 April 2020. http://bollettino.aia.it/Contenuti.aspx?CD_GruppoStampe=RS&CD_Specie=C4.

Bach, A. 2011. Associations between several aspects of heifer development and dairy cow survivability to second lactation. Journal of Dairy Science 94(2):1052-1057. https://doi.org/10.3168/jds.2010-3633.

Bach, A. 2012. Ruminant Nutrition Symposium: Optimizing performance of the offspring: Nourishing and managing the dam and postnatal calf for optimal lactation, reproduction, and immunity. Journal of Animal Science 90(6):1835-1845. https://doi.org/10.2527/jas.2011-4516.

Beggs, D. S., E. C. Jongman, P. H. Hemsworth, and A. D. Fisher. 2019. Lame cows on Australian dairy farms: A comparison of farmer-identified lameness and formal lameness scoring, and the position of lame cows within the milking order. Journal of Dairy Science 102(2):1522-1529. https://doi.org/10.3168/jds.2018-14847.

Bell, M. J. and P. Wilson. 2018. Estimated differences in economic and environmental performance of forage-based dairy herds across the UK. Food and Energy Security 7(1):e00127. https://doi.org/10.1002/fes3.127.

Benbrook, C., C. Carman, E. A. Clark, C. Daley, W. Fulwider, M. Hansen, C. Leifert, K. Martens, L. Paine, L. Petkewitz, G. Jodarski, F. Thicke, J. Velez, and G. Wegner. 2010. A dairy farm's footprint: Evaluating the impacts of conventional and organic farming systems. Critical Issue Report. The Organic Center, Washington D.C., USA.

Berry, D. P. 2015. Breeding the dairy cow of the future: What do we need? Animal Production Science 55(7):823-837. https://doi.org/10.1071/AN14835.

Berry, D. P., B. L. Harris, A. M. Winkelman, and W. Montgomerie. 2005. Phenotypic associations between traits other than production and longevity in New Zealand dairy cattle. Journal of Dairy Science 88(8):2962-2974. https://doi.org/10.3168/jds.S0022-0302(05)72976-3.

Berry, D. P., P. Lonergan, S. T. Butler, A. R. Cromie, T. Fair, F. Mossa, and A. C. O. Evans. 2008. Negative influence of high maternal milk production before and after conception on offspring survival and milk production in dairy cattle. Journal of Dairy Science 91(1):329-337. https://doi.org/10.3168/jds.2007-0438.

Bond, G. B., R. d. Almeida, A. Ostrensky, and C. F. M. Molento. 2012. Métodos de diagnóstico e pontos críticos de bem-estar de bovinos leiteiros. Ciência Rural 42(7):1286-1293. https://doi.org/10.1590/S0103-84782012005000044.

Bouffard, V., A. M. de Passillé, J. Rushen, E. Vasseur, C. G. R. Nash, D. B. Haley, and D. Pellerin. 2017. Effect of following recommendations for tiestall configuration on neck and leg lesions, lameness, cleanliness, and lying time in dairy cows. Journal of Dairy Science 100(4):2935-2943. https://doi.org/10.3168/jds.2016-11842.

Boulton, A. C., J. Rushton, and D. C. Wathes. 2017. An empirical analysis of the cost of rearing dairy heifers from birth to first calving and the time taken to repay these costs. Animal 11(8):1372-1380. https://doi.org/10.1017/S1751731117000064.

Brickell, J. S. and D. C. Wathes. 2011. A descriptive study of the survival of Holstein-Friesian heifers through to third calving on English dairy farms. Journal of Dairy Science 94(4):1831-1838. https://doi.org/10.3168/jds.2010-3710.

Britt, J. H., R. A. Cushman, C. D. Dechow, H. Dobson, P. Humblot, M. F. Hutjens, G. A. Jones, P. S. Ruegg, I. M. Sheldon, and J. S. Stevenson. 2018. Invited review: Learning from the future—

A vision for dairy farms and cows in 2067. Journal of Dairy Science 101(5):3722-3741. https://doi.org/10.3168/jds.2017-14025.

BRS – Bundesverband Rind und Schwein e.V. 2019. Rinder- und Schweineproduktion in Deutschland. BRS, Bonn, Germany.

Bruijnis, M. R. N., F. L. B. Meijboom, and E. N. Stassen. 2013. Longevity as an animal welfare issue applied to the case of foot disorders in dairy cattle. Journal of Agricultural and Environmental Ethics 26(1):191-205. https://doi.org/10.1007/s10806-012-9376-0.

Bryant, J. R., N. López-Villalobos, J. E. Pryce, C. W. Holmes, and D. L. Johnson. 2007. Quantifying the effect of thermal environment on production traits in three breeds of dairy cattle in New Zealand. New Zealand Journal of Agricultural Research 50(3):327-338. https://doi.org/10.1080/00288230709510301.

Burnett, T. A., A. M. L. Madureira, B. F. Silper, A. Tahmasbi, A. Nadalin, D. M. Veira, and R. L. A. Cerri. 2015. Relationship of concentrations of cortisol in hair with health, biomarkers in blood, and reproductive status in dairy cows. Journal of Dairy Science 98(7):4414-4426. https://doi.org/10.3168/jds.2014-8871.

Busanello, M., R. S. Rossi, L. D. Cassoli, J. C. F. Pantoja, and P. F. Machado. 2017. Estimation of prevalence and incidence of subclinical mastitis in a large population of Brazilian dairy herds. Journal of Dairy Science 100(8):6545-6553. https://doi.org/10.3168/jds.2016-12042.

Butler, W. R., J. J. Calaman, and S. W. Beam. 1996. Plasma and milk urea nitrogen in relation to pregnancy rate in lactating dairy cattle. Journal of Animal Science 74(4):858-865. https://doi.org/10.2527/1996.744858x.

Cardoso, C. S., M. J. Hötzel, D. M. Weary, J. A. Robbins, and M. A. G. von Keyserlingk. 2016. Imagining the ideal dairy farm. Journal of Dairy Science 99(2):1663-1671. https://doi.org/10.3168/jds.2015-9925.

Cavestany, D., A. B. El-Wishy, and R. H. Foote. 1985. Effect of season and high environmental temperature on fertility of Holstein cattle. Journal of Dairy Science 68(6):1471-1478. https://doi.org/10.3168/jds.S0022-0302(85)80985-1.

CDIC – Canadian Dairy Information Centre. 2020a. Culling and replacement rates in dairy herds in Canada. Accessed on 01 October 2020. https://www.dairyinfo.gc.ca/eng/dairy-statistics-and-market-information/dairy-animal-genetics/culling-and-replacement-rates-in-dairy-herds-in-canada/?id=1502475693224.

CDIC – Canadian Dairy Information Centre. 2020b. Number of farms, dairy cows and dairy heifers. Accessed on 20 April 2020. https://dairyinfo.gc.ca/eng/dairy-statistics-and-market-information/farm-statistics/farms-dairy-cows-and-dairy-heifers/?id=1502467423238&menupos=01.01.14.

Chester-Jones, H., B. J. Heins, D. Ziegler, D. Schimek, S. Schuling, B. Ziegler, M. B. de Ondarza, C. J. Sniffen, and N. Broadwater. 2017. Relationships between early-life growth, intake, and birth season with first-lactation performance of Holstein dairy cows. Journal of Dairy Science 100(5):3697-3704. https://doi.org/10.3168/jds.2016-12229.

Chuck, G. M., P. D. Mansell, M. A. Stevenson, and M. M. Izzo. 2018. Early-life events associated with first-lactation performance in pasture-based dairy herds. Journal of Dairy Science 101(4):3488-3500. https://doi.org/10.3168/jds.2017-12626.

Clasen, J. B., J. O. Lehmann, J. R. Thomasen, S. Østergaard, and M. Kargo. 2019. Combining extended lactation with sexed semen in a dairy cattle herd: Effect on genetic and total economic return. Livestock Science 223:176-183. https://doi.org/10.1016/j.livsci.2019.03.001.

Collier, R. J., G. E. Dahl, and M. J. VanBaale. 2006. Major advances associated with environmental effects on dairy cattle. Journal of Dairy Science 89(4):1244-1253. https://doi.org/10.3168/jds.S0022-0302(06)72193-2.

Compton, C. W. R., C. Heuer, P. T. Thomsen, T. E. Carpenter, C. V. C. Phyn, and S. McDougall. 2017. Invited review: A systematic literature review and meta-analysis of mortality and culling in dairy cattle. Journal of Dairy Science 100(1):1-16. https://doi.org/10.3168/jds.2016-11302.

CRV – Coöperatie Koninklijke CRV. 2020. Jaarstatistieken 2019 voor Nederland [Annual statistics 2019 for the Netherlands]. Arnhem, Netherlands. Accessed on 20 April 2020. https://cooperatiecrv-be6.kxcdn.com/wp-content/uploads/2020/02/Jaarstatistieken-2019-NL-totaal.pdf.

DairyNZ. 2020. Animal evaluation: Animal and herd averages. Accessed on 29 May 2020. https://www.dairynz.co.nz/animal/animal-evaluation/animal-and-herd-averages.

De Vries, A. 2013. Cow longevity economics: The cost benefit of keeping the cow in the herd. Pages 22-52 in Proc. Cow Longevity Conference. DeLaval, Hamra Farm, Tumba, Sweeden.

De Vries, A. 2017. Economic trade-offs between genetic improvement and longevity in dairy cattle. Journal of Dairy Science 100(5):4184-4192. https://doi.org/10.3168/jds.2016-11847.

De Vries, A. 2020. Symposium review: Why revisit dairy cattle productive lifespan? Journal of Dairy Science 103(4):3838-3845. https://doi.org/10.3168/jds.2019-17361.

De Vries, A. and M. I. Marcondes. 2020. Review: Overview of factors affecting productive lifespan of dairy cows. Animal 14(S1):s155-s164. https://doi.org/10.1017/S1751731119003264.

De Wit, A. A. C., M. L. F. Cesar, and T. A. M. Kruip. 2001. Effect of urea during *in vitro* maturation on nuclear maturation and embryo development of bovine cumulus-oocyte-complexes. Journal of Dairy Science 84(8):1800-1804. https://doi.org/10.3168/jds.S0022-0302(01)74618-8.

DeVries, T. J., M. G. Aarnoudse, H. W. Barkema, K. E. Leslie, and M. A. G. von Keyserlingk. 2012. Associations of dairy cow behavior, barn hygiene, cow hygiene, and risk of elevated somatic cell count. Journal of Dairy Science 95(10):5730-5739. https://doi.org/10.3168/jds.2012-5375.

Dhuyvetter, K., T. Kastens, M. Overton, and J. Smith. 2007. Cow culling decisions: Costs or economic opportunity? Pages 173-187 in Proc. Western Dairy Management Conference, Reno, Nevada, US.

Du, M., S. P. Ford, and M.-J. Zhu. 2017. Optimizing livestock production efficiency through maternal nutritional management and fetal developmental programming. Animal Frontiers 7(3):5-11. https://doi.org/10.2527/af.2017-0122.

Du, M., J. Tong, J. Zhao, K. R. Underwood, M. Zhu, S. P. Ford, and P. W. Nathanielsz. 2010. Fetal programming of skeletal muscle development in ruminant animals. Journal of Animal Science 88(suppl_13):E51-E60. https://doi.org/10.2527/jas.2009-2311.

Dubuc, J., T. F. Duffield, K. E. Leslie, J. S. Walton, and S. J. LeBlanc. 2011. Effects of postpartum uterine diseases on milk production and culling in dairy cows. Journal of Dairy Science 94(3):1339-1346. https://doi.org/10.3168/jds.2010-3758.

Ducrocq, V. 1994. Statistical analysis of length of productive life for dairy cows of the Normande breed. Journal of Dairy Science 77(3):855-866. https://doi.org/10.3168/jds.S0022-0302(94)77020-X.

Espejo, L. A., M. I. Endres, and J. A. Salfer. 2006. Prevalence of lameness in high-producing Holstein cows housed in freestall barns in Minnesota. Journal of Dairy Science 89(8):3052-3058. https://doi.org/10.3168/jds.S0022-0302(06)72579-6.

Essl, A. 1998. Longevity in dairy cattle breeding: A review. Livestock Production Science 57(1):79-89. https://doi.org/10.1016/S0301-6226(98)00160-2.

European Commission. 2016. Attitudes of Europeans towards animal welfare. Special Eurobarometer 442. European Commission, Brussels, Belgium. https://doi.org/10.2875/884639.

Fabian, J., R. A. Laven, and H. R. Whay. 2014. The prevalence of lameness on New Zealand dairy farms: A comparison of farmer estimate and locomotion scoring. The Veterinary Journal 201(1):31-38. https://doi.org/10.1016/j.tvjl.2014.05.011.

FAO – Food and Agriculture Organization of the United Nations. 2020. Production: Livestock primary. Accessed on 02 April 2020. http://www.fao.org/faostat/en/#data/QL.

Fetrow, J., K. V. Nordlund, and H. D. Norman. 2006. Invited Review: Culling: Nomenclature, definitions, and recommendations. Journal of Dairy Science 89(6):1896-1905. https://doi.org/10.3168/jds.S0022-0302(06)72257-3.

Firth, C. L., C. Laubichler, C. Schleicher, K. Fuchs, A. Käsbohrer, C. Egger-Danner, J. Köfer, and W. Obritzhauser. 2019. Relationship between the probability of veterinary-diagnosed bovine mastitis occurring and farm management risk factors on small dairy farms in Austria. Journal of Dairy Science 102(5):4452-4463. https://doi.org/10.3168/jds.2018-15657.

Gilbert, R. O., S. T. Shin, C. L. Guard, H. N. Erb, and M. Frajblat. 2005. Prevalence of endometritis and its effects on reproductive performance of dairy cows. Theriogenology 64(9):1879-1888. https://doi.org/10.1016/j.theriogenology.2005.04.022.

GIROLANDO – Associação Brasileira dos Criadores de Girolando. 2020. Evolução dos índices zootécnicos do gado girolando. Personal Communication. Accessed on April 29, 2020.

Giuliodori, M. J., R. P. Magnasco, D. Becu-Villalobos, I. M. Lacau-Mengido, C. A. Risco, and R. L. de la Sota. 2013. Metritis in dairy cows: Risk factors and reproductive performance. Journal of Dairy Science 96(6):3621-3631. https://doi.org/10.3168/jds.2012-5922.

Grandl, F., S. L. Amelchanka, M. Furger, M. Clauss, J. O. Zeitz, M. Kreuzer, and A. Schwarm. 2016. Biological implications of longevity in dairy cows: 2. Changes in methane emissions and efficiency with age. Journal of Dairy Science 99(5):3472-3485. https://doi.org/10.3168/jds.2015-10262.

Grandl, F., M. Furger, M. Kreuzer, and M. Zehetmeier. 2019. Impact of longevity on greenhouse gas emissions and profitability of individual dairy cows analysed with different system boundaries. Animal 13(1):198-208. https://doi.org/10.1017/S175173111800112X.

Haine, D., H. Delgado, R. Cue, A. Sewalem, K. Wade, R. Lacroix, D. Lefebvre, J. Arsenault, É. Bouchard, and J. Dubuc. 2017. Contextual herd factors associated with cow culling risk in Québec dairy herds: A multilevel analysis. Preventive Veterinary Medicine 144:7-12. https://doi.org/10.1016/j.prevetmed.2017.05.014.

Haskell, M. J., L. J. Rennie, V. A. Bowell, M. J. Bell, and A. B. Lawrence. 2006. Housing system, milk production, and zero-grazing effects on lameness and leg injury in dairy cows. Journal of Dairy Science 89(11):4259-4266. https://doi.org/http://dx.doi.org/10.3168/jds.S0022-0302(06)72472-9.

Haworth, G. M., W. P. Tranter, J. N. Chuck, Z. Cheng, and D. C. Wathes. 2008. Relationships between age at first calving and first lactation milk yield, and lifetime productivity and longevity in dairy cows. Veterinary Record 162(20):643-647. https://doi.org/10.1136/vr.162.20.643.

Heinrichs, A. J. and B. S. Heinrichs. 2011. A prospective study of calf factors affecting first-lactation and lifetime milk production and age of cows when removed from the herd. Journal of Dairy Science 94(1):336-341. https://doi.org/10.3168/jds.2010-3170.

Heise, J., Z. Liu, K. F. Stock, S. Rensing, F. Reinhardt, and H. Simianer. 2016. The genetic structure of longevity in dairy cows. Journal of Dairy Science 99(2):1253-1265. https://doi.org/10.3168/jds.2015-10163.

Hernandez-Mendo, O., M. A. G. von Keyserlingk, D. M. Veira, and D. M. Weary. 2007. Effects of pasture on lameness in dairy cows. Journal of Dairy Science 90(3):1209-1214. https://doi.org/10.3168/jds.S0022-0302(07)71608-9.

Higginson Cutler, J. H., J. Rushen, A. M. de Passillé, J. Gibbons, K. Orsel, E. Pajor, H. W. Barkema, L. Solano, D. Pellerin, D. Haley, and E. Vasseur. 2017. Producer estimates of prevalence and perceived importance of lameness in dairy herds with tiestalls, freestalls, and automated milking systems. Journal of Dairy Science 100(12):9871-9880. https://doi.org/10.3168/jds.2017-13008.

Hultgren, J. and C. Svensson. 2009. Heifer rearing conditions affect length of productive life in Swedish dairy cows. Preventive Veterinary Medicine 89(3):255-264. https://doi.org/10.1016/j.prevetmed.2009.02.012.

Hultgren, J. and C. Svensson. 2010. Calving interval in dairy cows in relation to heifer rearing conditions in southwest Sweden. Reproduction in Domestic Animals 45(1):136-141. https://doi.org/10.1111/j.1439-0531.2008.01273.x.

Huxley, J. N. 2013. Impact of lameness and claw lesions in cows on health and production. Livestock Science 156(1):64-70. https://doi.org/10.1016/j.livsci.2013.06.012.

ICBF - Irish Cattle Breeding Federation – Irish Cattle Breeding Federation. 2020. Age at first calving – Ireland. Personal Communication. Accessed on April 21, 2020.

idele – Institut de l'Elevage. 2019. Resultats de controle laitier des especes bovine, caprine et ovine: France 2018. Institut de l'Elevage & France Conseil Elevage, Paris, France. Accessed on 21 April 2020. http://idele.fr/no_cache/recherche/publication/idelesolr/recommends/resultats-de-controle-laitier-france-2018.html.

Knapp, J. R., G. L. Laur, P. A. Vadas, W. P. Weiss, and J. M. Tricarico. 2014. Invited review: Enteric methane in dairy cattle production: Quantifying the opportunities and impact of reducing emissions. Journal of Dairy Science 97(6):3231-3261. https://doi.org/10.3168/jds.2013-7234.

Komlósi, I., M. Wolfová, J. Wolf, B. Farkas, Z. Szendrei, and B. Béri. 2010. Economic weights of production and functional traits for Holstein-Friesian cattle in Hungary. Journal of Animal Breeding and Genetics 127(2):143-153. https://doi.org/10.1111/j.1439-0388.2009.00822.x.

Kowsar, R., F. Izadi, N. Sadeghi, A. Riasi, F. G. Zadegan, M. Hajian, M. H. Nasr-Esfahani, H. Farrokhpour, and A. Miyamoto. 2018. Urea changes oocyte competence and gene expression in resultant bovine embryo in vitro. Zygote 26(3):207-219. https://doi.org/10.1017/S0967199418000102.

Krömker, V., F. Pfannenschmidt, K. Helmke, R. Andersson, and N. T. Grabowski. 2012. Risk factors for intramammary infections and subclinical mastitis in post-partum dairy heifers. Journal of Dairy Research 79(3):304-309. https://doi.org/10.1017/S0022029912000222.

Kumar, N., A. Manimaran, A. Kumaresan, S. Jeyakumar, L. Sreela, P. Mooventhan, and M. Sivaram. 2017. Mastitis effects on reproductive performance in dairy cattle: A review. Tropical Animal Health and Production 49(4):663-673. https://doi.org/10.1007/s11250-017-1253-4.

Lactanet. 2020a. 2019 Annual Report & 2020 Business Plan. Accessed on 18 April 2020. http://canwestdhi.com/pdf_files/lactanet%20annual%20report.pdf.

Lactanet. 2020b. Herd benchmarks: Management centre. Accessed on 18 April 2020. http://canwestdhi.com/pdf_files/2019%20benchmarks/2019%20Management%20Centre%20Canada%20Lactanet%20Herds%20(English).pdf.

Langford, F. M. and A. W. Stott. 2012. Culled early or culled late: economic decisions and risks to welfare in dairy cows. Animal Welfare 21(1):41-55. https://doi.org/10.7120/096272812X13345905673647.

Mark, T. 2004. Applied genetic evaluations for production and functional traits in dairy cattle. Journal of Dairy Science 87(8):2641-2652. https://doi.org/10.3168/jds.S0022-0302(04)73390-1.

McDougall, S., D. Aberdein, A. Bates, and C. R. Burke. 2020. Prevalence of endometritis diagnosed by vaginal discharge scoring or uterine cytology in dairy cows and herds. Journal of Dairy Science. https://doi.org/10.3168/jds.2019-18048.

McDougall, S., D. Arthur, M. Bryan, J. Vermunt, and A. Weir. 2007. Clinical and bacteriological response to treatment of clinical mastitis with one of three intramammary antibiotics. New Zealand Veterinary Journal 55(4):161-170. https://doi.org/10.1080/00480169.2007.36762.

Molento, C. F. M. 2005. Bem-estar e produção animal: Aspectos econômicos. Archives of Veterinary Science 10(1):1-11. https://doi.org/10.5380/avs.v10i1.4078.

Moreira, T. F., R. R. Nicolino, R. M. Meneses, G. V. Fonseca, L. M. Rodrigues, E. J. Facury Filho, and A. U. Carvalho. 2019. Risk factors associated with lameness and hoof lesions in pasture-based dairy cattle systems in southeast Brazil. Journal of Dairy Science 102(11):10369-10378. https://doi.org/10.3168/jds.2018-16215.

Mulder, H. and G. Jansen. 2001. Derivation of economic values using lifetime profitability of Canadian Holstein cows. Parity 1(3):1-10.

Nielsen, B. H., P. T. Thomsen, and J. T. Sørensen. 2011. Identifying risk factors for poor hind limb cleanliness in Danish loose-housed dairy cows. Animal 5(10):1613-1619. https://doi.org/10.1017/S1751731111000905.

Nilforooshan, M. A. and M. A. Edriss. 2004. Effect of age at first calving on some productive and longevity traits in Iranian Holsteins of the Isfahan province. Journal of Dairy Science 87(7):2130-2135. https://doi.org/10.3168/jds.S0022-0302(04)70032-6.

Niozas, G., G. Tsousis, C. Malesios, I. Steinhöfel, C. Boscos, H. Bollwein, and M. Kaske. 2019a. Extended lactation in high-yielding dairy cows. II. Effects on milk production, udder health, and body measurements. Journal of Dairy Science 102(1):811-823. https://doi.org/10.3168/jds.2018-15117.

Niozas, G., G. Tsousis, I. Steinhöfel, C. Brozos, A. Römer, S. Wiedemann, H. Bollwein, and M. Kaske. 2019b. Extended lactation in high-yielding dairy cows. I. Effects on reproductive measurements. Journal of Dairy Science 102(1):799-810. https://doi.org/10.3168/jds.2018-15115.

Olde Riekerink, R. G. M., H. W. Barkema, D. F. Kelton, and D. T. Scholl. 2008. Incidence rate of clinical mastitis on Canadian dairy farms. Journal of Dairy Science 91(4):1366-1377. https://doi.org/10.3168/jds.2007-0757.

Olmos, G., L. Boyle, A. Hanlon, J. Patton, J. J. Murphy, and J. F. Mee. 2009a. Hoof disorders, locomotion ability and lying times of cubicle-housed compared to pasture-based dairy cows. Livestock Science 125(2):199-207. https://doi.org/10.1016/j.livsci.2009.04.009.

Olmos, G., J. F. Mee, A. Hanlon, J. Patton, J. J. Murphy, and L. Boyle. 2009b. Peripartum health and welfare of Holstein-Friesian cows in a confinement-TMR system compared to a pasture-based system. Animal Welfare 18(4):467-476.

Opsomer, G., M. Van Eetvelde, M. Kamal, and A. Van Soom. 2017. Epidemiological evidence for metabolic programming in dairy cattle. Reproduction, Fertility and Development 29(1):52-57. https://doi.org/10.1071/RD16410.

Pellerin, D., S. Adams, F. Bécotte, R. Cue, R. Moore, and R. Roy. 2014. Pour une vache, l'âge d'or c'est la 4e lactation! [For a cow, the golden age is the fourth lactation!]. Pages 133-147 in Proc. Symposium sur les bovins laitiers: Choix d'aujourd'hui pour les défis de demain. CRAAQ – Le Centre de référence en agriculture et agroalimentaire du Québec, Saint-Hyacinthe, QC, Canada.

PFHBIPM – Polska Federacja Hodowców Bydła i Producentów Mleka. 2020. Mlecznego ocena i hodowla bydła mlecznego: Dane za rok 2019. Warsaw, Poland. Accessed on 18 April 2020. http://pfhb.pl/fileadmin/user_upload/OCENA/publikacje/publikacje_2019/biuletyn_2018_calosc_na_www_02.pdf.

Pinedo, P. J., A. Daniels, J. Shumaker, and A. De Vries. 2014. Dynamics of culling for Jersey, Holstein, and Jersey × Holstein crossbred cows in large multibreed dairy herds. Journal of Dairy Science 97(5):2886-2895. https://doi.org/10.3168/jds.2013-7685.

Pinedo, P. J., A. De Vries, and D. W. Webb. 2010. Dynamics of culling risk with disposal codes reported by Dairy Herd Improvement dairy herds. Journal of Dairy Science 93(5):2250-2261. https://doi.org/10.3168/jds.2009-2572.

Polsky, L. and M. A. G. von Keyserlingk. 2017. Invited review: Effects of heat stress on dairy cattle welfare. Journal of Dairy Science 100(11):8645-8657. https://doi.org/10.3168/jds.2017-12651.

Pötzsch, C. J., V. J. Collis, R. W. Blowey, A. J. Packington, and L. E. Green. 2003. The Impact of Parity and Duration of Biotin Supplementation on White Line Disease Lameness in Dairy Cattle. Journal of Dairy Science 86(8):2577-2582. https://doi.org/10.3168/jds.S0022-0302(03)73852-1.

Pritchard, T., M. Coffey, R. Mrode, and E. Wall. 2013. Understanding the genetics of survival in dairy cows. Journal of Dairy Science 96(5):3296-3309. https://doi.org/10.3168/jds.2012-6219.

Rajala-Schultz, P. J., W. J. A. Saville, G. S. Frazer, and T. E. Wittum. 2001. Association between milk urea nitrogen and fertility in Ohio dairy cows. Journal of Dairy Science 84(2):482-489. https://doi.org/10.3168/jds.S0022-0302(01)74498-0.

Ranjbar, S., A. R. Rabiee, A. Gunn, and J. K. House. 2016. Identifying risk factors associated with lameness in pasture-based dairy herds. Journal of Dairy Science 99(9):7495-7505. https://doi.org/10.3168/jds.2016-11142.

Reneau, J. K., A. J. Seykora, B. J. Heins, M. I. Endres, R. J. Farnsworth, and R. F. Bey. 2005. Association between hygiene scores and somatic cell scores in dairy cattle. Journal of the American Veterinary Medical Association 227(8):1297-1301. https://doi.org/10.2460/javma.2005.227.1297.

Ribeiro, E. S., F. S. Lima, L. F. Greco, R. S. Bisinotto, A. P. A. Monteiro, M. Favoreto, H. Ayres, R. S. Marsola, N. Martinez, W. W. Thatcher, and J. E. P. Santos. 2013. Prevalence of periparturient diseases and effects on fertility of seasonally calving grazing dairy cows supplemented with concentrates. Journal of Dairy Science 96(9):5682-5697. https://doi.org/10.3168/jds.2012-6335.

Rocha, A., R. D. Randel, J. R. Broussard, J. M. Lim, R. M. Blair, J. D. Roussel, R. A. Godke, and W. Hansel. 1998. High environmental temperature and humidity decrease oocyte quality in *Bos taurus* but not in *Bos indicus* cows. Theriogenology 49(3):657-665. https://doi.org/10.1016/S0093-691X(98)00016-8.

Roche, S. M., D. L. Renaud, R. Genore, D. A. Shock, C. Bauman, S. Croyle, H. W. Barkema, J. Dubuc, G. P. Keefe, and D. F. Kelton. 2020. Canadian National Dairy Study: Describing Canadian dairy producer practices and perceptions surrounding cull cow management. Journal of Dairy Science 103(4):3414-3421. https://doi.org/10.3168/jds.2019-17390.

Rushen, J. and A. M. Passillé. 2013. The importance of improving cow longevity. Pages 03-21 in Proc. Cow Longevity Conference. DeLaval, Hamra Farm, Tumba, Sweeden.

Sant'anna, A. C. and M. J. Paranhos da Costa. 2011. The relationship between dairy cow hygiene and somatic cell count in milk. Journal of Dairy Science 94(8):3835-3844. https://doi.org/10.3168/jds.2010-3951.

Santman-Berends, I. M. G. A., J. M. Swinkels, T. J. G. M. Lam, J. Keurentjes, and G. van Schaik. 2016. Evaluation of udder health parameters and risk factors for clinical mastitis in Dutch dairy herds in the context of a restricted antimicrobial usage policy. Journal of Dairy Science 99(4):2930-2939. https://doi.org/10.3168/jds.2015-10398.

Schneider, M. d. P., E. Strandberg, U. Emanuelson, K. Grandinson, and A. Roth. 2007. The effect of veterinary-treated clinical mastitis and pregnancy status on culling in Swedish dairy cows. Preventive Veterinary Medicine 80(2):179-192. https://doi.org/10.1016/j.prevetmed.2007.02.006.

Schreiner, D. A. and P. L. Ruegg. 2003. Relationship between udder and leg hygiene scores and subclinical mastitis. Journal of Dairy Science 86(11):3460-3465. https://doi.org/10.3168/jds.S0022-0302(03)73950-2.

Sewalem, A., G. J. Kistemaker, F. Miglior, and B. J. Van Doormaal. 2006. Analysis of inbreeding and its relationship with functional longevity in Canadian dairy cattle. Journal of Dairy Science 89(6):2210-2216. https://doi.org/10.3168/jds.S0022-0302(06)72291-3.

Sewalem, A., F. Miglior, G. J. Kistemaker, P. Sullivan, and B. J. Van Doormaal. 2008. Relationship between reproduction traits and functional longevity in Canadian dairy cattle. Journal of Dairy Science 91(4):1660-1668. https://doi.org/10.3168/jds.2007-0178.

Sherwin, V. E., C. D. Hudson, A. Henderson, and M. J. Green. 2016. The association between age at first calving and survival of first lactation heifers within dairy herds. Animal 10(11):1877-1882. https://doi.org/10.1017/S1751731116000689.

Shook, G. E. 2006. Major advances in determining appropriate selection goals. Journal of Dairy Science 89(4):1349-1361. https://doi.org/10.3168/jds.S0022-0302(06)72202-0.

Somers, J. G. C. J., K. Frankena, E. N. Noordhuizen-Stassen, and J. H. M. Metz. 2003. Prevalence of claw disorders in Dutch dairy cows exposed to several floor systems. Journal of Dairy Science 86(6):2082-2093. https://doi.org/10.3168/jds.S0022-0302(03)73797-7.

Stojkov, J., M. A. G. von Keyserlingk, T. Duffield, and D. Fraser. 2020. Management of cull dairy cows: Culling decisions, duration of transport, and effect on cow condition. Journal of Dairy Science 103(3):2636-2649. https://doi.org/10.3168/jds.2019-17435.

Svensson, C. and P. Liberg. 2006. The effect of group size on health and growth rate of Swedish dairy calves housed in pens with automatic milk-feeders. Preventive Veterinary Medicine 73(1):43-53. https://doi.org/10.1016/j.prevetmed.2005.08.021.

Svensson, C., A. K. Nyman, K. P. Waller, and U. Emanuelson. 2006. Effects of housing, management, and health of dairy heifers on first-lactation udder health in southwest Sweden. Journal of Dairy Science 89(6):1990-1999. https://doi.org/10.3168/jds.S0022-0302(06)72266-4.

Thompson, A. J., D. M. Weary, J. A. Bran, R. R. Daros, M. J. Hötzel, and M. A. G. von Keyserlingk. 2019. Lameness and lying behavior in grazing dairy cows. Journal of Dairy Science 102(7):6373-6382. https://doi.org/10.3168/jds.2018-15717.

Vacek, M., L. Krpálková, J. Syrůček, M. Štípková, and M. Janecká. 2015. Relationships between growth and body condition development during the rearing period and performance in the first three lactations in Holstein cows. Czech Journal of Animal Science 60(9):417-425. https://doi.org/10.17221/8460-CJAS.

Valacta. 2019. The evolution of Valacta Atlantic dairy production: 2018 Stats & Tips.

van den Borne, B. H. P., G. van Schaik, T. J. G. M. Lam, and M. Nielen. 2010. Variation in herd level mastitis indicators between primi- and multiparae in Dutch dairy herds. Preventive Veterinary Medicine 96(1):49-55. https://doi.org/10.1016/j.prevetmed.2010.05.010.

Van Doormaal, B. 2009. A closer look at longevity. Accessed on 28 October 2020. https://www.cdn.ca/document.php?id=162.

van Pelt, M. 2017. Genetic improvement of longevity in dairy cows. Ph.D. thesis. Wageningen University, Wageningen, the Netherlands. https://doi.org/10.18174/404192.

Vanegas, J., M. Overton, S. L. Berry, and W. M. Sischo. 2006. Effect of rubber flooring on claw health in lactating dairy cows housed in free-stall barns. Journal of Dairy Science 89(11):4251-4258. https://doi.org/10.3168/jds.S0022-0302(06)72471-7.

Villettaz Robichaud, M., J. Rushen, A. M. de Passillé, E. Vasseur, D. Haley, K. Orsel, and D. Pellerin. 2018. Is the profitability of Canadian freestall farms associated with their performance on an animal welfare assessment? Journal of Dairy Science 101(3):2350-2358. https://doi.org/10.3168/jds.2017-13315.

Villettaz Robichaud, M., J. Rushen, A. M. de Passillé, E. Vasseur, D. Haley, and D. Pellerin. 2019a. Associations between on-farm cow welfare indicators and productivity and profitability on Canadian dairies: II. On tiestall farms. Journal of Dairy Science 102(5):4352-4363. https://doi.org/10.3168/jds.2018-14818.

Villettaz Robichaud, M., J. Rushen, A. M. de Passillé, E. Vasseur, K. Orsel, and D. Pellerin. 2019b. Associations between on-farm animal welfare indicators and productivity and profitability on Canadian dairies: I. On freestall farms. Journal of Dairy Science 102(5):4341-4351. https://doi.org/10.3168/jds.2018-14817.

von Keyserlingk, M. A. G., J. Rushen, A. M. de Passillé, and D. M. Weary. 2009. Invited review: The welfare of dairy cattle - Key concepts and the role of science. Journal of Dairy Science 92(9):4101-4111. https://doi.org/10.3168/jds.2009-2326.

Wall, E., M. P. Coffey, and G. E. Pollott. 2012. The effect of lactation length on greenhouse gas emissions from the national dairy herd. Animal 6(11):1857-1867. https://doi.org/10.1017/S1751731112000936.

Walter, A., R. Finger, R. Huber, and N. Buchmann. 2017. Opinion: Smart farming is key to developing sustainable agriculture. Proceedings of the National Academy of Sciences 114(24):6148. https://doi.org/10.1073/pnas.1707462114.

Washburn, S. P., S. L. White, J. T. Green, and G. A. Benson. 2002. Reproduction, mastitis, and body condition of seasonally calved Holstein and Jersey cows in confinement or pasture systems. Journal of Dairy Science 85(1):105-111. https://doi.org/10.3168/jds.S0022-0302(02)74058-7.

Weigel, K. A., R. W. Palmer, and D. Z. Caraviello. 2003. Investigation of factors affecting voluntary and involuntary culling in expanding dairy herds in Wisconsin using survival analysis. Journal of Dairy Science 86(4):1482-1486. https://doi.org/10.3168/jds.S0022-0302(03)73733-3.

Wolf, C. A. and G. T. Tonsor. 2017. Cow welfare in the U.S. dairy industry: Willingness-to-pay and willingness-to-supply. Journal of Agricultural and Resource Economics 42(2):164-179. https://doi.org/10.22004/ag.econ.257996.

2.10. Supplementary material

Supplementary Table 2.1. The sources of official statistics published by the top 21 high milk-producing countries regarding the number of dairy cows, total milk production, average milk yield per animal per year, and the number of slaughtered cows as well as data availability information.

Country	Number of dairy cows	Total milk production	Milk yield	Number of slaughtered cows	Specify the dairy category
United States of America	FAO (2020), USDA (2020a)	USDA (2020a)	USDA (2020a) and calculated	USDA (2020b)	Yes
India	---	---	---	Illegal to slaughter cows in most parts of the country (Narayanan, 2019)	---
Brazil	FAO (2020), IBGE (2020b)	FAO (2020), IBGE (2020a)	Calculated	IBGE (2020c)	No
Germany	Destatis (2020a)	BZL (2020), eurostat (2020b)	BZL (2020) and calculated	Destatis (2020b)	No
China, mainland	---	---	---	Reports the number of slaughtered cattle and buffaloes combined and does not specify cows	---
Russian Federation	---	---	---	Reports only the total live weight of slaughtered cattle and does not specify cows	---
France	eurostat (2020a)	eurostat (2020b)	Calculated	eurostat (2020c)	No
New Zealand	LIC & DairyNZ (2019), FAO (2020), NZ.Stat (2020b)	LIC & DairyNZ (2019), FAO (2020)	LIC & DairyNZ (2019) and calculated	NZ.Stat (2020a)	No
Turkey	---	---	---	Only available for years 2015 to 2019 (eurostat, 2020c)	---
Pakistan	---	---	---	Information not freely available	---
United Kingdom	---	---	---	Not reliable over the years (eurostat, 2020c)	---
Poland	eurostat (2020a), FAO (2020)	eurostat (2020b), FAO (2020)	Calculated	eurostat (2020c)	No
Netherlands	eurostat (2020a)	eurostat (2020b)	Calculated	eurostat (2020c)	No
Mexico	FAO (2020)	FAO (2020), SIACON (2020)	Calculated	USDA (2020c)	No
Italy	eurostat (2020a)	eurostat (2020b)	Calculated	eurostat (2020c)	No
Argentina	FAO (2020)	MAGyP (2020b)	Calculated	MAGyP (2020a)	No
Ukraine	---	---	Calculated	Reports the number of slaughtered cattle	---
Uzbekistan	---	---	Calculated	Reports not available	---
Australia	---	---	---	Reports the number of slaughtered cows and heifers together	---

Canada	Statistics Canada (2020b)	Statistics Canada (2020a)	Calculated	USDA (2020c)	No
Ireland	eurostat (2020a), FAO (2020)	eurostat (2020b)	Calculated	eurostat (2020c)	No

References from Supplementary Table 2.1

BZL – Bundesinformationszentrum Landwirtschaft. 2020. Milcherzeugung und -verwendung 2019. Accessed on 02 April 2020. https://www.ble.de/DE/BZL/Daten-Berichte/Milch-Milcherzeugnisse/_functions/TabelleMilcherzeugungVerwendung2019.html?nn=8906974.

Destatis – Statistisches Bundesamt. 2020a. Gehaltene tiere: Deutschland, jahre, tierarten. Accessed on 02 April 2020. https://www-genesis.destatis.de/genesis//online?operation=table&code=41311-0001&levelindex=0&levelid=1585877656772.

Destatis – Statistisches Bundesamt. 2020b. Geschlachtete tiere, schlachtmenge: Deutschland, jahre, tierarten, schlachtungsart. Accessed on 02 April 2020. https://www-genesis.destatis.de/genesis//online?operation=table&code=41331-0001&levelindex=0&levelid=1585878027052.

eurostat. 2020a. Bovine population (Annual data). Accessed on 13 April 2020. https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=apro_mt_lscatl&lang=en.

eurostat. 2020b. Production and utilization of milk on the farm (Annual data). Accessed on 02 April 2020. https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=apro_mk_farm&lang=en.

eurostat. 2020c. Slaughtering in slaughterhouses (Annual data). Accessed on 08 April 2020. https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=apro_mt_pann&lang=en.

FAO – Food and Agriculture Organization of the United Nations. 2020. Production: Livestock primary. Accessed on 02 April 2020. http://www.fao.org/faostat/en/#data/QL.

IBGE – Instituto Brasileiro de Geografia e Estatistica. 2020a. Pesquisa da pecuária municipal – Table 74. Accessed on 02 April 2020. https://sidra.ibge.gov.br/tabela/74.

IBGE – Instituto Brasileiro de Geografia e Estatistica. 2020b. Pesquisa da pecuária municipal – Table 94. Accessed on 02 April 2020. https://sidra.ibge.gov.br/tabela/94.

IBGE – Instituto Brasileiro de Geografia e Estatistica. 2020c. Pesquisa trimestral do abate de animais – Table 1092. Accessed on 02 April 2020. https://sidra.ibge.gov.br/tabela/1092.

LIC & DairyNZ. 2019. New Zealand dairy statistics 2018-19 Report. LIC – Livestock Improvement Corporation & DairyNZ, Hamilton, New Zealand. Accessed on 14 April 2020. https://www.dairynz.co.nz/publications/dairy-industry/new-zealand-dairy-statistics-2018-19/.

MAGyP – Ministerio de Agricultura, Ganadería y Pesca, Subsecretaría de Ganadería. 2020a. Ganadería – Indicadores bovinos anuales. Accessed on 14 April 2020. https://datos.agroindustria.gob.ar/dataset/indicadores-anuales-sector-bovino.

MAGyP – Ministerio de Agricultura, Ganadería y Pesca, Secretaría de Gobierno de Agroindustria de la Nación: Dirección Nacional Láctea. 2020b. Producción primaria. Accessed on 14 April 2020. https://www.magyp.gob.ar/sitio/areas/ss_lecheria/estadisticas/_01_primaria/index.php.

Narayanan, Y. 2019. Jugaad and informality as drivers of India's cow slaughter economy. Environment and Planning A: Economy and Space 51(7):1516-1535. https://doi.org/10.1177/0308518x19852640.

NZ.Stat. 2020a. Total New Zealand by kill by animal type (Annual-Dec). Accessed on 08 April 2020. http://archive.stats.govt.nz/infoshare/.

NZ.Stat. 2020b. Variable by total New Zealand (Annual-Jun). Accessed on 13 April 2020. http://archive.stats.govt.nz/infoshare/.

SIACON – Sistema de Información Agroalimentaria de Consulta. 2020. Pecuario estatal. Secretaría de Agricultura y Desarrollo Rural (SADER) & Servicio de Información Agroalimentaria y Pesquera (SIAP). Accessed on 14 April 2020. https://www.gob.mx/siap/documentos/siacon-ng-161430.

Statistics Canada. 2020a. Milk production and utilization. Accessed on 13 April 2020. https://doi.org/10.25318/3210011301-eng.

Statistics Canada. 2020b. Number of cattle, by class and farm type (x 1,000). Accessed on 13 April 2020. https://doi.org/10.25318/3210013001-eng.

USDA – United States Department of Agriculture. 2020a. Annual milk production and factors affecting supply (Annual). Accessed on 02 April 2020. https://www.ers.usda.gov/data-products/dairy-data/.

USDA – United States Department of Agriculture. 2020b. Livestock and poultry slaughter (Historical). Accessed on 02 April 2020. https://www.ers.usda.gov/data-products/livestock-meat-domestic-data/livestock-meat-domestic-data/#Livestock%20and%20poultry%20slaughter.

USDA – United States Department of Agriculture. 2020c. Production, supply and distribution (PSD) – Table livestock and poultry. Accessed on 08 April 2020. https://apps.fas.usda.gov/psdonline/downloads/psd_livestock_csv.zip.

Connecting statement 1

Avenues requiring further investigation to improve dairy cow longevity were highlighted in the critical review. Out of those, the long-term effect of early life conditions was shown to be the one least explored. Many early life aspects of a dairy cow could be considered. However, the lack of comprehensive data limits the possibilities. One alternative to overcome such limitation, while ensuring a significant sample size, is to use data collected by Dairy Herd Improvement (**DHI**) agencies. Lactanet, the Canadian DHI agency, routinely collects birth conditions related to calving ease, calf size, and twining. These indicators could be used to improve the understandings on the potential relationship between early life and subsequent cow longevity.

CHAPTER 3 – Birth conditions affect the longevity of Holstein offspring

Gabriel M. Dallago[1,*], Roger I. Cue[1], Kevin M. Wade[1], René Lacroix[2], and Elsa Vasseur[1]

[1] Department of Animal Science, McGill University, Sainte-Anne-de-Bellevue, Quebec, H9X 3V9, Canada

[2] Lactanet, Sainte-Anne-de-Bellevue, Quebec, H9X 3R4, Canada

* Corresponding author: gabriel.dallago@mail.mcgill.ca

Received: 25 January 2021; Accepted: 2 October 2021; Published: 16 November 2021

Journal of Dairy Science. 105(2):1255-1264. https://doi.org/ 10.3168/jds.2021-20214.

3.1. Abstract

Studies of dairy cow longevity usually focus on the animal life after first calving, with few studies considering early life conditions and their effects on longevity. The objective was to evaluate the effect of birth conditions routinely collected by Dairy Herd Improvement agencies on offspring longevity measured as length of life and length of productive life. Lactanet provided 712,890 records on offspring born in 5,425 Quebec dairy herds between January 1999 and November 2015 for length of life, and 506,066 records on offspring born in 5,089 Quebec dairy herds between January 1999 and December 2013 for length of productive life. Offspring birth conditions used in this study were calving ease (unassisted, pull, surgery, or malpresentation), calf size (small, medium, or large), and twinning (yes or no). Observations were considered censored if the culling reason was “exported,” “sold for dairy production,” or “rented out” as well as if the animals were not yet culled at the time of data extraction. If offspring were not yet culled when the data were extracted, the last test-day date was considered the censoring date. Conditional inference survival trees were used in this study to analyze the effect of offspring birth conditions on offspring longevity. The hazard ratio of culling between the groups of offspring identified by the survival trees was estimated using a Cox proportional hazard model with herd-year-season as a frailty term. Five offspring groups were identified with different length of life based on their birth condition. Offspring with the highest length of life [median = 3.61 year; median absolute deviation (MAD) = 1.86] were those classified as large or medium birth size and were also the result of an unassisted calving. Small offspring as a result of a twin birth had the lowest length of life (median = 2.20 year; MAD = 1.69) and were 1.52 times more likely to be culled early in life. Six groups were

identified with different length of productive life. Offspring that resulted from an unassisted or surgery calving and classified as large or medium when they were born were in the group with the highest length of productive life (median = 2.03 year; MAD = 1.63). Offspring resulting from a mal-presentation or pull in a twin birth were in the group with the lowest length of productive life (median = 1.15 year; MAD = 1.11) and were 1.70 times more likely to be culled early in life. In conclusion, birth conditions of calving ease, calf size, and twinning greatly affected offspring longevity, and such information could be used for early selection of replacement candidates.

3.2. Introduction

Dairy cow longevity is a complex functional trait that reflects the culling decisions made by a producer while taking into consideration many different aspects of the lifespan of a dairy animal (Van Doormaal, 2009). Cow longevity is associated with the overall sustainability of the dairy industry since short longevity is associated with financial losses on farms (Brickell and Wathes, 2011, Pellerin et al., 2014, Boulton et al., 2017), increased environmental footprint of milk production (Benbrook et al., 2010, Grandl et al., 2019), and impaired animal welfare status (Berry, 2015, De Vries, 2020). Even though the genetic potential for longevity has increased over time since the calculation of breeding values started to incorporate functional traits (De Vries, 2017, CRV, 2020, DairyNZ, 2020), the phenotypical expression of such potential has been identified as a problem requiring further research.

Longevity has been defined as an animal having an early first calving followed by a long productive life under profitable production levels (Dallago et al., 2021). Such definition was proposed since longevity is not only influenced by milk production, but also reproduction performance, incidence of health issues, and profitability (Compton et al., 2017, De Vries and Marcondes, 2020). Even though no metric exists covering all longevity facets outlined in the definition, length of life (**LL**) and length of productive life (**LPL**) are two of the most common (Dallago et al., 2021). While the former accounts for the entire length of time an animal stay in the herd (i.e., from birth to culling or death; Haworth et al., 2008), the latter measure the length of time between first calving and culling or death (Ducrocq, 1994, Schneider et al., 2007, Haworth et al., 2008).

The study of dairy cow longevity usually begins only after the animal's first calving, mainly focusing on LPL, and less attention has been given to early life conditions. For instance, culling has been extensively studied within lactations. Culling at the beginning of the lactation is often related to the incidence of injuries and diseases (Pinedo et al., 2010, Pinedo et al., 2014), while culling later in lactation is associated with reproduction problems and low milk production (Pinedo et al., 2010, Pinedo et al., 2014, Heise et al., 2016).

Although studied to a much lesser extent, a few studies indicated the existence of long-term effects of pre-partum and early life conditions that would affect cow longevity. Offspring born to dams exposed to heat stress during dry period had a compromised passive immune transfer and weighted less both at birth and weaning (Tao et al., 2012, Monteiro et al., 2016) and also had their survival and first lactation milk production negatively affected (Monteiro et al., 2016). The occurrences of health problems before the first calving were reported to negatively affect milk production (Svensson and Hultgren, 2008) and reproduction (Hultgren and Svensson, 2010) of dairy cows. In addition, a high average daily gain from birth to weaning was associated with early age at first calving (Vacek et al., 2015), high milk yield in the first lactation (Svensson and Hultgren, 2008, Vacek et al., 2015), and low occurrence of veterinarian-treated cases of mastitis at the beginning of the lactation of primiparous cows (Svensson et al., 2006). Since failure to reproduce, incidence of health issues such as mastitis and feet and leg problems, and low milk production are the main reported reasons for culling in high milk-producing countries (Dallago et al., 2021), altogether these studies provide evidence that early life conditions - especially during the rearing period - have an adverse effect on cow longevity.

Therefore, it was hypothesized that birth conditions have an effect on offspring longevity, and the objective of this study was to evaluate that effect, using birth conditions variables routinely collected by Dairy Herd Improvement (DHI) agencies.

3.3. Material and methods

3.3.1. Dataset

A retrospective longitudinal study was conducted using DHI data from dairy herds in the province of Quebec, Canada provided by the Canadian DHI agency Lactanet Inc and organized into Animal, Lactation, and Test Day data. The Animal data contained animal identification variables [offspring identification (**ID**), offspring registration ID, dam registration ID, and sire registration ID], breed (offspring, dam, and sire), birth date, and left herd date of 3,380,326 records on offspring born in 7,660 Quebec dairy herds between September 1961 and December 2015. The Lactation data had total lactation performance (i.e., milk, fat, and protein production), lactation start date, and birth conditions (calving ease, calf size, and twinning) of 4,698, 162 records on offspring that calved in 7,521 Quebec dairy herds between September 1991 and December 2015. The Test Day data contained offspring performance similar to the Lactation data but from each individual test day (N = 50,368,719 records on offspring with a test day between January 2000 and December 2015 on 7,525 herds).

Birth conditions used in this study were calving ease (unassisted, pull, surgery, or mal-presentation), calf size (small, medium, or large), and twinning (yes or no). Birth conditions were available in the Lactation data file, but they were identified by the dam ID. Offspring ID, on the other hand, was available in the Animal data, but associated with the dam registration ID. Therefore, to obtain a dam ID associated with the offspring ID, and extract the birth conditions, we merged offspring registration ID with dam registration ID from the Animal data (N = 1,684,222 records with no match and excluded). Duplicated observations on Animal (N = 248 records) and Lactation (N = 232,823 records) data files were excluded before both files were merged based on dam ID, year, and season of calving/birth. Next, records with missing observations on birth conditions were excluded (N = 88,240 records excluded). After the merge, there remained 1,082,122 records from offspring that were born between April 1997 and December 2015. For all cases of duplicated records, both records were excluded. Recorded freemartins (i.e., twin to a bull) were removed. In addition, observations were removed where the sex of one of the twins was missing to avoid the possible effect of not recorded freemartins. The data was then filtered to contain only information about female Holstein offspring and a minimum of three observations per Herd-Year-Season of birth. The Figure 3.1 shows a flowchart of the data preparation procedures.

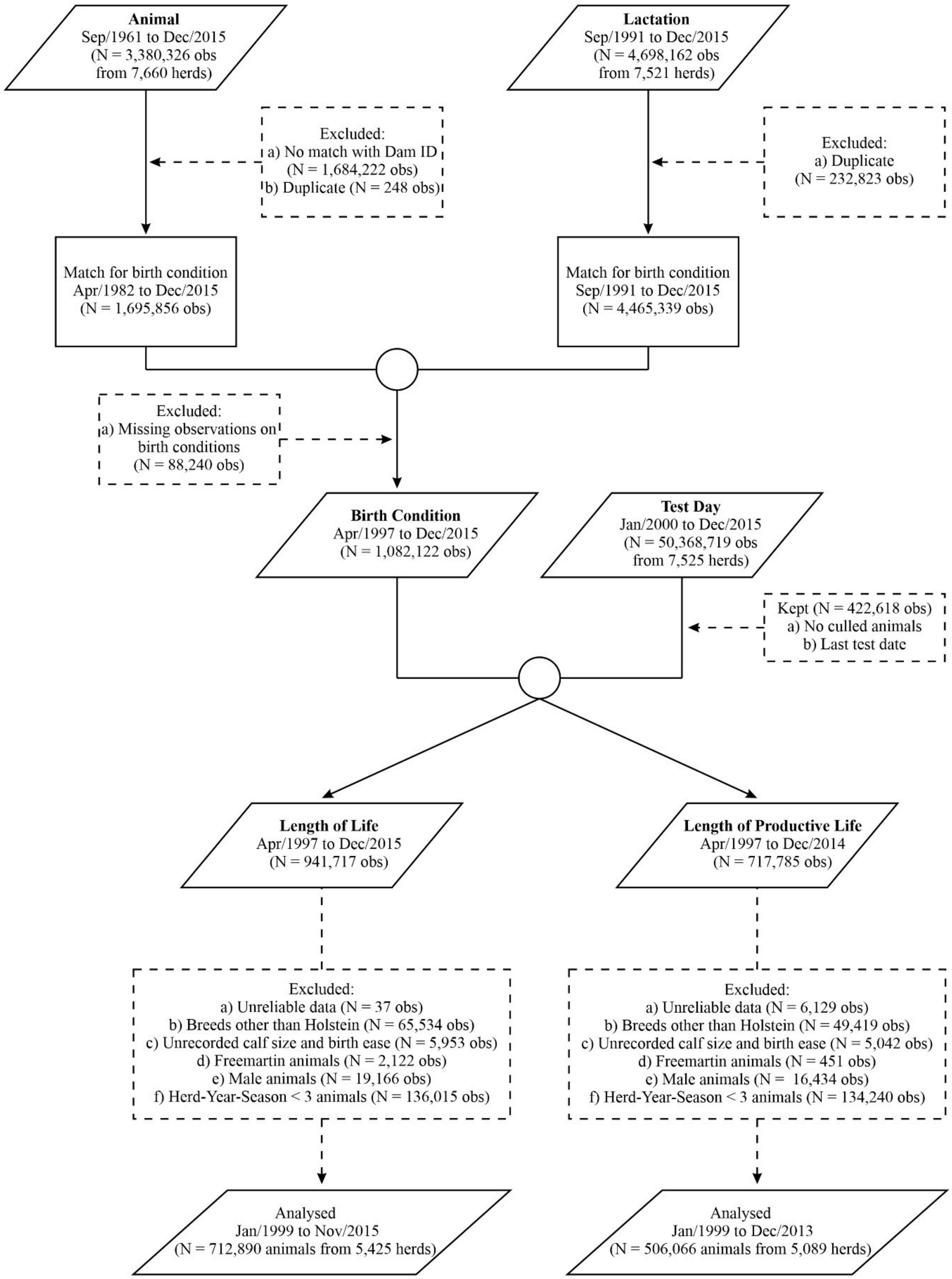


Figure 3.1. Flowchart of herds and animal selection. Parallelograms represent data input and output, solid-line rectangles represent intermediary. Obs = observations.

The LL and LPL were the longevity metrics evaluated in this study. The LL was calculated as the length of time (year) between birth and death/culling or censoring (Haworth et al., 2008) while the LPL was the length of time (year) between the first calving and death/culling or censoring (Ducrocq, 1994, Schneider et al., 2007). Time to event was defined as the time between birth and culling/death or censoring for LL and between first calving and culling/death or censoring for LPL. An observation was considered censored if, at the time that the data was extracted from the DHI database, the offspring was not yet culled (N = 321,483 for LL and N = 240,792 for LPL) or if the reported culling reason was "exported" (N = 25,297 for LL and N = 13,500 for LPL), "sold for dairy production" (N = 49,087 LL and N = 34,517 for LPL), or "rented out" (N = 1,496 for LL and N = 1,203 for LPL). If the offspring were not yet culled when the data were extracted, the last test day date was considered as the censoring date. For the other cases, offspring were considered censored at the date of the reported removal.

After the data handling and cleaning, there remained data from 712,890 female Holstein offspring born between January 1999 and November 2015 in 5,425 herds for LL and from 506,066 female Holstein offspring born between January 1999 and December 2013 in 5,089 herds for LPL.

3.3.2. Data analysis

All statistical analyses were carried out using R version 4.0.2 (https://www.r-project.org/) and its packages as follows: *coxme* (Therneau, 2020a), *partykit* (Hothorn et al., 2006, Hothorn and Zeileis, 2015), *survival* (Therneau, 2020b), and *survminer* (Kassambara et al., 2020). The statistical significance level was set at $\alpha < 0.05$.

Similar to Probo et al. (2018), conditional inference survival trees were used in this study to analyse the effect of birth conditions on offspring longevity since it has been shown to be less prone to overfitting and more reliable than other survival tree algorithms (Zhou and McArdle, 2015). In short, this algorithm aims at recursively partitioning the data into different nodes (i.e., group of offspring) based on the response variable. At each step, it selects independent variables with the highest ability to split the observations of offspring with different longevity. Therefore, the earlier a variable is used to split the data and originate a node, the higher it is its ability to identify offspring with different longevity. This is repeated until the independence hypothesis between the response and any of the independent variables cannot be rejected (Hothorn et al., 2006). By doing so, the algorithm is able to automatically group offspring separately within the

LL and LPL and the independent variables under consideration as well as detecting high-level interactions without the need to specify them (Ramezankhani et al., 2017). The trees were created having LL and LPL individually as the response variable and using the *ctree* function (Hothorn et al., 2006). The resulting groups of offspring identified by the survival tree were used as a categorical variable to estimate the hazard ratio (**HR**) of culling using a shared frailty Cox proportional hazard model. A frailty term for the combined effect of the Herd-Year-Season of birth (**HYS**) of the offspring [HYS ~ $N(0, \sigma^2)$] was included in the model to account for the shared frailty between offspring within the same HYS. The offspring longevity was assumed to be independent conditional on the HYS. The group with the lowest culling risk (i.e., highest longevity) was used as the reference level. In addition, Kaplan-Meier curves were constructed for each of the groups.

3.4. Results

Among studied offspring, 321,483 (45.1%) and 240,792 (47.6%) were censored for LL and LPL, respectively. Among offspring that were not censored, the overall median LL was 3.53 years [median absolute deviation (**MAD**)= 1.82; 95% percentile compatibility interval (**PCI**) = 3.53 – 3.54] and the overall median LPL was 1.98 years (MAD = 1.62; 95% PCI = 1.97 – 1.99). Descriptive statistics for the offspring birth conditions are presented in Table 3.1.

Table 3.1. Descriptive statistics showing the prevalence of distinct birth conditions, according to the different measures of longevity.

Birth condition	Length of life[1] (N = 712,890)		Length of productive life[2] (N = 506,066)	
	N	%	N	%
Calving ease				
Mal-presentation	10,068	1.41	7,110	1.40
Pull	243,220	34.12	173,303	34.25
Surgery	650	0.09	451	0.09
Unassisted	458,952	64.38	325,202	64.26
Calf size				
Large	119,957	16.83	87,054	17.20
Medium	492,852	69.13	351,607	69.48
Small	100,081	14.04	67,405	13.32
Twinning				
No	701,928	98.46	499,440	98.69
Yes	10,962	1.54	6,626	1.31

[1] Length of time between birth and death/culling (Haworth et al., 2008).
[2] Length of time between the first calving and death/culling (Ducrocq, 1994, Schneider et al., 2007).

3.4.1. Length of life

All three offspring birth conditions evaluated in this study were selected by the survival tree algorithm to group the offspring (Supplementary Figure 3.1). Calf size was found as the most important condition since it was at the top of the tree. Five groups were identified with different offspring LL based on their birth condition (P < 0.001; Supplementary Figure 3.1). Offspring with the highest LL were those that were classified as of large or medium size at birth and were the result of an unassisted calving (node 6; Supplementary Figure 3.1), whereas having a small size at birth and being the result of a twin birth (node 9; Supplementary Figure 3.1) were in the group with the lowest LL (Table 3.2). Therefore, the HR of culling was the highest for node 9 (Table 3.2) with node 6 as the reference group (Table 3.2). The Kaplan-Meier curves for these groups are shown in Figure 3.2A, where the curve for node 6 was the only one higher than the overall average curve, indicating that it covered a low-risk group for culling events.

Table 3.2. Descriptive statistics and hazard ratio of the patterns identified by the conditional inference survival tree for the length of life.

Node[1]	Cases and events[2]	Pattern description[3]	Length of life[4]		HR (95% CI)[5]	P-value
			Median (MAD)	95% PCI		
6	384,539; 212,990	Calf size = large or medium; calving ease = unassisted	3.61 (1.86)	3.59 – 3.62	Reference	---
4	8,683; 5,127	Calf size = large or medium; calving ease = mal-presentation	3.40 (1.68)	3.35 – 3.46	1.16 (1.12 – 1.19)	< 0.001
5	219,587; 120,225	Calf size = large or medium; calving ease = pull or surgery	3.52 (1.77)	3.51 – 3.54	1.02 (1.01 – 1.03)	< 0.001
8	93,254; 50,500	Calf size = small; twin = no	3.33 (1.80)	3.31 – 3.35	1.10 (1.09 – 1.12)	< 0.001
9	6,827; 2,556	Calf size = small; twin = yes	2.20 (1.69)	2.13 – 2.27	1.52 (1.45 – 1.59)	< 0.001

[1] Terminal node identified by the survival tree algorithm for the length of life
[2] Cases = Number of offspring in each node; events = number of offspring in each node with an observed culling event
[3] Simplified pattern description of each terminal node
[4] Median (Median absolute deviation; MAD) and 95% percentile compatibility interval (Greenland, 2019)
[5] HR (95%CI) = Hazard ratio and 95% compatibility interval (Greenland, 2019)

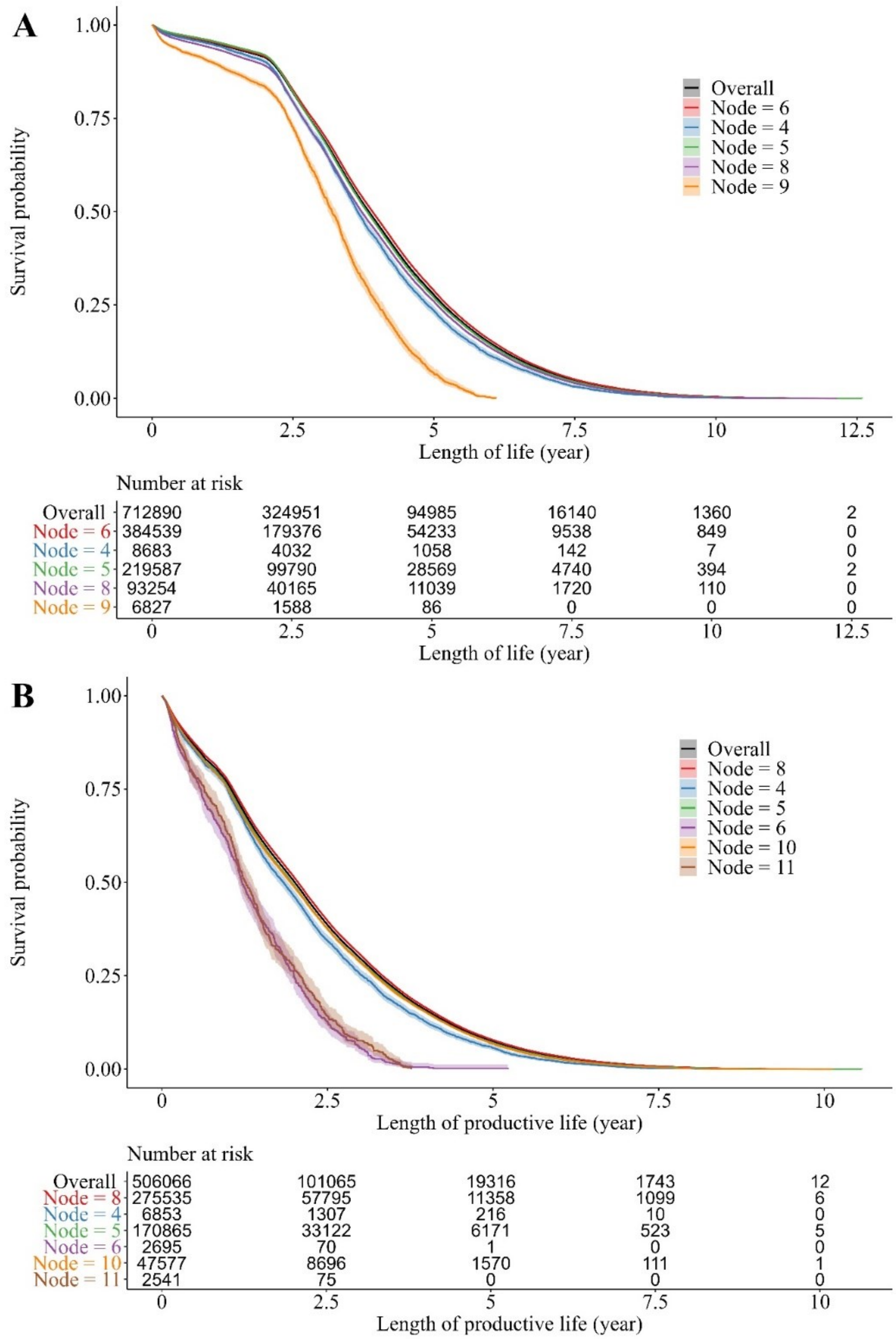


Figure 3.2. The Kaplan-Meier survival curves, followed by the 95% compatibility interval (Greenland, 2019), of the terminal nodes identified by the survival tree algorithm for the length of life (A) and the length of productive life (B). The risk table at the bottom of both plots indicates the number of offspring at risk of being culled at different times. In both plots, the solid black line (overall) shows the survival curve of the total population, and the other lines show the survival curve of each terminal node. The terminal nodes of graph A represent the following patterns: Node = 6: large or medium calf size and unassisted calving ease; Node = 4: large or medium calf size and mal-presentation calving ease; Node = 5: large or medium calf size and pull or surgery calving ease; Node = 8: small calf size and no twinning; Node = 9: Small calf size and twinning. The terminal nodes of graph B represent the following patterns: Node = 8: Surgery or unassisted calving ease and large or medium calf size; Node = 4: mal-presentation calving ease and no twin birth; Node = 5: pull calving ease and no twin birth; Node = 6: mal-presentation or pull calving ease and twining; Node = 10: surgery or unassisted calving ease, small calf size, and no twinning; Node = 11: surgery or unassisted calving ease, small calf size, and twinning.

3.4.2. Length of productive life

Similar to LL, all three offspring birth conditions evaluated in this study were also selected by the survival tree algorithm to group the offspring based on LPL (Supplementary Figure 3.2). However, calving ease was found as the most important condition in this case and 6 groups were identified with different LPL ($P < 0.001$; Supplementary Figure 3.2). Offspring that resulted from an unassisted or surgery calving and were classified as of large or medium size at birth (Node 8; Supplementary Figure 3.2) were in the group with the highest LPL, whereas offspring resulting from a mal-presentation or pull in a twin birth (Node 6; Supplementary Figure 3.2) ere in the group with the lowest LPL (Table 3.3). Accordingly, the HR of culling was the highest for node 6, with node 8 as the reference group (Table 3.3). The Kaplan-Meier curves for these groups are shown in Figure 3.2, where the curve for node 8 was also the only one higher than the overall average curve, indicating that it covered a low-risk group for culling events.

Table 3.3. Descriptive statistics and hazard ratio of the patterns identified by the conditional inference survival tree for the length of productive life.

Node[1]	Cases and events[2]	Pattern description[3]	Length of productive life[4]		HR (95% CI)[5]	P-value
			Median (MAD)	95% PCI		
8	275,535; 147,301	Calving ease = surgery or unassisted; calf size = large or medium	2.03 (1.63)	2.02 – 2.04	Reference	---
4	6,853; 3,859	Calving ease = mal-presentation; twin = no	1.80 (1.49)	1.72 – 1.87	1.15 (1.11 – 1.20)	< 0.001
5	170,865; 89,410	Calving ease = pull; twin = no	1.93 (1.62)	1.92 – 1.94	1.05 (1.04 – 1.06)	< 0.001
6	2,695; 590	Calving ease = mal-presentation or pull; twin = yes	1.15 (1.11)	1.03 – 1.21	1.70 (1.55 – 1.87)	< 0.001
10	47,577; 23,552	Calving ease = surgery or unassisted; calf size = small; twin = no	1.91 (1.62)	1.88 – 1.94	1.05 (1.04 – 1.07)	< 0.001
11	2,541; 562	Calving ease = surgery or unassisted; calf size = small; twin = yes	1.18 (1.15)	1.10 – 1.28	1.63 (1.48 – 1.79)	< 0.001

[1] Terminal node identified by the survival tree algorithm for the length of productive life.
[2] Cases = Number of offspring in each node; Events = Number of offspring in each node with an observed culling event.
[3] Simplified pattern description of each terminal node.
[4] Median (Median absolute deviation) and 95% percentile compatibility interval (Greenland, 2019).
[5] HR (95%CI) = Hazard ratio and 95% compatibility interval (Greenland, 2019).

3.5. Discussion

Different indicators of birth condition were selected by the machine-learning algorithm to first split the observations into groups of Offspring with different LL (Supplementary Figure 3.1) and LPL (Supplementary Figure 3.2). The first split contains the indicator with the strongest association with the response variable (Hothorn et al., 2006). The results of our study indicate that size at birth is the birth condition with the strongest association with the LL (Supplementary Figure 3.1). Since the main difference between LL and LPL is the inclusion of the length of the non-productive life stage of the animal in the former, our results suggest that size at birth had the strongest association with the ability of the offspring to avoid herd removal up to the first calving. In turn, this would be influenced by the mortality rate since it is higher during the early life stages compared to lactating cows (Compton et al., 2017).

Calf weight at birth, as a proxy for calf size, has been reported to be associated with calf mortality. For instance, clinically normal calves that died at birth or within 24 hrs after birth were, on average, 6 kg lighter than calves that died due to difficult calving (Berglund et al., 2003), which could be because smaller calves would have an undeveloped immune system and low body energy reserve, making them more susceptible to hypothermia (Wathes et al., 2008). The growth of small calves is also compromised, which negatively influences the age at first calving and the ability to remain in the herd. Bodyweight gain is positively associated with birth weight, in which the higher the birth weight, the higher the body weight at 6, 9, and 15 months old (Swali and Wathes, 2006). Brickell et al. (2007) reported that cows with a higher body weight at 6 months old (183 kg; SD = 36 kg) had an age at first calving lower than 23 months compared to animals with a lower body weight also at 6 months old (162 kg; SD = 35 kg). Additionally, a positive relationship was reported between first lactation milk yield and the average daily gain from weaning to first service (Svensson and Hultgren, 2008) or conception (Vacek et al., 2015) as well as between the average daily gain during the first two months of life and completing the first lactation (Bach, 2011).

Although the same size at birth, calves with difficult calving had a higher HR of being culled compared with offspring with an unassisted birth for both LL (Table 3.2) and LPL (Table 3.3). In addition to increasing the risk of stillbirths (Berglund et al., 2003, Lombard et al., 2007) and mortality during the first 21 days of life (Wells et al., 1996), difficult calving has been associated with failure of passive-immunity transfer (Renaud et al., 2020) and with low calf vitality (Barrier et al., 2012). Difficult calving also increases the risk of developing respiratory and digestive diseases between birth and 120 days of age (Lombard et al., 2007). The negative consequences of a difficult calving are likely to have long-term effects. Difficult calving was associated with increased age at first calving (Heinrichs et al., 2005), which in turn makes animals less likely to complete a first lactation (Bach, 2011, Sherwin et al., 2016). In addition, difficult calving was associated with a reduction in cumulative milk yield in the first lactation of the offspring (Eaglen et al., 2011), which would increase the chances of early culling.

Even though results from this study indicate that calving assistance has negative effects on offspring longevity, delivery assistance alone is not likely to be a problem. Systematically assisting the delivery of calves that do not present clear signs of dystocia improves the calf vigour and does not influence the likelihood of stillbirth (Villettaz Robichaud et al., 2017a) as well as the occurrence of pneumonia, diarrhea, and the survival of the offspring up to weaning at 7 weeks old (Villettaz Robichaud et al., 2017b). Even though further studies are needed to evaluate the long-term effects of systematically assisting the delivery on offspring longevity, this would indicate that calving assistance as a consequence of dystocia would have a negative effect on offspring longevity since this is likely the most prevalent scenario in our study given the rate of assistance provided by Canadian farmers (Villettaz Robichaud et al., 2016).

The occurrence of twinning had a negative effect on offspring longevity measured as LL (Table 3.2 and Supplementary Figure 3.1) or LPL (Table 3.3 and Supplementary Figure 3.2). The overall twinning rate in dairy cows was estimated to be 4.2% in the US, but it increased with the parity of the dam as well as overtime between 1996 and 2004 (Silva del Río et al., 2007). This overall prevalence was higher than the values reported here (Table 3.1) most likely because freemartins were removed from our study. The negative consequences of twining on the dam's survival are well documented. For example, it increases the hazard ratio of multiparous cows being culled within 120 days in milk (Probo et al., 2018). However, twining has also been reported to have negative consequences on the offspring. The occurrence of twin births was identified as a mortality risk factor within the first 24 hours after birth (Lombard et al., 2007, Silva del Río et al., 2007) as well as up to 21 days of life (Wells et al., 1996). Twining is also associated with low body weight (Windeyer et al., 2014) and average daily gain (Shivley et al., 2018) compared to single births. Such consequences are likely to have a cascade effect that compromises the ability of the offspring to remain in the herd, resulting in reduced offspring longevity as suggested by the results in this study.

The overall average longevity observed in the present study was low but in agreement with results published previously. Hertl et al. (2018) reported similar results with an average LPL of approximately 2 years based, on 24,831 Holstein cows from 5 herds in New York State, while

Hultgren and Svensson (2009) reported a median length of productive life of 2.1 years among 2,124 culled dairy cows from 109 Swedish herds. On the other hand, the genetic potential for longevity has increased over time since the calculation of breeding values started to incorporate functional traits (De Vries, 2017, CRV, 2020, DairyNZ, 2020). Based on a combination of survival data and various non-production traits known to be associated with longevity (Beavers and Van Doormaal, 2017), the genetic potential for LL of the average Canadian Holstein cow born in 2018, for example, was estimated to be of 8.5 years (Lactanet, 2021). This would represent an LPL of 6.3 years if an average age at first calving of 27 months is assumed. However, the culling rate in Canada was 32.36% in 2020 among herds enrolled on a milk recording program (CDIC, 2021). This represents an LPL of 3.09 years, which is about half of the genetic potential of the cows. Results from the above literature indicate that dairy cows are not expressing their full potential for longevity, and the present study provides insights on the adverse effect of birth conditions (i.e., calving ease, birth size, and twinning) on offspring longevity, which could be used in replacement and culling strategies to select the best candidates for early herd removal.

Early life indicators of longevity are required to support farmers' culling and replacement strategies. Age at first calving has been proposed as an indicator of longevity since animals that calve for the first time at a younger age are less likely to be culled (Bach, 2011, Sherwin et al., 2016) and more likely to have a long productive life (Hultgren and Svensson, 2009). However, much earlier indicators would be desirable since age at first calving only happens at approximately two years after birth. The results of the present study indicate that birth conditions could be used as early indicators of offspring longevity by identifying offspring during the rearing period that are less likely to express their maximum potential and support farmers in the early selection of replacement animals.

The present study focused on indicators of birth conditions that are available in DHI databases. Our results are likely to be influenced by the fact that 26.7% of Canadian farmers always assist their cows during calving, and 38% always assist all heifers regardless of the necessity to do so (Villettaz Robichaud et al., 2016). In addition, a limitation of our approach is the inherited subjectivity of the records reported by dairy farmers. For instance, classifying calf size into small, medium, and large is, to some extent, subjective to the farmer's criteria, which is likely to vary among different farmers. Even though a certain standardization is present, some heterogeneity is present in the way people apply these standards due to their perceptions. Our results provide a strong incentive for farmers to record actual weight at birth, making birth-size information more precise given its potential to help identify offspring that may stay longer in the herd and optimize replacement and culling strategies.

3.6. Conclusion

In conclusion, our results indicate the existence of a long-term effect of birth conditions on offspring longevity. Calving ease, birth size, and twinning are associated with a reduction in both length of life and length of productive life and affects the ability of the offspring to remain in the herd. This information provides insights to optimize replacement and culling strategies in dairy herds and propose early indicators for the selection of the best candidates for early herd removal.

3.7. References

Bach, A. 2011. Associations between several aspects of heifer development and dairy cow survivability to second lactation. Journal of Dairy Science 94(2):1052-1057. https://doi.org/10.3168/jds.2010-3633.

Barrier, A. C., E. Ruelle, M. J. Haskell, and C. M. Dwyer. 2012. Effect of a difficult calving on the vigour of the calf, the onset of maternal behaviour, and some behavioural indicators of pain in the dam. Preventive Veterinary Medicine 103(4):248-256. https://doi.org/10.1016/j.prevetmed.2011.09.001.

Beavers, L. and B. Van Doormaal. 2017. Herd life and voluntary disposals. Accessed on 23 September 2021. https://www.cdn.ca/document.php?id=467.

Benbrook, C., C. Carman, E. A. Clark, C. Daley, W. Fulwider, M. Hansen, C. Leifert, K. Martens, L. Paine, L. Petkewitz, G. Jodarski, F. Thicke, J. Velez, and G. Wegner. 2010. A dairy farm's footprint: Evaluating the impacts of conventional and organic farming systems. Critical Issue Report. The Organic Center, Washington D.C., USA.

Berglund, B., L. Steinbock, and M. Elvander. 2003. Causes of stillbirth and time of death in Swedish Holstein calves examined post mortem. Acta Veterinaria Scandinavica 44(3):111. https://doi.org/10.1186/1751-0147-44-111.

Berry, D. P. 2015. Breeding the dairy cow of the future: What do we need? Animal Production Science 55(7):823-837. https://doi.org/10.1071/AN14835.

Boulton, A. C., J. Rushton, and D. C. Wathes. 2017. An empirical analysis of the cost of rearing dairy heifers from birth to first calving and the time taken to repay these costs. Animal 11(8):1372-1380. https://doi.org/10.1017/S1751731117000064.

Brickell, J. S., N. Bourne, Z. Cheng, and D. C. Wathes. 2007. Influence of plasma IGF-I oncentrations and body weight at 6 months on age at first calving in dairy heifers on comercial farms. Reproduction in Domestic Animals 42(Suppl. 2):118-119. https://doi.org/10.1111/j.1439-0531.2007.00910.x.

Brickell, J. S. and D. C. Wathes. 2011. A descriptive study of the survival of Holstein-Friesian heifers through to third calving on English dairy farms. Journal of Dairy Science 94(4):1831-1838. https://doi.org/10.3168/jds.2010-3710.

CDIC – Canadian Dairy Information Centre. 2021. Culling and replacement rates in dairy herds in Canada. Accessed on 07 July 2021. https://agriculture.canada.ca/en/canadas-agriculture-sectors/animal-industry/canadian-dairy-information-centre/dairy-statistics-and-market-information/dairy-animal-genetics/culling-and-replacement-rates-dairy-herds-canada.

Compton, C. W. R., C. Heuer, P. T. Thomsen, T. E. Carpenter, C. V. C. Phyn, and S. McDougall. 2017. Invited review: A systematic literature review and meta-analysis of mortality and culling in dairy cattle. Journal of Dairy Science 100(1):1-16. https://doi.org/10.3168/jds.2016-11302.

CRV – Coöperatie Koninklijke CRV. 2020. Jaarstatistieken 2019 voor Nederland [Annual statistics 2019 for the Netherlands]. Arnhem, Netherlands. Accessed on 20 April 2020. https://cooperatiecrv-be6.kxcdn.com/wp-content/uploads/2020/02/Jaarstatistieken-2019-NL-totaal.pdf.

DairyNZ. 2020. Animal evaluation: Animal and herd averages. Accessed on 29 May 2020. https://www.dairynz.co.nz/animal/animal-evaluation/animal-and-herd-averages.

Dallago, G. M., K. M. Wade, R. I. Cue, J. T. McClure, R. Lacroix, D. Pellerin, and E. Vasseur. 2021. Keeping dairy cows for longer: A critical literature review on dairy cow longevity in high milk-producing countries. Animals 11(3):808. https://doi.org/10.3390/ani11030808.

De Vries, A. 2017. Economic trade-offs between genetic improvement and longevity in dairy cattle. Journal of Dairy Science 100(5):4184-4192. https://doi.org/10.3168/jds.2016-11847.

De Vries, A. 2020. Symposium review: Why revisit dairy cattle productive lifespan? Journal of Dairy Science 103(4):3838-3845. https://doi.org/10.3168/jds.2019-17361.

De Vries, A. and M. I. Marcondes. 2020. Review: Overview of factors affecting productive lifespan of dairy cows. Animal 14(S1):s155-s164. https://doi.org/10.1017/S1751731119003264.

Ducrocq, V. 1994. Statistical analysis of length of productive life for dairy cows of the Normande breed. Journal of Dairy Science 77(3):855-866. https://doi.org/10.3168/jds.S0022-0302(94)77020-X.

Eaglen, S. A. E., M. P. Coffey, J. A. Woolliams, R. Mrode, and E. Wall. 2011. Phenotypic effects of calving ease on the subsequent fertility and milk production of dam and calf in UK Holstein-Friesian heifers. Journal of Dairy Science 94(11):5413-5423. https://doi.org/10.3168/jds.2010-4040.

Grandl, F., M. Furger, M. Kreuzer, and M. Zehetmeier. 2019. Impact of longevity on greenhouse gas emissions and profitability of individual dairy cows analysed with different system boundaries. Animal 13(1):198-208. https://doi.org/10.1017/S175173111800112X.

Greenland, S. 2019. Valid p-values behave exactly as they should: Some misleading criticisms of p-values and their resolution with s-values. The American Statistician 73(sup. 1):106-114. https://doi.org/10.1080/00031305.2018.1529625.

Haworth, G. M., W. P. Tranter, J. N. Chuck, Z. Cheng, and D. C. Wathes. 2008. Relationships between age at first calving and first lactation milk yield, and lifetime productivity and longevity in dairy cows. Veterinary Record 162(20):643-647. https://doi.org/10.1136/vr.162.20.643.

Heinrichs, A. J., B. S. Heinrichs, O. Harel, G. W. Rogers, and N. T. Place. 2005. A prospective study of calf factors affecting age, body size, and body condition score at first calving of Holstein dairy heifers. Journal of Dairy Science 88(8):2828-2835. https://doi.org/10.3168/jds.S0022-0302(05)72963-5.

Heise, J., Z. Liu, K. F. Stock, S. Rensing, F. Reinhardt, and H. Simianer. 2016. The genetic structure of longevity in dairy cows. Journal of Dairy Science 99(2):1253-1265. https://doi.org/10.3168/jds.2015-10163.

Hertl, J. A., Y. H. Schukken, L. W. Tauer, F. L. Welcome, and Y. T. Gröhn. 2018. Does clinical mastitis in the first 100 days of lactation 1 predict increased mastitis occurrence and shorter herd life in dairy cows? Journal of Dairy Science 101(3):2309-2323. https://doi.org/10.3168/jds.2017-12615.

Hothorn, T., K. Hornik, and A. Zeileis. 2006. Unbiased recursive partitioning: A conditional inference framework. Journal of Computational and Graphical Statistics 15(3):651-674. https://doi.org/10.1198/106186006X133933.

Hothorn, T. and A. Zeileis. 2015. partykit: A modular toolkit for recursive partytioning in R. Journal of Machine Learning Research 16(118):3905-3909.

Hultgren, J. and C. Svensson. 2009. Heifer rearing conditions affect length of productive life in Swedish dairy cows. Preventive Veterinary Medicine 89(3):255-264. https://doi.org/10.1016/j.prevetmed.2009.02.012.

Hultgren, J. and C. Svensson. 2010. Calving interval in dairy cows in relation to heifer rearing conditions in southwest Sweden. Reproduction in Domestic Animals 45(1):136-141. https://doi.org/10.1111/j.1439-0531.2008.01273.x.

Kassambara, A., M. Kosinski, and P. Biecek. 2020. survminer: Drawing survival curves using 'ggplot2'.Version 0.4.8. Accessed 16 December 2020 https://CRAN.R-project.org/package=survminer.

Lactanet. 2021. National genetic trends by birth year for Holstein cows. Accessed on 23 September 2021. https://www.cdn.ca/files_ge_articles.php.

Lombard, J. E., F. B. Garry, S. M. Tomlinson, and L. P. Garber. 2007. Impacts of dystocia on health and survival of dairy calves. Journal of Dairy Science 90(4):1751-1760. https://doi.org/10.3168/jds.2006-295.

Monteiro, A. P. A., S. Tao, I. M. T. Thompson, and G. E. Dahl. 2016. In utero heat stress decreases calf survival and performance through the first lactation. Journal of Dairy Science 99(10):8443-8450. https://doi.org/https://doi.org/10.3168/jds.2016-11072.

Pellerin, D., S. Adams, F. Bécotte, R. Cue, R. Moore, and R. Roy. 2014. Pour une vache, l'âge d'or c'est la 4e lactation! [For a cow, the golden age is the fourth lactation!]. Pages 133-147 in Proc. Symposium sur les bovins laitiers: Choix d'aujourd'hui pour les défis de demain. CRAAQ – Le Centre de référence en agriculture et agroalimentaire du Québec, Saint-Hyacinthe, QC, Canada.

Pinedo, P. J., A. Daniels, J. Shumaker, and A. De Vries. 2014. Dynamics of culling for Jersey, Holstein, and Jersey × Holstein crossbred cows in large multibreed dairy herds. Journal of Dairy Science 97(5):2886-2895. https://doi.org/10.3168/jds.2013-7685.

Pinedo, P. J., A. De Vries, and D. W. Webb. 2010. Dynamics of culling risk with disposal codes reported by Dairy Herd Improvement dairy herds. Journal of Dairy Science 93(5):2250-2261. https://doi.org/10.3168/jds.2009-2572.

Probo, M., O. B. Pascottini, S. LeBlanc, G. Opsomer, and M. Hostens. 2018. Association between metabolic diseases and the culling risk of high-yielding dairy cows in a transition management facility using survival and decision tree analysis. Journal of Dairy Science 101(10):9419-9429. https://doi.org/10.3168/jds.2018-14422.

Ramezankhani, A., M. Tohidi, F. Azizi, and F. Hadaegh. 2017. Application of survival tree analysis for exploration of potential interactions between predictors of incident chronic kidney disease: a 15-year follow-up study. Journal of Translational Medicine 15(1):240. https://doi.org/10.1186/s12967-017-1346-x.

Renaud, D. L., K. M. Waalderbos, L. Beavers, T. F. Duffield, K. E. Leslie, and M. C. Windeyer. 2020. Risk factors associated with failed transfer of passive immunity in male and female dairy calves: A 2008 retrospective cross-sectional study. Journal of Dairy Science 103(4):3521-3528. https://doi.org/10.3168/jds.2019-17397.

Schneider, M. d. P., E. Strandberg, U. Emanuelson, K. Grandinson, and A. Roth. 2007. The effect of veterinary-treated clinical mastitis and pregnancy status on culling in Swedish dairy cows. Preventive Veterinary Medicine 80(2):179-192. https://doi.org/10.1016/j.prevetmed.2007.02.006.

Sherwin, V. E., C. D. Hudson, A. Henderson, and M. J. Green. 2016. The association between age at first calving and survival of first lactation heifers within dairy herds. Animal 10(11):1877-1882. https://doi.org/10.1017/S1751731116000689.

Shivley, C. B., J. E. Lombard, N. J. Urie, C. A. Kopral, M. Santin, T. J. Earleywine, J. D. Olson, and F. B. Garry. 2018. Preweaned heifer management on US dairy operations: Part VI. Factors associated with average daily gain in preweaned dairy heifer calves. Journal of Dairy Science 101(10):9245-9258. https://doi.org/10.3168/jds.2017-14022.

Silva del Río, N., S. Stewart, P. Rapnicki, Y. M. Chang, and P. M. Fricke. 2007. An observational analysis of twin births, calf sex ratio, and calf mortality in Holstein dairy cattle. Journal of Dairy Science 90(3):1255-1264. https://doi.org/10.3168/jds.S0022-0302(07)71614-4.

Svensson, C. and J. Hultgren. 2008. Associations between housing, management, and morbidity during rearing and subsequent first-lactation milk production of dairy cows in southwest Sweden. Journal of Dairy Science 91(4):1510-1518. https://doi.org/10.3168/jds.2007-0235.

Svensson, C., A. K. Nyman, K. P. Waller, and U. Emanuelson. 2006. Effects of housing, management, and health of dairy heifers on first-lactation udder health in southwest Sweden. Journal of Dairy Science 89(6):1990-1999. https://doi.org/10.3168/jds.S0022-0302(06)72266-4.

Swali, A. and D. C. Wathes. 2006. Influence of the dam and sire on size at birth and subsequent growth, milk production and fertility in dairy heifers. Theriogenology 66(5):1173-1184. https://doi.org/10.1016/j.theriogenology.2006.03.028.

Tao, S., A. P. A. Monteiro, I. M. Thompson, M. J. Hayen, and G. E. Dahl. 2012. Effect of late-gestation maternal heat stress on growth and immune function of dairy calves. Journal of Dairy Science 95(12):7128-7136. https://doi.org/10.3168/jds.2012-5697.

Therneau, T. M. 2020a. coxme: Mixed effects Cox models.Version 2.2-16. Accessed 16 December 2020 https://CRAN.R-project.org/package=coxme.

Therneau, T. M. 2020b. survival: A package for survival analysis in R.Version 3.1-12. Accessed 16 December 2020 https://CRAN.R-project.org/package=survival.

Vacek, M., L. Krpálková, J. Syrůček, M. Štípková, and M. Janecká. 2015. Relationships between growth and body condition development during the rearing period and performance in the first three lactations in Holstein cows. Czech Journal of Animal Science 60(9):417-425. https://doi.org/10.17221/8460-CJAS.

Van Doormaal, B. 2009. A closer look at longevity. Accessed on 28 October 2020. https://www.cdn.ca/document.php?id=162.

Villettaz Robichaud, M., A. M. de Passillé, D. L. Pearl, S. J. LeBlanc, S. M. Godden, D. Pellerin, E. Vasseur, J. Rushen, and D. B. Haley. 2016. Calving management practices on Canadian dairy farms: Prevalence of practices. Journal of Dairy Science 99(3):2391-2404. https://doi.org/10.3168/jds.2015-9641.

Villettaz Robichaud, M., D. L. Pearl, S. M. Godden, S. J. LeBlanc, and D. B. Haley. 2017a. Systematic early obstetrical assistance at calving: I. Effects on dairy calf stillbirth, vigor, and passive immunity transfer. Journal of Dairy Science 100(1):691-702. https://doi.org/10.3168/jds.2016-11213.

Villettaz Robichaud, M., D. L. Pearl, S. M. Godden, J. Rushen, S. J. LeBlanc, and D. B. Haley. 2017b. Systematic early obstetrical assistance at calving: II. Effects on dairy heifer calf growth, health, and survival to weaning. Journal of Dairy Science 100(1):703-712. https://doi.org/10.3168/jds.2016-11214.

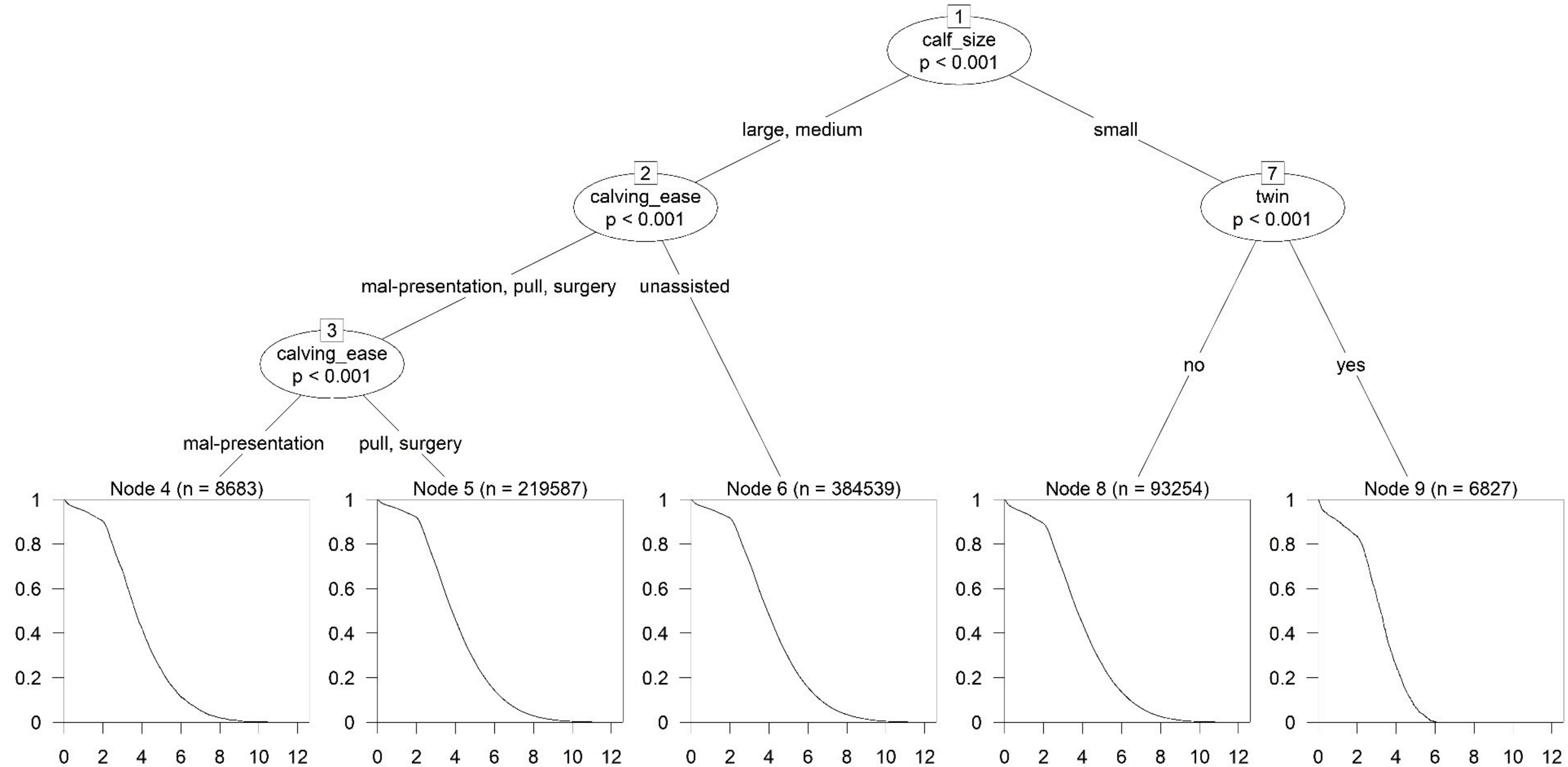
Wathes, D. C., J. S. Brickell, N. E. Bourne, A. Swali, and Z. Cheng. 2008. Factors influencing heifer survival and fertility on commercial dairy farms. Animal 2(8):1135-1143. https://doi.org/10.1017/S1751731108002322.

Wells, S. J., D. A. Dargatz, and S. L. Ott. 1996. Factors associated with mortality to 21 days of life in dairy heifers in the United States. Preventive Veterinary Medicine 29(1):9-19. https://doi.org/10.1016/S0167-5877(96)01061-6.

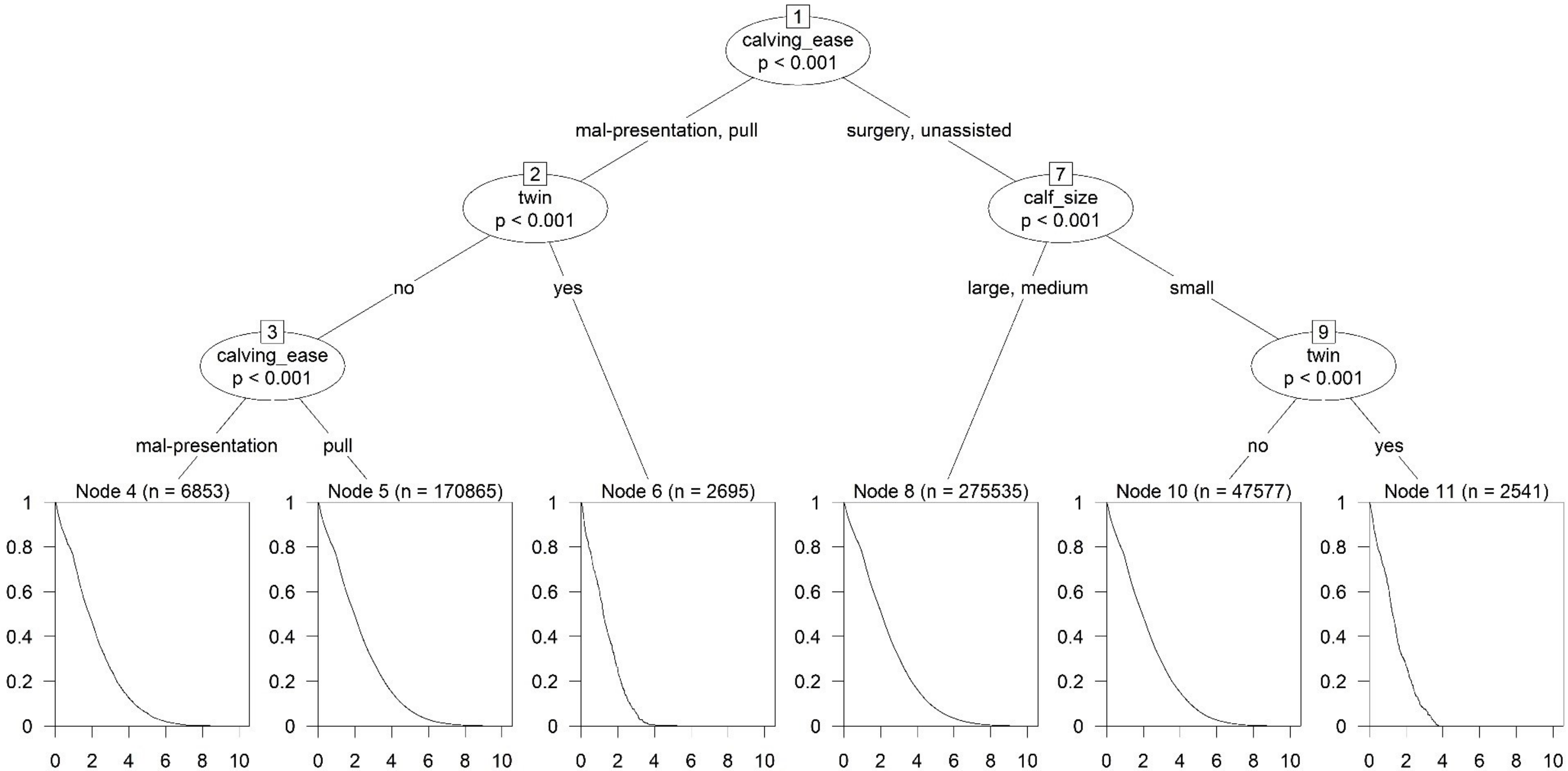
Windeyer, M. C., K. E. Leslie, S. M. Godden, D. C. Hodgins, K. D. Lissemore, and S. J. LeBlanc. 2014. Factors associated with morbidity, mortality, and growth of dairy heifer calves up to 3 months of age. Preventive Veterinary Medicine 113(2):231-240. https://doi.org/10.1016/j.prevetmed.2013.10.019.

Zhou, Y. and J. J. McArdle. 2015. Rationale and applications of survival tree and survival ensemble methods. Psychometrika 80(3):811-833. https://doi.org/10.1007/s11336-014-9413-1.

3.8. Supplementary material



Supplementary Figure 3.1. Conditional inference survival tree of culling events showing the effects of birth conditions on the offspring length of life (total N = 712,890). Circles represent both the birth conditions and the cut-off values found by the algorithm for splitting the offspring into smaller groups. Numbers in the squares on top of the circles indicate the node number. Squares at the bottom represent the terminal nodes. They contain Kaplan-Meier survival curves of each group and the sample size on top (n).



Supplementary Figure 3.2. Conditional inference survival tree of culling events showing the effects of birth conditions on the offspring length of productive life (total N = 506,066). Circles represent both the birth conditions and the cut-off values found by the algorithm for splitting the offspring into smaller groups. Numbers in the squares on top of the circles indicate the node number. Squares at the bottom represent the terminal nodes. They contain Kaplan-Meier survival curves of each group and the sample size on top (n).

Connecting statement 2

Calving ease, calf size, and twinning were shown to have a long-term effect on offspring longevity, which builds on to the scarce body of literature available on the topic. However important, they only indicate how a dairy calf comes into existence. After calving, the calf is subjected to many other early life conditions that require further investigation. In addition, animal lifetime productivity and animal's lifetime profitability are intertwined with longevity, making them important to be considered in simultaneously. The study described in the next chapter was conducted with the objective to investigate some of these aspects. It focused on animal-level early life outcomes and their relationship with indicators of longevity, productivity, and profitability were analysed.

CHAPTER 4 – The effect of early life animal outcomes on subsequent longevity, productivity, and profitability: An animal-based cohort study

Gabriel M. Dallago[1,*], Ibrahim Elsohaby[2], J T. McClure[3], René Lacroix[4], and Elsa Vasseur[1]

[1] Department of Animal Science, McGill University, Sainte-Anne-de-Bellevue, Quebec, H9X 3V9, Canada

[2] Department of Animal Medicine, Faculty of Veterinary Medicine, Zagazig University, Zagazig City 44511, Sharkia Province, Egypt

[3] Department of Health Management, Atlantic Veterinary College, University of Prince Edward Island, Charlottetown, Prince Edward Island, C1A 4P3, Canada

[4] Lactanet, Sainte-Anne-de-Bellevue, Quebec, H9X 3R4, Canada

* Corresponding author: gabriel.dallago@mail.mcgill.ca

Manuscript submitted to the Journal of Dairy Science: JDS.2022-21962

4.1. Abstract

Milk production efficiency and overall profitability can be improved by raising resilient animals that are more likely to reach their potential. Identify such animals based on early life outcomes is a possibility requiring further investigation. The objective of this study was to analyze the associations between early life animal outcomes and their subsequent length of life (LL), length of productive life (LPL), lifetime cumulative energy-corrected milk (ECM), and lifetime cumulative milk value. Data from two cohorts of animals were analyzed based on the availability of lifetime variables. After data cleaning, cohort 1 was composed of 367 calves with LL data that were born between June 2014 and November 2015 in 8 dairy herds from New Brunswick (Canada). Cohort 2 was composed of a subset of 273 calves from cohort 1 with data on LPL, lifetime cumulative ECM, and lifetime cumulative milk value. In addition to herd identification, birth year, and birth season, the following early life animal outcomes were evaluated: birth weight, weaning weight, weaning age, weaning average daily gain, concentration of serum immunoglobulin G, occurrence of navel infection, of scours, and of pneumonia, and if animals received antibiotic treatment between birth and weaning. Multiple imputation was used to handle missing data and an 80:20 split ratio was used to create the training and validation data sets, respectively. For both cohorts and lifetime variables, different machine learning algorithms were trained with the training

data sets and using 5-fold cross-validation repeated 10 times. The best models, identified based on the lowest prediction error in the validation data sets, were used to estimate both variable importance and accumulated local effect (ALE) between early life animal outcomes and each lifetime variable. On average, prediction errors were relatively high, with a mean arctangent absolute percentage error ranging from 33.1% on LL to 49.8% on ECM. Herd was shown to be the most important variable and had an average impact of 12%, 11%, 15%, and 14% on LL, LPL, lifetime cumulative ECM, and lifetime cumulative milk value predictions, respectively. The remaining indicators had, on average, an impact equal to or below to 6% on all lifetime variables. Though there was a great variability between herds, the higher the ALEs for LPL, the higher the ALEs for both lifetime cumulative ECM and lifetime cumulative milk value. These results highlight the importance of herd management decisions at different stage of the animal's life to improve dairy farming efficiency.

4.2. Introduction

Milk production efficiency in dairy farms needs to increase to meet the expected demand for dairy products. Driven by population growth, urbanization, and economic growth (Vos and Bellù, 2019), the consumption of milk and dairy products is expected to increased by 46% by 2050 compared to 2005/2007 in low- and middle-income countries, and by 19% in high-income countries over the same period (Alexandratos and Bruinsma, 2012). Strategic genetic selection, improved health practices, improved nutrition, and management of environmental conditions have all contributed to increase milk yield in the past (Collier et al., 2006, Shook, 2006). Between 1961 and 2018, the total increase of milk production per animal per year ranged from 1,055 kg to 7,393 kg in the different high milk-producing countries (Dallago et al., 2021b). However, both the scarcity of resources and the pressure for sustainable intensification impose constraints on further milk production increase. Therefore, it is important to continue optimizing management practices (Pretty and Bharucha, 2019) to maximize the efficiency in which resources are used to sustainably increase milk production and the profitability of dairy farming.

Efficiency could be improved by identifying and raising resilient animals. Resilient animals have a high adaptability to challenges and cumulative good health and fertility, resulting in a greater longevity (Adriaens et al., 2020) because cow longevity is associated with the animal's ability to avoid early culling due to health and reproduction issues (Dallago et al., 2021b), which would imply being productive for longer and, therefore, more profitable (Brickell and Wathes, 2011, Boulton et al., 2017, Habel et al., 2021). Studies on cow longevity often focused on lactating cows. However, there is a growing body of literature on the long term effects of early life conditions on adult animal performance and productivity. Calving difficulty, birth size, and twining were associated with the longevity of Canadian Holstein offspring (Dallago et al., 2021a). In addition, the occurrence of calfhood diarrhea had a negative effect on first lactation 305-day energy-corrected milk (**ECM**) in Swedish dairy cows (Svensson and Hultgren, 2008). A negative effect of diarrhea during weaning on days to conception and age at first calving was also observed (Aghakeshmiri et al., 2017). In turn, age at first calving was shown to be associated with the ability of animals to remain in the herd at least until the end of the first lactation (Sherwin et al., 2016).

Therefore, we hypothesized that early life animal outcomes are associated with the subsequent longevity, productivity, and profitability of dairy animals. The objective of this study was to analyze the associations between body weight, health events, and immune status of dairy calves (i.e., early life animal outcomes) on their subsequent length of life (**LL**) length of productive life (**LPL**), lifetime cumulative ECM, and lifetime cumulative milk value.

4.3. Material and methods

A cross-sectional study was conducted based on early life indicators of a cohort of animals born in eight dairy herds from New Brunswick (Canada). Longevity, production, and profitability indicators were provided by the Canadian DHI Agency Lactanet (Sainte-Anne-de-Bellevue, Quebec, Canada). All statistical analyses were carried out using the R software, version 4.1.1 (https://www.r-project.org/). The R code that supports the findings of this study can be found in the following public GitHub repository: https://github.com/CowLifeMcGill/PEI_cohort_study.

4.3.1. Dataset

Early life animal outcomes were available for 463 Holstein claves born in eight dairy herds from New Brunswick (Canada), between June 2014 and November 2015. Calves involved in the study were monitored from birth to weaning. Blood samples were collected within 24 to 48 h after calving according to a protocol approved by the University of Prince Edward Island (UPEI) Research Ethics Board (protocol #6006206). Calf serum immunoglobulin G (**IgG**; mg/dL) was measured using infrared spectroscopy (Elsohaby et al., 2014), and body weight (kg) at birth and around weaning was measured using an electronic scale. In addition to serum IgG and body weight, the following early life indicators were collected and reported by the farmers: Animal ID, herd ID, calving date, weaning date, and occurrence of navel infection (yes or no), scours (yes or no), pneumonia (yes or no), and if animals received antibiotic treatment (yes or no) between birth and weaning.

The DHI outcomes of milk production (kg), fat production (kg), protein production (kg), milk value ($CAD), culling date as well as animal and herd identification (**ID**) were provided by Lactanet (Sainte-Anne-de-Bellevue, Quebec, Canada) for 439 Holstein cows born in the same eight dairy herds from New Brunswick (Canada) between June 2014 and November 2015, and which participated in the study.

4.3.2. Data handling and cleaning

Cows that were not yet culled at the time of the DHI data extraction were removed (N = 65), since it was not possible to calculate longevity metrics, lifetime cumulative milk production, nor milk value for those animals. Similar to Dallago et al. (2021a), LL (year) and LPL (year) were the longevity metrics considered in this study. The LL was calculated as the length of time (year) between birth and culling while the LPL was the length of time (year) between the first calving and culling. Milk, fat, and protein production as well as milk value were aggregated by calculating the sum over the lifetime of the animals. Lastly, lifetime cumulative ECM (kg) was calculated as ECM (lifetime cumulative kg) = 12.55 × fat (lifetime cumulative kg) + 7.39 × protein (lifetime cumulative kg) + 0.2595 × milk yield (lifetime cumulative kg).

Duplicate observations on early life indicators (i.e., repeated animal ID, but having different entry values) were removed (N = 7). In addition, negative values on serum IgG concentration (N = 2) were treated as missing data (Figure 4.1). Body weight was measured twice for each animal: it was first measured at birth and a second time around the weaning date (about 2 months old). Consequently, the model developed by Cue et al. (2012) for Canadian Holstein cows was used to estimate the body weight of the animals at the exact weaning date. For animals that were weighted at weaning (N = 63), the original recorded body weight was kept. In addition, weaning average daily gain (**ADG**; kg/day) and weaning age (month) were also estimated.

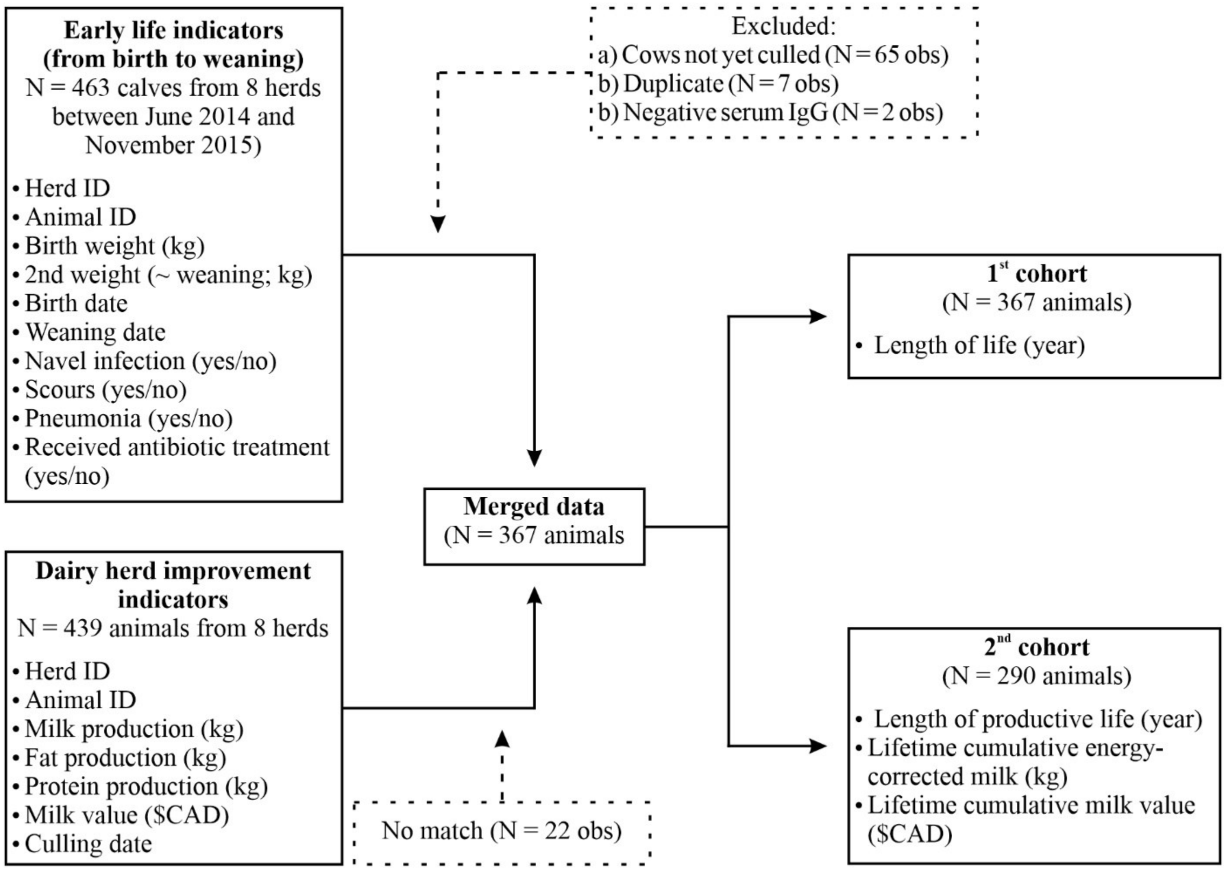


Figure 4.1. Flowchart indicating data handling and cleaning steps to create the two cohorts of data submitted for statistical analysis. The solid-line rectangles represent the data and the dashed-line rectangles represent processes. Obs = Observation.

4.3.3. Creating working data set

The DHI lifetime variables and early life indicators were merged based on animal ID and herd ID (N = 367). It was not possible to calculate LPL, lifetime cumulative ECM, nor lifetime cumulative milk value for all animals because some of them were culled before calving for the first time (N = 77). Therefore, two cohorts of data were created. The first cohort included data of all animals in which there was a match between files (N = 367), and it was used to analyze the LL. The second cohort only included information from animals that were not culled before calving for the first time (N = 290). Some animals were culled shortly after calving for the first time and did not have data on milk production and milk value (N = 17). Data from those animals were also removed from the second cohort. Descriptive statistics on both cohorts are shown on Table 4.1 and Table 4.2.

Table 4.1. Descriptive statistics for numerical indicators and outcome variables of two cohorts of animals from eight New Brunswick (Canada) dairy farms. Cohorts were defined based on the availability of the outcome variables length of life (Cohort 1), length of productive life (Cohort 2), lifetime cumulative energy-corrected milk (ECM; Cohort 2), and lifetime cumulative milk value (Cohort 2).

Variable	Cohort 1[1] (N = 367)					Cohort 2[1] (N = 273)				
	NA	Mean	SD	Min	Max	NA	Mean	SD	Min	Max
Early life indicators										
Birth weight (kg)	17	41.5	5.21	27.0	59.0	9	41.7	5.03	27.0	59.0
Weaning weight (kg)	42	91.7	13.2	60.0	130.8	22	92.3	13.6	62.0	130.8
Weaning age (month)	35	2.13	0.38	1.38	3.65	18	2.13	0.40	1.38	3.65
Weaning ADG (kg)[2]	44	0.769	0.099	0.365	1.041	23	0.773	0.102	0.365	1.04
Serum IgG (mg/dL)	12	1,607.7	833.1	30.6	4,779.4	0	1,604.7	840.8	30.6	4,779.4
Outcomes										
Length of life (year)	0	3.78	1.63	0.03	6.85	---	---	---	---	---
Length of productive life (year)	---	---	---	---	---	0	2.42	1.23	0.11	4.88
Lifetime cumulative ECM (kg)	---	---	---	---	---	0	27,788.6	15,979.9	757.9	67,930.4
Lifetime cumulative milk value ($ CAD)	---	---	---	---	---	0	19,474.0	11,399.1	528.0	48,447.8

[1] NA = Missing observations; SD = Standard deviation; Min = Minimum; Max = Maximum.
[2] Average weight daily gain from birth to weaning.

Table 4.2. Descriptive statistics of categorical indicators of two cohorts of animals from eight New Brunswick (Canada) dairy farms. Cohorts were defined based on the availability of the outcome variables length of life (Cohort 1), length of productive life (Cohort 2), lifetime cumulative energy-corrected milk (Cohort 2), and lifetime cumulative milk value (Cohort 2).

Early life indicator		**Cohort 1 (N = 376)**		**Cohort 2 (N = 273)**	
		N	**%**	**N**	**%**
Navel infection	No	345	94.01	260	95.24
	Yes	7	1.91	5	1.83
	Missing	15	4.09	8	2.93
Scours	No	343	93.46	257	94.14
	Yes	9	2.45	8	2.93
	Missing	15	4.09	8	2.93
Pneumonia	No	343	93.46	261	95.60
	Yes	9	2.45	4	1.47
	Missing	15	4.09	8	2.93
Antibiotics[1]	No	259	70.57	200	73.26
	Yes	48	13.08	31	11.36
	Missing	60	16.35	42	15.38
Year	2014	300	81.7	221	80.95
	2015	67	18.3	52	19.05
Season	Fall	160	43.6	121	44.32
	Summer	121	33.0	85	31.14
	Winter	86	23.4	67	24.54
Herd	1	41	11.17	34	12.45
	2	15	4.09	14	5.13
	3	68	18.53	55	20.15
	4	52	14.17	34	12.45
	5	44	11.99	31	11.36
	6	96	26.16	66	24.18
	7	29	7.90	25	9.16
	8	22	5.99	14	5.13

[1] During the weaning period.

Multiple imputation was used to handle missing observations instead of case wise deletion to maximize the number of observations. After data cleaning and merging, the percentage of missing observations ranged from 3.27% on serum IgG concentration (n = 12) to 16.35% on received antibiotic treatment (n = 60). The average percentage of missing observations per herd was 4.12% (Standard deviation; SD = 3.56%) and ranged from 1.0 to 12.0%. The function *mice* from the package *mice* (van Buuren and Groothuis-Oudshoorn, 2011) was used to create 10 multiple imputed versions of the data using the random forest method. In short, for each numeric or categorical indicator with missing data, the remaining indicators were used to create a random forest model to impute the missing values. The distribution of the imputed data was visually compared to the data with non-missing values. In addition to multiple imputation being a better approach compared to other methodologies in order to increase power and accuracy of the data analysis (van Buuren, 2019), random forest is able to handle complex interactions between variables even in conditions where there is a high number of missing observations (Tang and Ishwaran, 2017).

4.3.4. Data analysis

The associations between early life animal outcomes and longevity, productivity, and profitability were analyzed using the machine-learning algorithms recursive partitioning and regression tree (**RPART**), gradient boosting machine (**GBM**), random forest (**RF**), and support vector machine (**SVM**) with a radial basis kernel. For each of the imputed versions of the data, an 80:20 ratio was used to create the training and validation data sets, respectively, based on the distribution of each lifetime variable (i.e., LL, LPL, ECM, or milk value). All models were trained with 5-fold cross-validation repeated 10 times on the training data sets using the *caret* package (Kuhn, 2020) by specifying the methods *rpart2*, *gbm*, *ranger*, and *svmRadialSigma*, respectively for the algorithms RPART, GBM, RF, and SVM. Hyperparameters for these models were tuned using adaptive resampling, which resamples the hyperparameter tuning grid by concentrating on values closer to the identified optimal setting (Kuhn, 2014, 2020). The models were evaluated based on the coefficient of determination (**R^2**), root mean squared error (**RMSE**), mean absolute error (**MAE**), and mean arctangent absolute percentage error (**MAAPE**; Kim and Kim, 2016) calculated using the validation data sets. The best model was defined as having the highest R^2 and the lowest RMSE, MAE, and MAAPE averages across all the imputed versions of the data.

Using the best model, variable importance and accumulated local effect (**ALE**) were used to describe the association between early life animal outcomes (i.e., either numeric or categorical) and longevity, productivity, and profitability. Variable importance was calculated using permutation. In short, this is a model agnostic approach that measures the prediction error of the model after shuffling the values for the early life indicators, which changes the relationship between the outcome and early life indicators. Shuffling the values of important early life indicators would result in an increase of the error while the error would remain unchanged for early life indicators that are not important (Molnar, 2019). Additionally, ALE indicates the average influence of the early life indicators in predicting the outcome variables (Molnar, 2019). Variable importance and ALE were estimated using the functions *FeatureImp* and *FeatureEffect*, respectively, from the R package *iml* (Molnar et al., 2018) and based on each of the complete imputed data sets (i.e., training and validation data set combined). Average and standard deviations were calculated to aggregate the results of each early life indicator across all imputed version of the data.

4.4. Results

A weak relationship was observed between early life animal outcomes and their subsequent longevity (i.e., LL and LPL), productivity (i.e., lifetime cumulative ECM), and profitability (i.e., lifetime cumulative milk value). Overall, the prediction errors were relatively high, with an average MAAPE ranging from 33.1% on LL to 49.8% on lifetime cumulative ECM in the training data set whereas it ranged from 39.2% on LL to 53.7% on lifetime cumulative milk value in the validation data set. (Supplementary Table 4.1). Based on the training data set, the LL, LPL, lifetime cumulative ECM, and lifetime cumulative milk value were best predicted by RF models (Supplementary Table 4.1). In the validation data set, however, only LL was best predicted by a RF model, with the remaining lifetime variables having no model outperforming the others in most of the evaluation metrics (Supplementary Table 4.1). Therefore, the effect of the early life outcome measures on all lifetime variables were further described based on the RF models since it had the best performance on the training data set.

Herd, as a categorical variable, was the most important indicator associated with all four lifetime variables. It had an average impact of 12%, 11%, 15%, and 14% on LL, LPL, lifetime cumulative ECM, and lifetime cumulative milk value predictions respectively, whereas the remaining indicators had, on average, an impact equal to or below to 6% on all lifetime variables (Figure 4.2).The incidence of diseases were the least important indicators while weight related indicators and serum IgG were of intermediate importance (Figure 4.2).

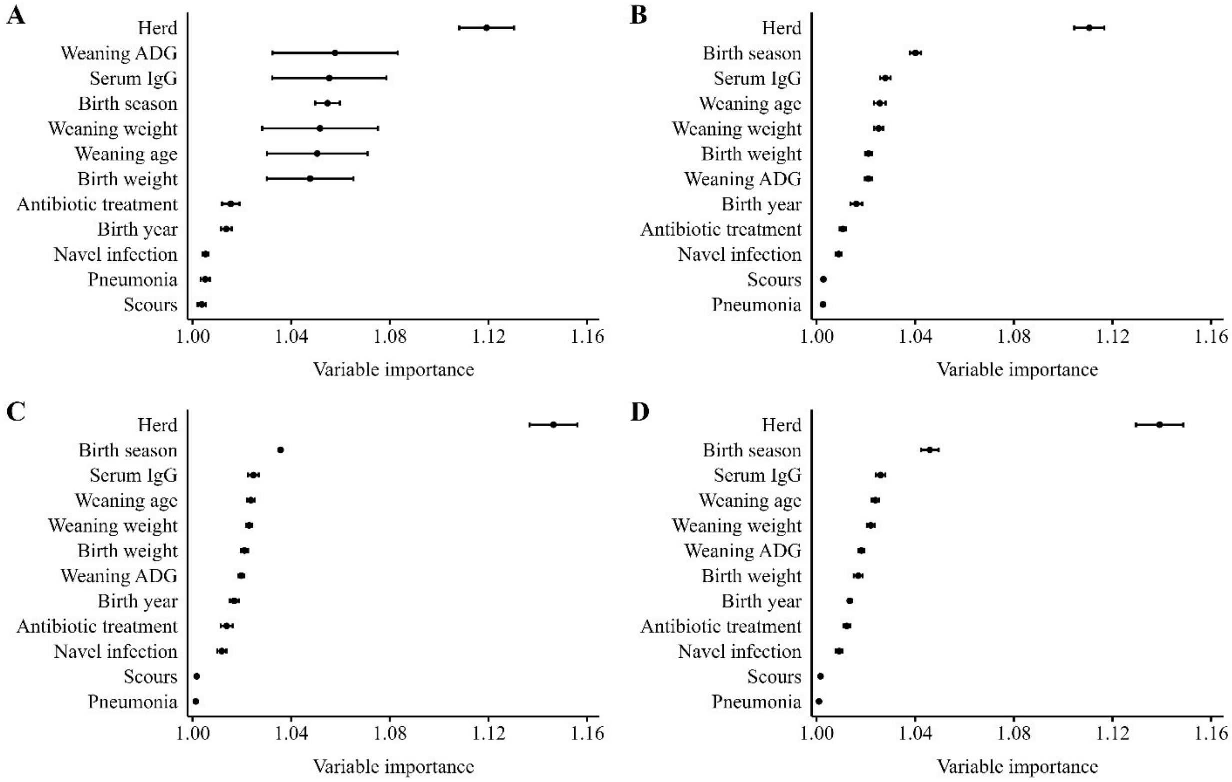


Figure 4.2. The importance (x axis) of indicator variables (y axis; from highest to lowest) to predict length of life (LL; **A**), length of productive life (LPL; **B**), lifetime cumulative energy-corrected milk (ECM; **C**), and lifetime cumulative milk value (**D**). Mean (•) and standard deviation (bar) were calculated across all 10 multiple imputed versions of the data. Variable importance indicates the increase in model error prediction. It is measured as the ratio of root mean squared error of the original model and after shuffling the values of the indicators (Molnar, 2019).

The ALE estimated for categorical early life animal outcomes are shown on Table 4.3. Surprisingly, the occurrence of navel infection, scours, and pneumonia during the pre-weaning period were associated with higher LL, LPL, lifetime cumulative ECM and lifetime cumulative milk value. On the other hand, all outcomes evaluated in this study had a lower ALE for animals born in the winter compared to animals born in the fall. Though there was a great variability between herds, the higher the ALE for LPL, the higher the ALEs for both lifetime cumulative ECM and lifetime cumulative milk value. Such results were not observed for LL, in which high ALEs in LL were not followed by high ALEs in the other lifetime variables. However, the herd with the lowest ALE for LL also had the lowest ALE for LPL, lifetime cumulative ECM, and lifetime cumulative milk value, whereas the herd with the highest LL also had the highest LPL, lifetime cumulative ECM, and lifetime cumulative milk value.

Table 4.3. Accumulated local effect (ALE) of categorical indicators on each of the outcomes evaluated. Means ± standard deviations were calculated across all 10 multiple imputed versions of the data.

Early life indicator		Outcome			
		Length of life (year) (N = 376)	**Length of productive life (year) (N = 273)**	**Lifetime cumulative energy-corrected milk (kg) (N = 273)**	**Lifetime cumulative milk value ($CAD) (N = 273)**
Navel infection	No	-0.01 ± 0.001	-0.01 ± 0.001	-103.5 ± 8.67	-53.1 ± 8.14
	Yes	0.38 ± 0.05	0.38 ± 0.05	5,475.2 ± 689.62	2,818.1 ± 522.26
Scours	No	-0.01 ± 0.002	-0.004 ± 0.0002	-43.5 ± 3.42	-19.7 ± 0.88
	Yes	0.26 ± 0.07	0.12 ± 0.005	1,442.3 ± 113.15	653.8 ± 29.31
Pneumonia	No	-0.001 ± 0.002	-0.002 ± 0.0001	-30.0 ± 1.09	-1.2 ± 0.61
	Yes	0.05 ± 0.07	0.13 ± 0.004	2,019.0 ± 72.99	79.4 ± 41.33
Antibiotics[1]	No	0.002 ± 0.01	0.001 ± 0.003	-11.9 ± 34.68	52.3 ± 29.46
	Yes	-0.01 ± 0.03	-0.01 ± 0.02	93.2 ± 261.02	-392.1 ± 215.52
Birth year	2014	-0.01 ± 0.001	-0.02 ± 0.001	-227.0 ± 9.29	-79.9 ± 5.04
	2015	0.03 ± 0.01	0.08 ± 0.004	964.6 ± 39.47	339.6 ± 21.41
Birth season	Fall	0.03 ± 0.01	0.03 ± 0.002	291.9 ± 19.71	273.6 ± 13.62
	Summer	-0.03 ± 0.01	-0.01 ± 0.004	-39.1 ± 28.36	522.5 ± 22.19
	Winter	-0.01 ± 0.01	-0.04 ± 0.004	-477.6 ± 44.46	-1,157.0 ± 46.16
Herd	1	-0.01 ± 0.03	0.01 ± 0.01	1,213.8 ± 79.36	408.5 ± 54.15
	2	0.16 ± 0.06	0.05 ± 0.01	470.8 ± 138.06	166.8 ± 76.00
	3	0.02 ± 0.03	-0.16 ± 0.01	-1,340.1 ± 219.61	-315.4 ± 130.49
	4	-0.14 ± 0.03	-0.17 ± 0.01	-984.7 ± 95.48	-1,338.0 ± 78.78
	5	-0.13 ± 0.03	0.25 ± 0.01	5,385.0 ± 125.37	3,448.4 ± 86.59
	6	-0.1 ± 0.01	-0.11 ± 0.01	-4,694.9 ± 119.49	-3,302.9 ± 62.69
	7	0.69 ± 0.04	0.53 ± 0.01	5,995.1 ± 195.30	4,825.8 ± 82.22
	8	-0.04 ± 0.04	-0.01 ± 0.01	3,741.0 ± 325.53	2,647.0 ± 127.81

[1] During the weaning period.

Animals born with a low or high weight were associated with a reduction on LL, LPL, lifetime cumulative ECM, and lifetime cumulative milk value (Figure 4.3). Except for lifetime cumulative ECM, a low birth weight had, on average, a more negative ALE as opposed to a high birth weight. On the other hand, the higher the weaning weight, the lower the average ALE observed for all lifetime variables (Figure 4.4). A high weaning age was associated with a reduction on both LL and lifetime cumulative milk value whereas the opposite was observed for LPL and lifetime cumulative ECM (Figure 4.5). We found a reduction on all lifetime variables starting at the low end of the weaning ADG and going up to 0.800 kg/day, whereas an increase was observed thereafter, except for lifetime cumulative milk value (Figure 4.6). Lastly, we found that the higher the concentration of serum IgG, the higher the LL, LPL, lifetime cumulative ECM, and lifetime cumulative milk value (Figure 4.7). However, the shape of the association was not linear and the ALEs at the high end of the serum IgG values were more than 100% higher than what we found at the low end (Figure 4.7).

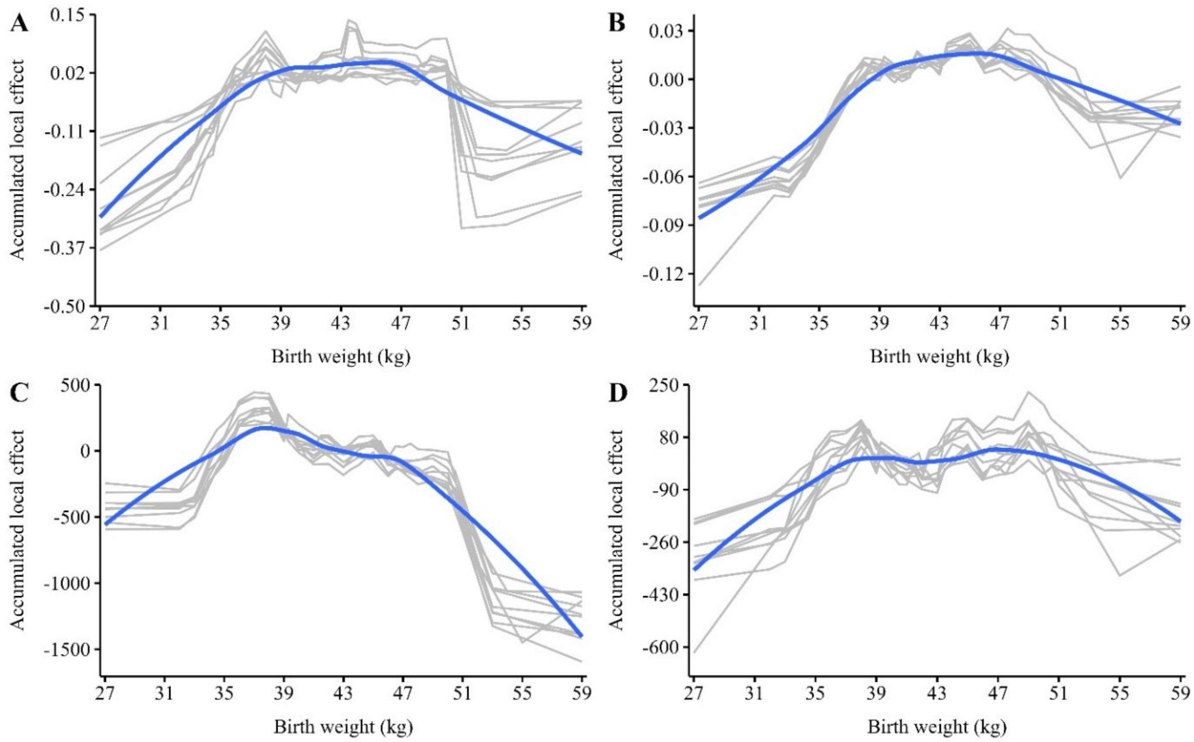


Figure 4.3. The accumulated local effect (ALE) of birth weight on length of life (**A**; LL), length of productive life (**B**; LPL), lifetime cumulative energy-corrected milk (**C**; ECM), and lifetime cumulative milk value (**D**) across all 10 multiple imputed versions of the data (grey lines) followed by a loess trend line (blue).

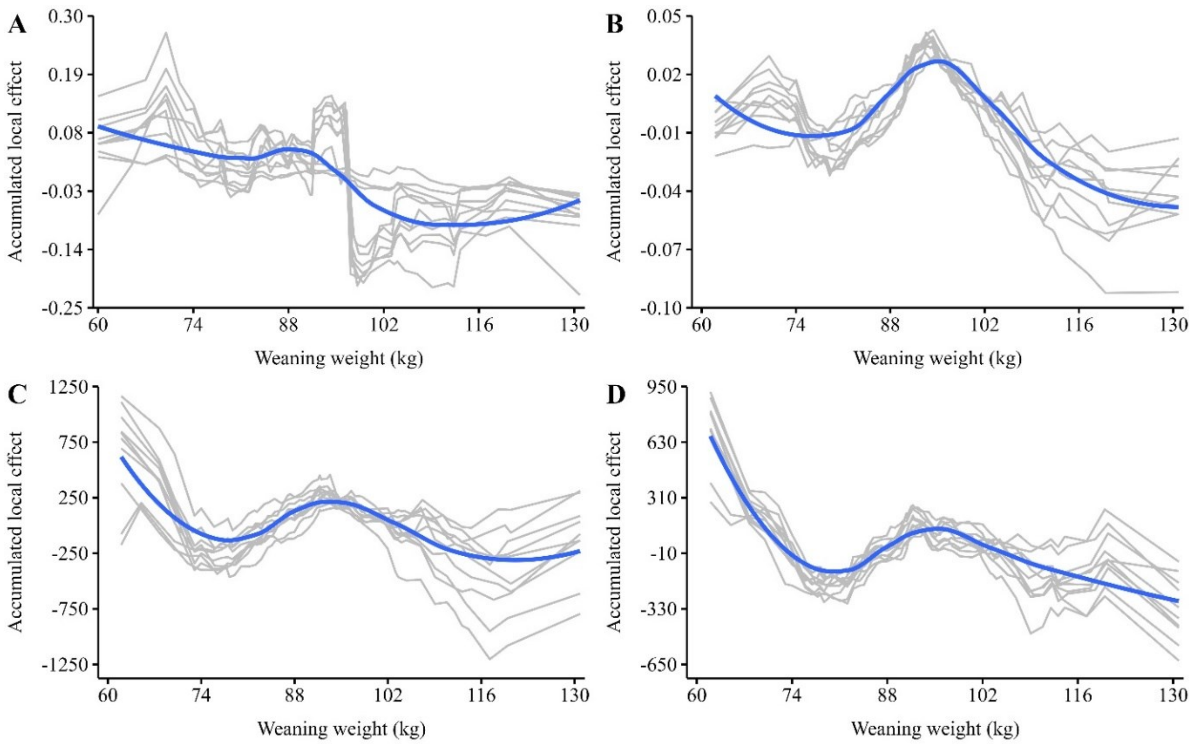


Figure 4.4. The accumulated local effect (ALE) of weaning weight on length of life (**A**; LL), length of productive life (**B**; LPL), lifetime cumulative energy-corrected milk (**C**; ECM), and lifetime cumulative milk value (**D**) across all 10 multiple imputed versions of the data (grey lines) followed by a loess trend line (blue).

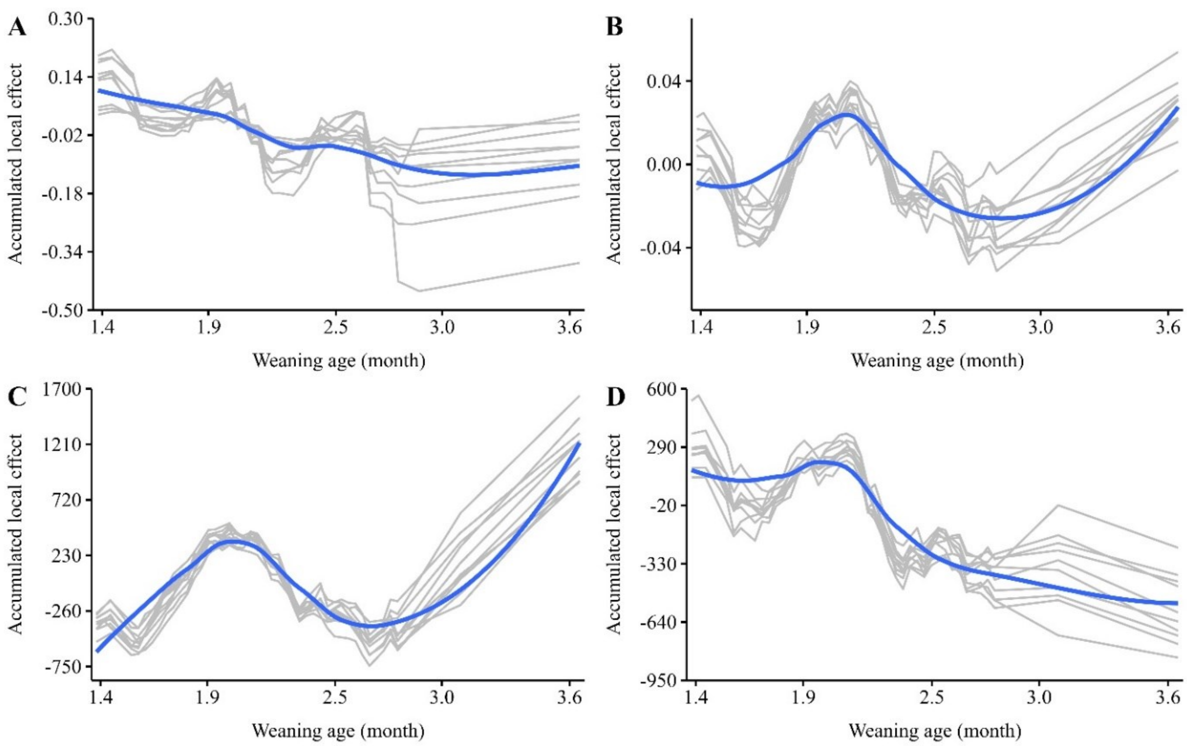


Figure 4.5. The accumulated local effect (ALE) of weaning age on length of life (**A**; LL), length of productive life (**B**; LPL), lifetime cumulative energy-corrected milk (**C**; ECM), and lifetime cumulative milk value (**D**) across all 10 multiple imputed versions of the data (grey lines) followed by a loess trend line (blue).

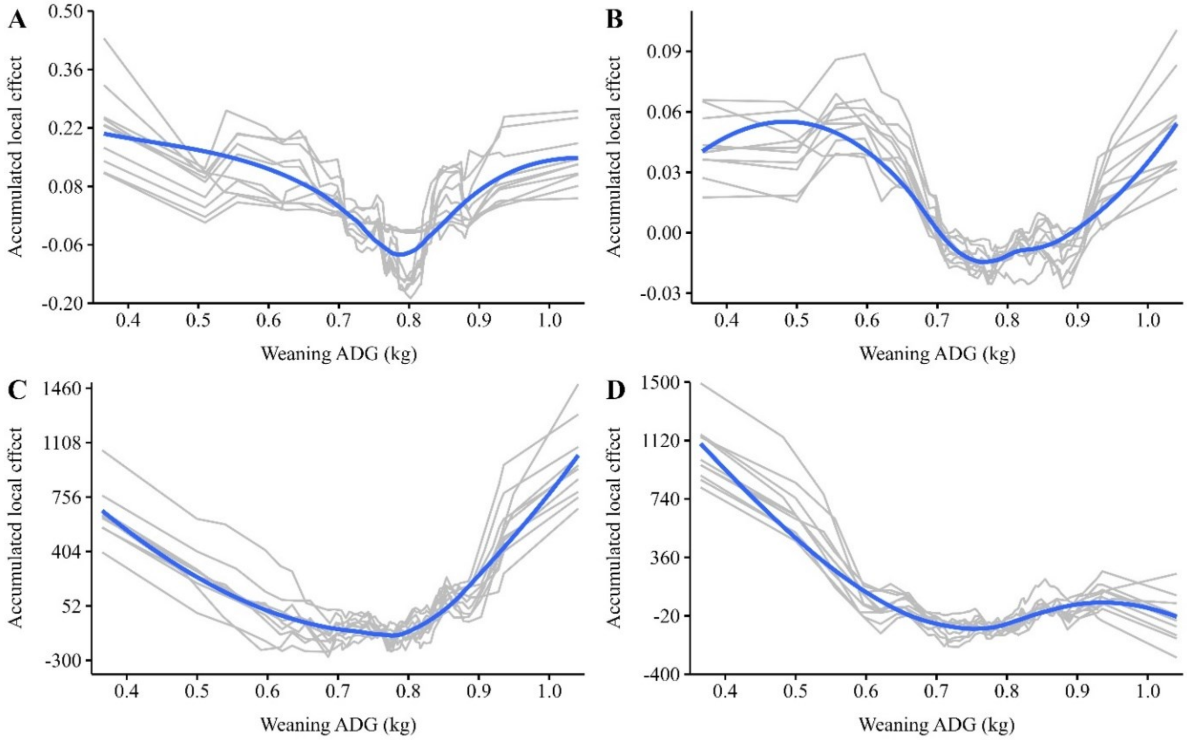


Figure 4.6. The accumulated local effect (ALE) of average daily gain (ADG) of weight from birth to weaning on length of life (**A**; LL), length of productive life (**B**; LPL), lifetime cumulative energy-corrected milk (**C**; ECM), and lifetime cumulative milk value (**D**) across all 10 multiple imputed versions of the data (grey lines) followed by a loess trend line (blue).

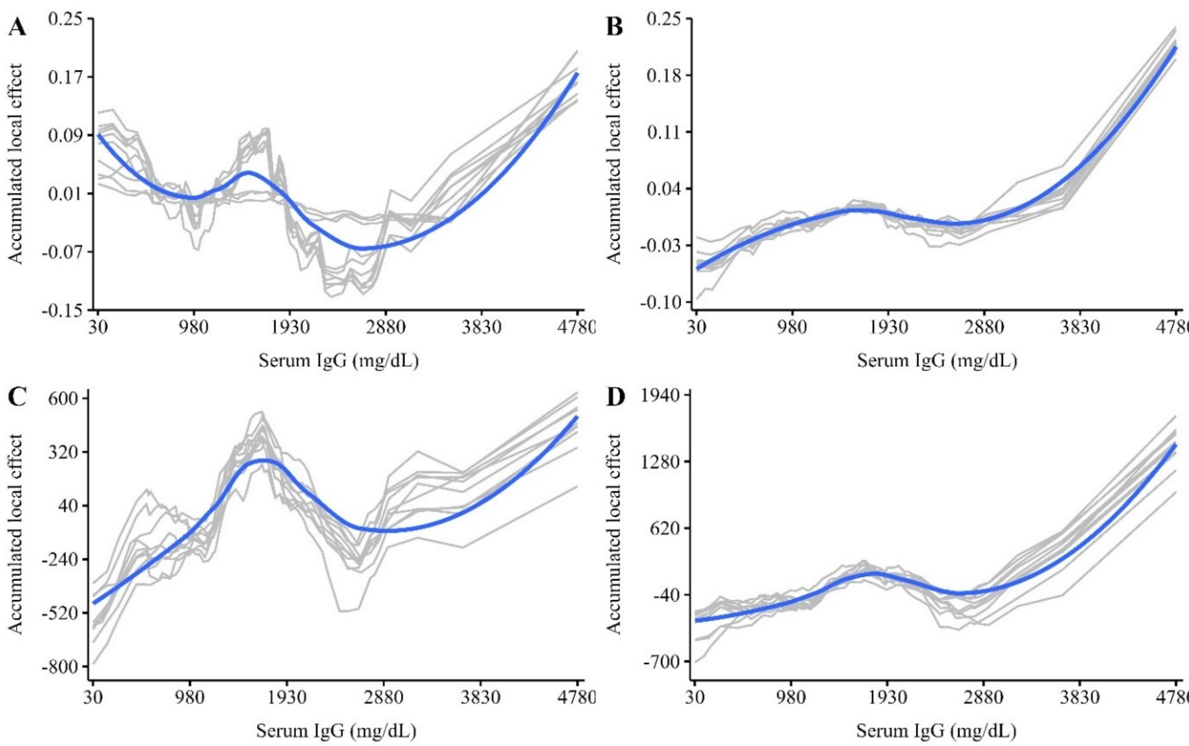


Figure 4.7. The accumulated local effect (ALE) of serum IgG on length of life (**A**; LL), length of productive life (**B**; LPL), lifetime cumulative energy-corrected milk (**C**; ECM), and lifetime cumulative milk value (**D**) across all 10 multiple imputed versions of the data (grey lines) followed by a loess trend line (blue).

4.5. Discussion

The associations between early life animal outcomes and longevity (i.e., LL and LPL), productivity (i.e., lifetime cumulative ECM), and profitability (i.e., lifetime cumulative milk value) were evaluated based on two cohorts of animals from eight dairy herds. Different machine learning models were trained and the best ones were used to describe the associations under consideration in this study.

Herd was the most important indicator associated with animal longevity, productivity, and profitability. The longevity of dairy cows in high milk producing systems is mostly dictated by culling decisions carried out by farmers. Even though longevity has decreased in most high milk producing countries in the past decades (Dallago et al., 2021b), increasing cow longevity is a strategy to improve dairy farming sustainability since it is associated with its economic efficiency (Delgado et al., 2017, Habel et al., 2021) environmental impact (Grandl et al., 2019), and social acceptability (Bruijnis et al., 2013, Rushen and Passillé, 2013). Some herds were associated with greater longevity in our study, but not all of them, which is in agreement with previous studies (Beaudeau et al., 2000). In turn, this highlights unique management styles between farms and the consideration of different underlining factors to carry out culling decisions. Previous events are often neglected in culling decisions. The occurrence of costly health issues in previous lactations of lactating dairy cows, for example, are not usually taken into account by farmers (Beaudeau et al., 2000), even though they could compromise the profitability of the current lactation. It was not surprising, therefore, to find a weak relationship between the remaining early life animal outcomes and subsequent longevity, productivity, and profitability in our study.

The prevalence of health events observed in the present cohort of animals was low. For instance, Medrano-Galarza et al. (2018) reported a 17% prevalence of bovine respiratory diseases based on 17 Canadian dairy farms while Karle et al. (2019) reported an overall average of 7% based on three farms from California (USA) and Closs Jr and Dechow (2017) reported a 13% and 9% prevalence of pneumonia and scours, respectively, in a dairy farm in Pennsylvania (USA). This could indicate farmers' inability to properly diagnose health events in dairy calves in our study as it was reported in the assessment of other health issues in adult cattle such as lameness (Higginson Cutler et al., 2017, Beggs et al., 2019).

Contrary to previous studies, the occurrence of all three health issues was associated with higher longevity, productivity, and profitability. On Swedish dairy cows, the occurrence of severe calfhood respiratory disease was associated with an increase on calving interval of cows on first to third or greater lactations (Hultgren and Svensson, 2010) and navel infections was associated with negative effect on survival heifer survival up to 96 month of age (Britney et al., 1984). Our results might indicate, however, that animals who were properly diagnosed and treated made a full recovery, which in turn did not compromise their subsequent adult performance as opposed to animals that might have been misdiagnosed. For instance, receiving two or more treatment for disease had a negative effect on culling age while no effect was observed in animals receiving only one treatment (Closs Jr and Dechow, 2017). In addition, both the occurrence of health events and the use of antibiotic treatments were recorded as binary variables (i.e., yes or no) and the effect of duration, co-occurrence, and severity of diseases on lifetime traits such as length of productive life and lifetime cumulative ECM remains to be evaluated. Reproductive performance, which is one of the most prevalent reasons for culling dairy cows (Compton et al., 2017, CDIC, 2021), could also provide additional insights on the underlying reasons behind our results in a future study.

Body weight indicators were previously shown to be associated with animal performance and production. Similar to our results, birth size, as a proxy for birth weight, was selected as an important birth condition associated with the LL and LPL of Canadian Holstein calves (Dallago et al., 2021a). In addition, ADG during weaning was reported to have a positive effect on first lactation milk yield was also reported (Svensson and Hultgren, 2008, Soberon and Van Amburgh, 2013, Vacek et al., 2015). Our results suggest that such positive effect of a high ADG could also be extended to the animal lifetime since it was associated with a greater lifetime cumulative ECM. However, a potential tradeoff exists between ADG and weight at weaning since we found a negative association between lifetime cumulative ECM and weaning weight. Monitoring and using body weight as a weaning criterion could be a strategy to avoid negative consequences on adult LL, LPL, and lifetime cumulative ECM.

The efficacy of passive immunity transfer in dairy calves has been shown to be associated with animal performance and production. The concentration of IgG in the serum of animals is used to determine failure to transfer passive immunity (**FTP**). In our study, serum IgG was measured using Fourier-transform infrared spectroscopy, a rapid and reagent-free methodology (Elsohaby et al., 2014), and calves with serum IgG concentration lower than 1,000 mg/dL were considered to be FTP (Lombard et al., 2020). FTP animals were shown to be not only more prone to develop diseases and have a high mortality (Cuttance et al., 2018, Lora et al., 2018, Urie et al., 2018), but also to have an impaired growth during weaning (Windeyer et al., 2014, Hang et al., 2017) and a reduction in both mature equivalent milk and fat production (DeNise et al., 1989). Since we found that increased longevity, productivity, and profitability were associated with higher levels of serum IgG, our results also indicate the presence of a long-term effect of FTP observed later in the adult cattle.

4.6. Conclusion

Associations were found between early life animal outcomes of body weight, health events, and health status and subsequent longevity (length of life and length of productive life), productivity (lifetime cumulative energy-corrected milk), and profitability (lifetime cumulative milk value). Herd was shown to be the most important variable. Additionally, herds associated with the highest longevity were also associated with the highest productivity and profitability, highlighting the importance of managerial decisions at different stage of the animal's life to improve dairy farming efficiency and productivity.

4.7. References

Adriaens, I., N. C. Friggens, W. Ouweltjes, H. Scott, B. Aernouts, and J. Statham. 2020. Productive life span and resilience rank can be predicted from on-farm first-parity sensor time series but not using a common equation across farms. Journal of Dairy Science 103(8):7155-7171. https://doi.org/10.3168/jds.2019-17826.

Aghakeshmiri, F., M. Azizzadeh, N. Farzaneh, and M. Gorjidooz. 2017. Effects of neonatal diarrhea and other conditions on subsequent productive and reproductive performance of heifer calves. Veterinary Research Communications 41(2):107-112. https://doi.org/10.1007/s11259-017-9678-9.

Alexandratos, N. and J. Bruinsma. 2012. World agriculture towards 2030/2050: The 2012 revision. Agrifood Economics Division (ESA) working paper No. 12-03. FAO, Rome, Italy. Accessed on 30 November 2021. https://www.fao.org/publications/card/en/c/de5f0205-/.

Beaudeau, F., H. Seegers, V. Ducrocq, C. Fourichon, and N. Bareille. 2000. Effect of health disorders on culling in dairy cows: a review and a critical discussion. Ann. Zootech. 49(4):293-311. https://doi.org/10.1051/animres:2000102.

Beggs, D. S., E. C. Jongman, P. H. Hemsworth, and A. D. Fisher. 2019. Lame cows on Australian dairy farms: A comparison of farmer-identified lameness and formal lameness scoring, and the position of lame cows within the milking order. Journal of Dairy Science 102(2):1522-1529. https://doi.org/10.3168/jds.2018-14847.

Boulton, A. C., J. Rushton, and D. C. Wathes. 2017. An empirical analysis of the cost of rearing dairy heifers from birth to first calving and the time taken to repay these costs. Animal 11(8):1372-1380. https://doi.org/10.1017/S1751731117000064.

Brickell, J. S. and D. C. Wathes. 2011. A descriptive study of the survival of Holstein-Friesian heifers through to third calving on English dairy farms. Journal of Dairy Science 94(4):1831-1838. https://doi.org/10.3168/jds.2010-3710.

Britney, J. B., S. W. Martin, J. B. Stone, and R. A. Curtis. 1984. Analysis of early calfhood health status and subsequent dairy herd survivorship and productivity. Preventive Veterinary Medicine 3(1):45-52. https://doi.org/10.1016/0167-5877(84)90023-0.

Bruijnis, M. R. N., F. L. B. Meijboom, and E. N. Stassen. 2013. Longevity as an animal welfare issue applied to the case of foot disorders in dairy cattle. Journal of Agricultural and Environmental Ethics 26(1):191-205. https://doi.org/10.1007/s10806-012-9376-0.

CDIC – Canadian Dairy Information Centre. 2021. Culling and replacement rates in dairy herds in Canada. Accessed on 07 July 2021. https://agriculture.canada.ca/en/canadas-agriculture-sectors/animal-industry/canadian-dairy-information-centre/dairy-statistics-and-market-information/dairy-animal-genetics/culling-and-replacement-rates-dairy-herds-canada.

Closs Jr, G. and C. Dechow. 2017. The effect of calf-hood pneumonia on heifer survival and subsequent performance. Livestock Science 205:5-9. https://doi.org/10.1016/j.livsci.2017.09.004.

Collier, R. J., G. E. Dahl, and M. J. VanBaale. 2006. Major advances associated with environmental effects on dairy cattle. Journal of Dairy Science 89(4):1244-1253. https://doi.org/10.3168/jds.S0022-0302(06)72193-2.

Compton, C. W. R., C. Heuer, P. T. Thomsen, T. E. Carpenter, C. V. C. Phyn, and S. McDougall. 2017. Invited review: A systematic literature review and meta-analysis of mortality and culling in dairy cattle. Journal of Dairy Science 100(1):1-16. https://doi.org/10.3168/jds.2016-11302.

Cue, R. I., D. Pietersma, D. Lefebvre, R. Lacroix, K. Wade, D. Pellerin, A. M. de Passillé, and J. Rushen. 2012. Growth modeling of dairy heifers in Québec based on random regression. Canadian Journal of Animal Science 92(1):33-47. https://doi.org/10.4141/cjas2011-083.

Cuttance, E. L., W. A. Mason, R. A. Laven, and C. V. C. Phyn. 2018. The relationship between failure of passive transfer and mortality, farmer-recorded animal health events and body weights of calves from birth until 12 months of age on pasture-based, seasonal calving dairy farms in New Zealand. The Veterinary Journal 236:4-11. https://doi.org/10.1016/j.tvjl.2018.04.005.

Dallago, G. M., R. I. Cue, K. M. Wade, R. Lacroix, and E. Vasseur. 2021a. Birth conditions affect the longevity of Holstein offspring. Journal of Dairy Science 105. https://doi.org/10.3168/jds.2021-20214.

Dallago, G. M., K. M. Wade, R. I. Cue, J. T. McClure, R. Lacroix, D. Pellerin, and E. Vasseur. 2021b. Keeping dairy cows for longer: A critical literature review on dairy cow longevity in high milk-producing countries. Animals 11(3):808. https://doi.org/10.3390/ani11030808.

Delgado, H. A., R. I. Cue, D. Haine, A. Sewalem, R. Lacroix, D. Lefebvre, J. Dubuc, E. Bouchard, and K. M. Wade. 2017. Profitability measures as decision-making tools for Québec dairy herds. Canadian Journal of Animal Science 98(1):18-31. https://doi.org/10.1139/cjas-2016-0202.

DeNise, S. K., J. D. Robison, G. H. Stott, and D. V. Armstrong. 1989. Effects of passive immunity on subsequent production in dairy heifers. Journal of Dairy Science 72(2):552-554. https://doi.org/10.3168/jds.S0022-0302(89)79140-2.

Elsohaby, I., C. B. Riley, S. Hou, J. T. McClure, R. A. Shaw, and G. P. Keefe. 2014. Measurement of serum immunoglobulin G in dairy cattle using Fourier-transform infrared spectroscopy: A reagent free approach. The Veterinary Journal 202(3):510-515. https://doi.org/10.1016/j.tvjl.2014.09.014.

Grandl, F., M. Furger, M. Kreuzer, and M. Zehetmeier. 2019. Impact of longevity on greenhouse gas emissions and profitability of individual dairy cows analysed with different system boundaries. Animal 13(1):198-208. https://doi.org/10.1017/S175173111800112X.

Habel, J., V. Uhlig, S. Hoischen-Tauber, E.-M. Schwabenbauer, T. Rumphorst, L. Ebert, D. Möller, and A. Sundrum. 2021. Income over service life cost – Estimation of individual profitability of dairy cows at time of death reveals farm-specific economic trade-offs. Livestock Science 254:104765. https://doi.org/10.1016/j.livsci.2021.104765.

Hang, B. P. T., J. Dicksved, K. S. Sjaunja, and E. Wredle. 2017. Colostrum quality, IgG absorption and daily weight gain of calves in small-scale dairy production systems in Southern Vietnam.

Tropical Animal Health and Production 49:1143-1147. https://doi.org/10.1007/s11250-017-1308-6.

Higginson Cutler, J. H., J. Rushen, A. M. de Passillé, J. Gibbons, K. Orsel, E. Pajor, H. W. Barkema, L. Solano, D. Pellerin, D. Haley, and E. Vasseur. 2017. Producer estimates of prevalence and perceived importance of lameness in dairy herds with tiestalls, freestalls, and automated milking systems. Journal of Dairy Science 100(12):9871-9880. https://doi.org/10.3168/jds.2017-13008.

Hultgren, J. and C. Svensson. 2010. Calving interval in dairy cows in relation to heifer rearing conditions in southwest Sweden. Reproduction in Domestic Animals 45(1):136-141. https://doi.org/10.1111/j.1439-0531.2008.01273.x.

Karle, B. M., G. U. Maier, W. J. Love, S. A. Dubrovsky, D. R. Williams, R. J. Anderson, A. L. Van Eenennaam, T. W. Lehenbauer, and S. S. Aly. 2019. Regional management practices and prevalence of bovine respiratory disease in California's preweaned dairy calves. Journal of Dairy Science 102(8):7583-7596. https://doi.org/10.3168/jds.2018-14775.

Kim, S. and H. Kim. 2016. A new metric of absolute percentage error for intermittent demand forecasts. International Journal of Forecasting 32(3):669-679. https://doi.org/10.1016/j.ijforecast.2015.12.003.

Kuhn, M. 2014. Futility analysis in the cross-validation of machine learning models. arXiv:1405.6974v1.

Kuhn, M. 2020. caret: Classification and regression training.Version 6.0-86. https://CRAN.R-project.org/package=caret.

Lombard, J., N. Urie, F. Garry, S. Godden, J. Quigley, T. Earleywine, S. McGuirk, D. Moore, M. Branan, M. Chamorro, G. Smith, C. Shivley, D. Catherman, D. Haines, A. J. Heinrichs, R. James, J. Maas, and K. Sterner. 2020. Consensus recommendations on calf- and herd-level passive immunity in dairy calves in the United States. Journal of Dairy Science 103(8):7611-7624. https://doi.org/10.3168/jds.2019-17955.

Lora, I., F. Gottardo, B. Contiero, B. Dall Ava, L. Bonfanti, A. Stefani, and A. Barberio. 2018. Association between passive immunity and health status of dairy calves under 30 days of age. Preventive Veterinary Medicine 152:12-15. https://doi.org/10.1016/j.prevetmed.2018.01.009.

Medrano-Galarza, C., S. J. LeBlanc, A. Jones-Bitton, T. J. DeVries, J. Rushen, A. Marie de Passillé, M. I. Endres, and D. B. Haley. 2018. Associations between management practices and within-pen prevalence of calf diarrhea and respiratory disease on dairy farms using automated milk feeders. Journal of Dairy Science 101(3):2293-2308. https://doi.org/10.3168/jds.2017-13733.

Molnar, C. 2019. Interpretable Machine Learning. Christoph Molnar.

Molnar, C., G. Casalicchio, and B. Bischl. 2018. iml: An R package for interpretable machine learning. Journal of Open Source Software 3(26):786. https://doi.org/10.21105/joss.00786.

Pretty, J. and Z. P. Bharucha. 2019. Sustainable livestock and animal-sourced food. Pages 225-232 in Sustainable food and agriculture: An integrated approach. C. Campanhola and S. Pandey, ed. The Food and Agriculture Organization of the Uniteed Nations & Elsevier Inc. https://doi.org/10.1016/B978-0-12-812134-4.00022-4.

Rushen, J. and A. M. Passillé. 2013. The importance of improving cow longevity. Pages 03-21 in Proc. Cow Longevity Conference. DeLaval, Hamra Farm, Tumba, Sweeden.

Sherwin, V. E., C. D. Hudson, A. Henderson, and M. J. Green. 2016. The association between age at first calving and survival of first lactation heifers within dairy herds. Animal 10(11):1877-1882. https://doi.org/10.1017/S1751731116000689.

Shook, G. E. 2006. Major advances in determining appropriate selection goals. Journal of Dairy Science 89(4):1349-1361. https://doi.org/10.3168/jds.S0022-0302(06)72202-0.

Soberon, F. and M. E. Van Amburgh. 2013. Lactation Biology Symposium: The effect of nutrient intake from milk or milk replacer of preweaned dairy calves on lactation milk yield as adults: A meta-analysis of current data. Journal of Animal Science 91(2):706-712. https://doi.org/10.2527/jas.2012-5834.

Svensson, C. and J. Hultgren. 2008. Associations between housing, management, and morbidity during rearing and subsequent first-lactation milk production of dairy cows in southwest Sweden. Journal of Dairy Science 91(4):1510-1518. https://doi.org/10.3168/jds.2007-0235.

Tang, F. and H. Ishwaran. 2017. Random forest missing data algorithms. Stat Anal Data Min 10(6):363 – 377. https://doi.org/10.1002/sam.11348.

Urie, N. J., J. E. Lombard, C. B. Shivley, C. A. Kopral, A. E. Adams, T. J. Earleywine, J. D. Olson, and F. B. Garry. 2018. Preweaned heifer management on US dairy operations: Part V. Factors associated with morbidity and mortality in preweaned dairy heifer calves. Journal of Dairy Science 101(10):9229-9244. https://doi.org/10.3168/jds.2017-14019.

Vacek, M., L. Krpálková, J. Syrůček, M. Štípková, and M. Janecká. 2015. Relationships between growth and body condition development during the rearing period and performance in the first three lactations in Holstein cows. Czech Journal of Animal Science 60(9):417-425. https://doi.org/10.17221/8460-CJAS.

van Buuren, S. 2019. Flexible imputation of missing data. 2 ed. Chapman & Hall/CRC, New York. https://doi.org/10.1201/9780429492259.

van Buuren, S. and K. Groothuis-Oudshoorn. 2011. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software 45(3):1 - 67. https://doi.org/10.18637/jss.v045.i03.

Vos, R. and L. G. Bellù. 2019. Global trends and challenges to food and agriculture into the 21st century. Pages 11-30 in Sustainable food and agriculture: An integrated approach. C. Campanhola and S. Pandey, ed. The Food and Agriculture Organization of the Uniteed Nations & Elsevier Inc. https://doi.org/10.1016/B978-0-12-812134-4.00002-9.

Windeyer, M. C., K. E. Leslie, S. M. Godden, D. C. Hodgins, K. D. Lissemore, and S. J. LeBlanc. 2014. Factors associated with morbidity, mortality, and growth of dairy heifer calves up to 3 months of age. Preventive Veterinary Medicine 113(2):231-240. https://doi.org/10.1016/j.prevetmed.2013.10.019.

4.8. Supplementary material

Supplementary Table 4.1. Prediction performance of recursive partitioning and regression tree (RPART), gradient boosting machine (GBM), random forest (RF), and support vector machine (SVM) with a radial basis kernel models on each of the outcomes evaluated. Performance is shown both on training and validation data sets. Means ± standard deviations were calculated across all 10 multiple imputed versions of the data. Best results within outcome, data set, and metric are bolded.

Outcome[1]	Model	Training[2]				Validation[2]			
		R^2	RMSE	MAE	SMAPE	R^2	RMSE	MAE	MAAPE
LL	RPART	0.05 ± 0.00	1.59 ± 0.00	1.33 ± 0.00	38.7 ± 0.00	0.02 ± 0.00	1.63 ± 0.00	1.38 ± 0.00	40.0 ± 0.00
	GBM	0.18 ± 0.07	1.47 ± 0.07	1.23 ± 0.06	36.8 ± 1.24	0.04 ± 0.02	1.61 ± 0.02	1.35 ± 0.02	39.4 ± 0.40
	RF	**0.39 ± 0.09**	**1.27 ± 0.09**	**1.06 ± 0.08**	**33.1 ± 1.65**	**0.06 ± 0.02**	**1.60 ± 0.01**	**1.34 ± 0.01**	**39.2 ± 0.26**
	SVM	0.11 ± 0.06	1.53 ± 0.05	1.25 ± 0.06	37.1 ± 1.6	0.02 ± 0.03	1.63 ± 0.03	1.35 ± 0.03	39.5 ± 0.66
LPL	RPART	0.05 ± 0.00	1.19 ± 0.00	1.00 ± 0.00	44.8 ± 0.00	0.059 ± 0.00	1.22 ± 0.00	**1.00 ± 0.00**	**45.2 ± 0.00**
	GBM	0.16 ± 0.01	1.12 ± 0.01	0.93 ± 0.01	43.0 ± 0.31	0.005 ± 0.02	1.25 ± 0.01	1.02 ± 0.01	45.7 ± 0.53
	RF	**0.30 ± 0.004**	**1.03 ± 0.003**	**0.87 ± 0.003**	**41.4 ± 0.11**	**0.060 ± 0.01**	**1.21 ± 0.004**	1.02 ± 0.003	45.9 ± 0.11
	SVM	0.10 ± 0.003	1.16 ± 0.002	0.97 ± 0.001	44.0 ± 0.07	0.037 ± 0.002	1.23 ± 0.002	1.02 ± 0.003	46.0 ± 0.17
ECM	RPART	0.06 ± 0.00	15,445.1 ± 0.00	12,876.4 ± 0.00	49.8 ± 0.00	0.03 ± 0.00	15,907.2 ± 0.00	12,960.7 ± 0.00	**48.6 ± 0.00**
	GBM	0.19 ± 0.01	14,295.4 ± 49.03	11,911.7 ± 61.55	47.9 ± 0.2	0.05 ± 0.01	15,746.2 ± 114.52	**12,850.1 ± 99.34**	48.8 ± 0.29
	RF	**0.32 ± 0.004**	**13,091.3 ± 36.39**	**10,867.8 ± 37.12**	**45.5 ± 0.11**	0.08 ± 0.01	**15,493.9 ± 48.69**	12,926.1 ± 45.41	49.3 ± 0.15
	SVM	0.14 ± 0.003	14,779.4 ± 21.68	12,086.0 ± 23.37	48.0 ± 0.1	**0.04 ± 0.004**	15,862.0 ± 35.61	13,236.8 ± 33.94	50.0 ± 0.14
MV	RPART	0.06 ± 0.00	10,964.0 ± 0.00	8,989.3 ± 0.00	49.1 ± 0.00	0.04 ± 0.00	11,483.8 ± 0.00	9,855.3 ± 0.00	53.7 ± 0.00
	GBM	0.26 ± 0.10	9,678.7 ± 694.56	7,895.9 ± 614.51	45.7 ± 2.22	0.06 ± 0.03	11,426.8 ± 177.02	9,724.4 ± 167.40	52.6 ± 1.14
	RF	**0.31 ± 0.003**	**9,360.8 ± 21.22**	**7,693.9 ± 17.48**	45.4 ± 0.07	**0.08 ± 0.01**	**11,257.3 ± 69.88**	9,704.1 ± 71.43	53.3 ± 0.30

SVM	0.20 ± 0.04	10,123.8 ± 238.33	8,024.8 ± 283.37	**45.2 ± 1.1**	0.07 ± 0.02	11,297.5 ± 91.38	**9,620.3 ± 76.18**	**52.4 ± 0.35**

[1] LL = Length of life (year); LPL = Length of productive life (year); ECM = Lifetime cumulative energy-corrected milk (kg); MV = Lifetime cumulative milk value ($CAD).
[2] R^2 = Coefficient of determination; RMSE = Root mean squared error; MAE = Mean absolute error; MAAPE = Mean arctangent absolute percentage error (Kim and Kim, 2016).

Connecting statement 3

Animal outcomes were found to be linked with animal subsequent longevity, productivity, and profitability. A strong herd effect was also found, which highlighted the importance of farmers as decision-making agents. By choosing which practices to use, farmers are also responsible for deciding how to raise dairy animal. Early life management practices adopted during the rearing period build up to the conditions to which the calves are exposed, which could not only influence early life animal outcomes, but also be reflected in lifetime indicators of longevity, productivity, and profitability. The study described in the next chapter was carried out to explore this potential influence focusing on the management practices adopted by Quebec dairy herds and herd-level indicators of longevity, productivity, and profitability were analysed.

CHAPTER 5 – Early life management practices and their association with herd longevity, productivity, and profitability

Gabriel M. Dallago*, and Elsa Vasseur

Department of Animal Science, McGill University, Sainte-Anne-de-Bellevue, Quebec, H9X 3V9, Canada

* Corresponding author: gabriel.dallago@mail.mcgill.ca

Manuscript in preparation to be submitted to the Journal of Dairy Science

5.1. Abstract

Improving the management of lactating cows to reduce the occurrence of health issues could result in increased cow longevity. Though a similar rationale could be extrapolated to calves, the potential long-term effect of early life management practices remains unknown. The objectives of this study were to characterize dairy farms based on their early life management practices and analyze the association between management practices and herd longevity, productivity, and profitability. Early life management practices regarding colostrum feeding (11 questions), milk feeding (9 questions), solid feed and weaning (13 questions), and housing (12 questions) were collected in Quebec (Canada) dairy farms using a questionnaire. The Dairy Herd Improvement (**DHI**) metrics of herd average length of productive life (**LPL**) and percentage of cows on third or greater lactation (**3+ lactation**) were used as longevity indicators while herd lifetime cumulative energy-corrected milk (**ECM**) and lifetime cumulative milk value were used as indicators of productivity and profitability, respectively. After handling and cleaning, there remained answers from 2,004 Quebec dairy farms obtained between February 2020 and February 2021 (**WD1**). The WD1 was merged with the DHI outcomes and contained data from 1,658 farms (**WD2**). Cluster analysis was performed using the WD1 to characterize farms based on their early life management practices. Cluster stability assessment was used to determine both the best clustering algorithm and number of clusters. Different machine learning algorithms were then trained using the WD2 to evaluate the associations between DHI outcomes and early life management practices that best described the clusters. Multiple imputation was used to handle missing data and an 80:20 split ratio was used to create the training and validation data sets, respectively. Models were trained with the training data sets and using 5-fold cross-validation repeated 10 times. Two distinct clusters were identified: production-oriented farms (N = 1,285; 64%) were characterized by using powdered

milk replacer, either acidified or non-acidified, and to automatically feed calves housed in groups both before and after weaning. Longevity-oriented farms (N = 719; 36%) were characterized by not measuring colostrum Immunoglobulin G and using non-pasteurized or non-acidified milk (whole or waste) to feed their calves. Both colostrum and milk were fed using buckets without teats, and weaned calves were housed individually tied-up in stalls. In addition, practices adopted by production-oriented farms were associated with reduced longevity but increased productivity and profitability, whereas practices adopted by longevity-oriented farms were associated with increased longevity but reduced productivity and profitability. Our results highlighted that early life management practices are linked with herd outcomes of longevity, profitability, and productivity, but adopting the best practices is not necessarily associated with better herd outcomes.

5.2. Introduction

Cow longevity is a central component of dairy farming sustainability (Dallago et al., 2021). Higher longevity indicates a high proportion of mature cows in more profitable lactations, diluting the rearing costs through a longer period (Essl, 1998). Mature cows also produce more milk, which in turn reduces the environmental footprint of dairy farms (Benbrook et al., 2010, Grandl et al., 2019). Though cow longevity is frequently determined by the farmers, since they are responsible for deciding when to cull a cow, such a decision is a consequence of underlying cow factors and of the farmer's management style. The occurrence of diseases such as mastitis and lameness impact both animal productivity and profitability (Puerto et al., 2021a and 2021b), and they are important animal factors associated with culling decisions (Pinedo et al., 2010, CDIC, 2021a). However, heath issues can be treated, and the decision process between culling or treating is often not objective (Adriaens et al., 2020) and dependents upon the farmer's production priorities (Rilanto et al., 2022). Therefore, cow longevity is a functional trait which has the potential to be maximized to improve the sustainability of dairy farming.

Improving management practices to reduce the occurrence of health problems could be a possibility to improve cow longevity. The associations between management practices and the occurrence of health issues have been extensively studied. The occurrence of mastitis, for example, is associated with hygiene practices. Increasing the cleaning frequency in dairy barns improves the hygiene scores for the udder (DeVries et al., 2012), contributing to a decrease in somatic cell counts (Schreiner and Ruegg, 2003, Reneau et al., 2005) and a reduction of clinical mastitis (Santman-Berends et al., 2016). On the other hand, concrete floors (Somers et al., 2003, Vanegas et al., 2006) and bed comfort (McPherson and Vasseur, 2020) are some of the risk factors associated with the occurrence of lameness.

Since improving the management of lactating cows could reduce the occurrence of health issues and potentially increase cow longevity, it would not be unreasonable to argue that early life management practices used in raising calves could have a similar ramification. However, associations between early life management practices and animal health and performance are often analyzed in the short-term. For instance, calves receiving 2 colostrum meals have a greater average daily gain (**ADG**) before weaning and are less likely to have failure of transfer of passive immunity (**FTP**) and to be treated for diseases compared to calves receiving 1 colostrum meal (Abuelo et al., 2021). A few studies indicate the existence of long-term effects of some rearing factors on cow longevity such as the age when calves eat a high amount of concentrate feeding (Heinrichs and Heinrichs, 2011), the type of heifer housing (Hultgren and Svensson, 2009), and the age and body condition score at first breeding (Hultgren et al., 2011). The long-term link between a broader range of early life management practices, such as colostrum feeding and FTP, and adult cow performance has not been demonstrated.

Therefore, the objectives of this study were to characterize dairy farms based on their early life management practices and to analyze the association between management practices and herd longevity, productivity, and profitability. We hypothesized that the adoption of early life management practices varies between herds and that these practices are associated with farm outcomes.

5.3. Material and methods

A cross-sectional study was conducted using data on early life management practices as well as Dairy Herd Improvement (**DHI**) indicators of longevity, productivity, and profitability from Quebec (Canada) dairy farms. All statistical analyses were carried out using the R statistical software, version 4.1.1 (https://www.r-project.org/). The R code that supports the findings of this study can be found in the following public GitHub repository: https://github.com/CowLifeMcGill/early_pract_longevity.

5.3.1. Datasets

Early life management practices were retrospectively collected at farm level using a questionnaire. The questionnaire (Annex 5.1) comprised 45 questions and sub-questions divided into 4 groups of management practices: colostrum feeding (11 questions), milk feeding (9 questions), solid feed and weaning (13 questions), and housing (12 questions). Field technicians from Lactanet (Sainte-Anne-de-Bellevue, Quebec, Canada), the Canadian DHI agency, conducted the survey over phone calls on 2,087 dairy farms from Quebec (Canada), between February 2020 and February 2021. For each question, farmers were asked to consider the majority of the female calves (75%) rather than exceptions. The answers (data) were qualitative nominal (e.g., multiple choice), qualitative ordinal (e.g., scale of answers from 1 = less than 50 days to 4 = more than 90 days), or numeric (e.g., liters of milk fed per day).

The DHI indicators of milk production (kg), fat production (kg), protein production (kg), and milk value ($CAD) as well as the longevity metrics of length of productive life (**LPL**; years) and percentage of animals on third or greater lactations (**3+ lactation**) were provided by Lactanet. Productions and milk value were provided for 209,749 animals from 3,242 farms while LPL and 3+ lactation were provided for 64,041 animals from 3,242 farms and 32,937 animals from 3,240 farms, respectively. All DHI indicators were extracted for a 12-month period prior to the date when an on-farm welfare assessment was conducted (between December 2015 and December 2019) as part of the Animal Care module of the proAction® Quality Assurance Program (DFC, 2021).

5.3.2. Data handling and cleaning

Questionnaires from organic farms (N = 70) and questions describing how the management of male calves differed from female (N = 2) were excluded. The content of fat (%) and protein (%) of milk replacer (Questions 7.1.a and 7.1.b., respectively) had a high percentage of missing values, since not all farms use powdered milk replacer. The value zero was imputed to farms that do not use powdered milk replacer (N = 723). In addition, farms that do not use powdered milk replacer were not differentiated on the binary (i.e., yes or no) question regarding milk replacer being medicated (Questions 7.1.c). Therefore, the new level “No milk replacer” was used to indicate such differentiation. Since frequency for adding bedding was answered on a daily, weekly, or monthly basis, all answers were standardized to a monthly frequency. One farm was excluded since they answered that no colostrum was offered to calves. Identified erroneous answers for how much colostrum was offered in the first meal (Question 3.4.1; N = 2) were replaced by missing data.

The following additional changes were made because the frequency of some responses was low. The answers “From 7 to 12 hours” (N = 84) and “More than 12 hours” (N = 7) for the question “How many hours after birth does the calf receive it’s first meal?” (Question 2) were combined into the new level “More than 7 hours”. Next, the milk feeding systems (Question 5) “Group bucket with teats” (N = 46), “Floating teat” (N = 10), “Group bucket” (N = 6), and “Nursing cow” (N = 4) were combined to create the level “Other”. Calf milk state (Question 7) “Non-pasteurized / Acidified” (N = 8) and “Pasteurized / Acidified” (N = 1) were combined into the new level “Others”. Lastly, weaning ages (Question 12) “From 71 to 90 days” (N = 220) and “More than 90 days” (N = 43) were grouped into the new level “More than 70”.

Herd average 3+lactation and LPL as well as cumulative milk production and milk value were calculated based on animal level data. Animals with negative LPL were removed (N = 33). Next, animals with either missing (N = 12,515) observations on production (milk, fat, and protein) and milk value indicators were excluded. For animals that moved between herds, both productions and milk value were associated with the herd in which the cow ended each of her lactations for the calculation of the cumulative milk, fat, protein, and milk value. Cumulative energy-corrected milk (**ECM**; kg) was then calculated as ECM (lifetime cumulative kg) = 12.55 × fat (lifetime cumulative kg) + 7.39 × protein (lifetime cumulative kg) + 0.2595 × milk yield (lifetime cumulative kg). Herd averages were calculated for 3+lactation, LPL, lifetime cumulative ECM, and lifetime cumulative milk value.

5.3.3. Working data sets

Two working data sets were used for statistical analysis. The working data 1 (**WD1**) was composed of the questionnaire on early life management practices. After handling and cleaning, there remained data from 2,004 Quebec dairy farms that answered the survey between February 2020 and February 2021. The percentage of missing data (i.e., questions without answers) ranged from 0.1% (N = 2) on when calves are removed after birth to 16.0% (N = 320) on age when fermented feed is fist offered. No further steps were necessary to handle the missing data in the WD1 as described in the data analyses section.

The working data 2 (**WD2**) was created by merging WD1 with DHI outcomes based on herd ID. After the merge, data from farms with the level "Other" for both colostrum feeding system (Question 3; N = 6) and colostrum source (Question 3.2; N = 3) questions as well as "Frozen pasteurized" (N = 3) for colostrum state (Question 3.2.1) were removed due to their low frequency. The final WD2 contained data from 1,658 farms with a percentage of missing data (i.e., questions without answers) ranging from 0.1% (N = 1) on when calves are removed after birth to 15.4% (N = 255) on age when fermented feed is fist offered. Multiple imputation was used to handle missing data in the WD2 instead of case wise deletion to maximize the number of observations. The function *mice* from the package *mice* (van Buuren and Groothuis-Oudshoorn, 2011) was used to create 10 multiple imputed versions of the data using the random forest method. In short, for each early life management practice (i.e., either numeric or categorical) with missing data, the remaining practices were used to create a random forest model to impute the missing values. The distribution of the imputed data was visually compared to the data with non-missing values. In addition to multiple imputation being a better approach compared to other methodologies in order to increase power and accuracy of the data analysis(van Buuren, 2019), random forest is able to handle complex interactions between variables even in conditions where there is a high number of missing data (Tang and Ishwaran, 2017).

5.3.4. Data analyzes

5.3.4.1. Cluster analysis

Cluster analysis was performed to characterize farms based on their early life management practices using the WD1. The Gower distance (Gower, 1971) was used to calculate the dissimilarity matrix since there were both qualitative and numerical type of answers. Though there were missing data on WD1, the function *daisy* from the R package cluster (Maechler et al., 2021), which was used to calculate the dissimilarity matrix, automatically handled missing data by not including pairwise missing values in the calculation of the dissimilarities matrix. Clustering algorithms and the number of clusters were both defined based on cluster quality assessment. Cluster quality was evaluated based on the stability of the hierarchical clustering with Ward's minimum variance linkage, partitioning around medoids (**PAM**), and normal mixture model-based clustering algorithms. Cluster stability was assessed based on bootstrapped Jaccard mean distance obtained after 100 resamples and with cluster number varying from two to seven (Supplementary

Table 5.1). This was done using the function *clusterboot* from the R package *fpc* (Hennig, 2020). The cluster algorithm and number of clusters with the highest overall stability were used for further analysis.

Clusters were described following the methodology described by Husson et al. (2017) and using the function *catdes* from the R package *FactoMineR* (Sebastien Le et al., 2008). In short, we used the v-test as well as chi-squared test and one-way ANOVA to describe the clusters. The v-test provides a descriptive measure of deviation between practices in the clusters compared to the overall distribution, in which the higher the value, the greater the deviation (Husson et al., 2017). Therefore, the v-test was used to identify and rank the top 10 qualitative and numerical early life management practices that were statistically significant according to the chi-squared test ($\alpha < 0.05$) and the one-way ANOVA ($\alpha < 0.05$) respectively, to describe each cluster separately. These were also further used to analyze the association between early life management practices and herd outcomes based on machine learning algorithms.

5.3.4.2. Machine learning

The associations between early life management practices and herd longevity (i.e., LPL, 3+Lactation), lifetime cumulative ECM, and lifetime cumulative milk value were analyzed using the machine-learning algorithms recursive partitioning and regression tree (**RPART**), gradient boosting machine (**GBM**), random forest (**RF**), and support vector machine (**SVM**) with a radial basis kernel. For each of the imputed versions of the data, an 80:20 ratio was used to create the training and validation data sets, respectively, based on the herd outcome (i.e., LPL, 3+Lact, lifetime cumulative ECM, or lifetime cumulative milk value) distribution. All models were trained with 5-fold cross-validation repeated 10 times on the training data sets using the *caret* R package (Kuhn, 2020) by specifying the methods *rpart*, *gbm*, *ranger*, and *svmRadialSigma*, respectively for the algorithms RPART, GBM, RF, and SVM. Hyperparameters for these models were tuned using adaptive resampling, which resample the hyperparameter tuning grid by concentrating on values closer to the identified optimal setting (Kuhn, 2014, 2020). The models were evaluated based on the coefficient of determination (**R^2**), root mean squared error (**RMSE**), mean absolute error (**MAE**), and mean arctangent absolute percentage error (**MAAPE**; Kim and Kim, 2016) calculated using the validation data sets. The best model was defined as having the highest R^2 and the lowest RMSE, MAE, and MAAPE averages across all the imputed versions of the data.

Using the best model, variable importance was estimated while accumulative local effect (**ALE**) was used to describe the association between the outcomes and the top 5 most important early life management practices. Variable importance was calculated using permutation. In short, this is a model agnostic approach that measures the prediction error of the model after shuffling the values of the early life indicators, which changes the relationship between the outcome and early life indicators. Shuffling the values of important early life indicators would result in an increase of the error while the error would remain unchanged for early life indicators that are not important (Molnar, 2019). On the other hand, ALE indicates the average influence of the early life indicators in predicting the outcome variables (Molnar, 2019). Variable importance and ALE were estimated using the functions *FeatureImp* and *FeatureEffect* respectively, from the R package *iml* (Molnar et al., 2018) and based on each of the complete imputed data sets (i.e., training and validation data set combined). Average and standard deviation were calculated to aggregate the results across all imputed version of the data.

5.4. Results

5.4.1. Early life management practice clusters

Farms were grouped into two stable clusters based on their adoption of early life management practices using the normal mixture model-based clustering algorithm (Figure 5.1). Farms in cluster 1 (**production-oriented**; N = 1,285 farms), composed of 64% of the farms, were characterized as adopting more modern early life practices. Farms in this cluster used powdered milk replacer, either non-acidified or acidified, through an automatic feeding system to feed their calves, which in turn were housed in groups both before and after weaning (Table 5.1). On the other hand, farms in cluster 2 (**longevity-oriented**; N = 719 farms) were mainly characterized by adopting more traditional practices. They used non-pasteurized or non-acidified milk (whole or waste) to feed their calves, both colostrum and milk were fed using buckets without teats, they did not measure the concentration of Immunoglobulin G in the colostrum, and weaned calves were housed individually tied-up in stalls (Table 5.2). Overall, farms in cluster 2 fed calves less and offered both water and concentrate feeding later than farms in cluster 1 (Table 5.3).

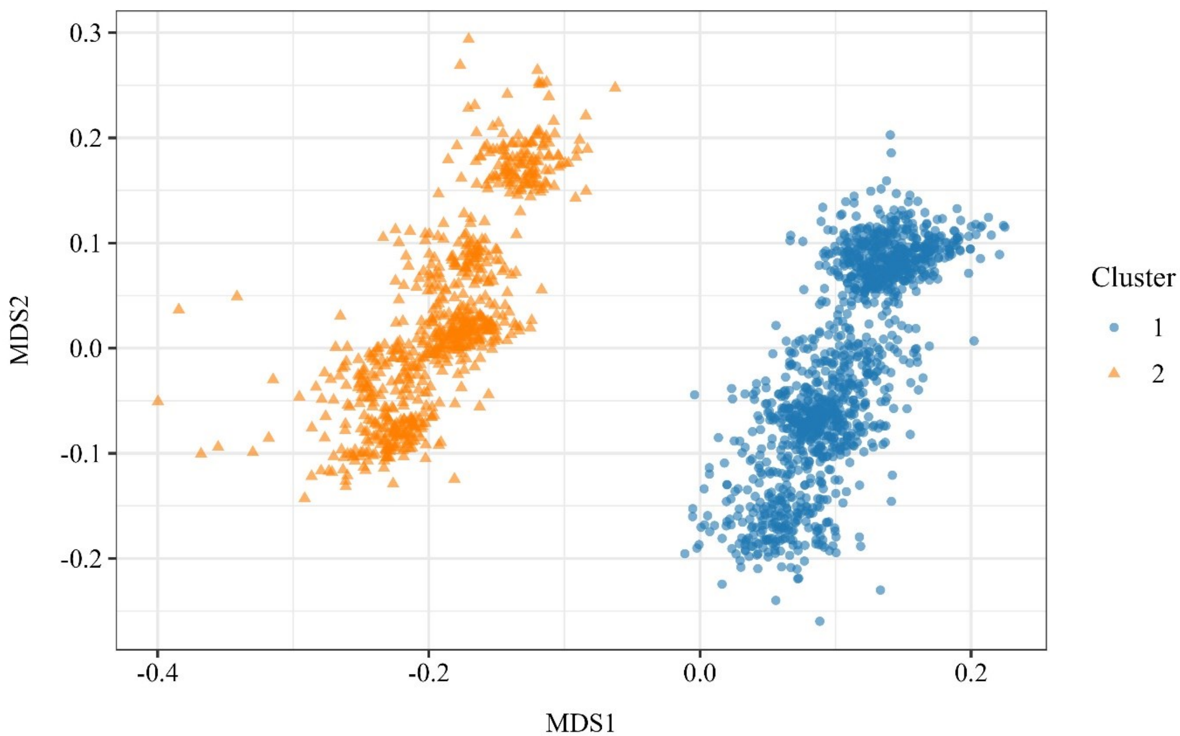


Figure 5.1. Multidimensional scaling (MDS) visualization of clusters based on early life management practices from 2,017 Quebec – Canada dairy herds. Clusters were created using mixed normal model-based algorithm since it produced more stable results. Herds are represented by the points and clusters are represented by a combination of color and point shape. Cluster 1 (production-oriented) = 1,285 herds, Cluster 2 (longevity-oriented) = 719 herds.

Table 5.1. Top qualitative early life management practices that best describe the production-oriented Quebec – Canada dairy farms on cluster 1 (N = 1,285). Practices were selected based on the v-test and chi-squared test (α < 0.05).

Early life management practice	Overall[1]	V-test	Cla/Mod[2]	Mod/Cla[3]	P-value[4]
Medicated milk replacer = No	59.2	+Inf	100.0	92.4	< 0.001
Milk state = Non-acidified	51.4	+Inf	100.0	80.2	< 0.001
Milk source = Powdered milk replacer	64.1	+Inf	100.0	100	< 0.001
Milk state = Partially acidified	7.6	11.7	100.0	11.8	< 0.001
Milk feeding system = Automatic	11.7	11.6	94.5	17.3	< 0.001
Non-weaned housing = Group	39.0	9.4	76.6	46.5	< 0.001
Milk state = Acidified	5.0	9.3	100.0	7.8	< 0.001
Medicated milk replacer = Yes	4.9	9.2	100.0	7.6	< 0.001
Non-weaned housing detail = Pen	38.2	8.8	76.0	45.3	< 0.001
Weaned housing = Group	67.2	6.8	69.3	72.6	< 0.001

[1] Proportion of farms adopting each one of the listed early life management practices based on all 2,004 farms.
[2] Proportion of farms adopting each one of the listed early life management practices in cluster 1 out of the overall number of farms that adopt the same practice (e.g., 100.0% out of 2,004 × 34.3% ≈ 1,186 farms that feed calves with non-medicated milk replacer are in cluster 1).
[3] Proportion of farms in Cluster 1 adopting each one of the listed early life management practices out of the number of farms in cluster 1 (e.g., 92.4% out of 1,285 farms in cluster 1 feed calves with no-medicated milk replacer).
[4] Overall vs cluster (α < 0.05).

Table 5.2. Top qualitative early life management practices that best describe the longevity-oriented Quebec – Canada dairy farms on cluster 2 (N = 719). Practices were selected based on the v-test and chi-squared test (α < 0.05).

Early life management practice	Overall[1]	V-test	Cla/Mod[2]	Mod/Cla[3]	P-value[4]
Medicated milk replacer = No milk replacer	35.9	+Inf	100.0	100.0	< 0.001
Milk state = Non-pasteurized / non-acidified	34.2	+Inf	100.0	95.3	< 0.001
Milk source = Whole milk	33.1	+Inf	100.0	92.2	< 0.001
Milk source = Waste milk	2.8	10.6	100.0	7.8	< 0.001
Milk feeding system = Individual buckets without teats	23.9	9.5	54.3	36.2	< 0.001
Weaned housing = Individual tied-up	21.6	7.0	50.3	30.3	< 0.001
Milk state = Pasteurized / Non-acidified	1.3	6.6	96.3	3.6	< 0.001
Weaned housing detail = Stall	19.0	6.2	49.9	26.4	< 0.001
Colostrum IgG assessment = Not evaluated	69.0	6.1	40.2	77.3	< 0.001
Colostrum feeding system = Buckets without teats	8.7	5.9	56.9	13.8	< 0.001

[1] Proportion of farms adopting each one of the listed early life management practices based on all 2,017 farms.
[2] Proportion of farms adopting each one of the listed early life management practices in cluster 2 out of the overall number of farms that adopt the same practice (e.g., 100% out of 719 × 35.9% ≈ 258 farms that do not feed calves with powdered milk replacer are in cluster 2).
[3] Proportion of farms in Cluster 2 adopting each one of the listed early life management practices out of the number of farms in cluster 2 (e.g., 100%% out of 719 farms in cluster 2 do not feed calves with powdered milk replacer).
[4] Overall vs cluster (α < 0.05).

Table 5.3. Top numeric early life management practices that best describe Quebec – Canada dairy farms on cluster 1 (N = 1,285) and cluster 2 (N = 719). Practices were selected based on the v-test and one-way ANOVA (α < 0.05).

Early life management practice	Overall[1]	Cluster 1 Production-oriented			Cluster 2 Longevity-oriented		
		v-test	Mean ± SD[2]	p-value[3]	v-test	Mean ± SD[2]	p-value[3]
Milk replacer protein (%)	15.2 ± 11.84	43.1	23.8 ± 4.01	< 0.001	-43.1	0.0 ± 0.00	< 0.001
Milk replacer fat (%)	12.0 ± 9.49	42.2	18.6 ± 3.97	< 0.001	-42.2	0.0 ± 0.00	< 0.001
Daily feeding (L)	8.4 ± 2.28	9.4	8.8 ± 2.33	< 0.001	-9.4	7.8 ± 2.05	< 0.001
Starter protein (%)	20.6 ± 2.10	2.2	20.6 ± 1.90	0.03	-2.2	20.4 ± 2.4	0.03
Age water is first offered (day)	7.4 ± 10.47	-3.9	6.7 ± 9.61	< 0.001	3.9	8.6 ± 11.77	< 0.001
Frequency of adding bedding (month)	33.2 ± 22.42	-4.3	31.5 ± 22.41	< 0.001	4.3	36.1 ± 22.15	< 0.001
Age concentrate feeding is first offered (day)	9.2 ± 9.73	-5.3	8.4 ± 9.03	< 0.001	5.3	10.8 ± 10.68	< 0.001

[1] Mean ± standard deviation calculated based on all farms.
[2] Mean ± standard deviation calculated based on farms in this clusters.
[3] Overall vs cluster (α < 0.05).

5.4.2. Longevity, productivity, and profitability

Descriptive statistics for the herd DHI outcomes evaluated in this study are presented in Table 5.4. Overall, the prediction errors were within acceptable levels. The highest MAAPE was observed for LPL (17.5%) and the lowest was observed for both ECM (10.5%)and milk value (10.5%) in the validation data set (Supplementary Table 5.2). The LPL was best predicted by random forest whereas GBM models had the best performance carrying out prediction of 3+lactation, cumulative ECM, and cumulative milk value (Supplementary Table 5.2). Therefore, these models were used to assess the associations between early life management practices and herd longevity, productivity, and profitability.

Table 5.4. Descriptive statistics of longevity, production, and profitability considered in this study from 1,658 Quebec – Canada dairy farms.

Herd outcome	Mean	Standard deviation	Minimum	Maximum
Length of productive life (years)	3.3	0.76	1.5	7.9
3+ lactation (%)[1]	41.5	8.12	9.8	75.6
ECM (kg)[2]	11,300.4	1,574.72	4,244.0	19,010.5
Milk value ($CAD)[3]	7,916.4	1,148.48	1,393.2	15,212.0

[1] Herd average percentage of cows on third or greater lactations.
[2] Herd average animal lifetime cumulative energy-corrected milk.
[3] Herd average animal lifetime cumulative milk value.

The importance of the early life management practices in predicting herd LPL, 3+lactation, lifetime cumulative ECM, and lifetime cumulative milk value is depicted in Figure 5.2. All practices had a small importance on average. However, calf milk feeding system was the most important practice in predicting both longevity outcomes, while details about housing of calves before weaning and age when concentrate feed was first offered were the most important practices associated with both lifetime cumulative ECM and lifetime cumulative milk value.

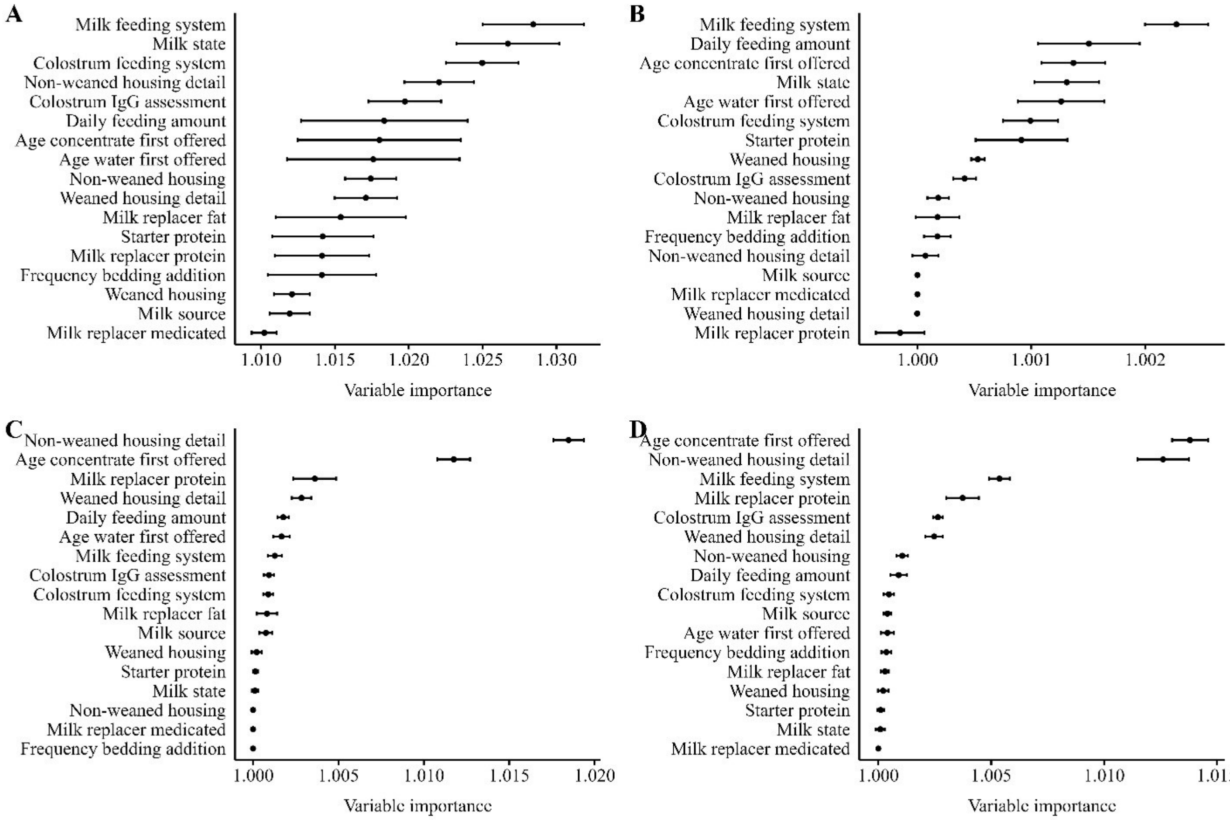


Figure 5.2. The importance (x axis) of indicator variables (y axis; from highest to lowest) to predict herd length of productive life (LPL; **A**), percentage of cows on third or greater lactation (3+lactation; **B**), energy-corrected milk (ECM; **C**), and milk value (**D**). Mean (•) and standard deviation (bar) were calculated across all 10 multiple imputed versions of the data. Variable importance indicates the increase in model error prediction. It is measured as the ratio of root mean squared error of the original model and after shuffling the values of the indicators (Molnar, 2019).

The ALEs between the top important qualitative early life management practices and the herd outcomes evaluated in this study are shown in Table 5.5. Feeding colostrum with a bottle or through an esophageal tube were both associated with decreased LL. Milk feeding calves using individual bucket with teats was associated with high both LPL and 3+ lactation while free feeding in a feed line was associated with high lifetime cumulative milk value. Regarding the housing of non-weaned calves, using hutch was the only practice associated with high lifetime cumulative ECM and lifetime cumulative milk value, whereas pen was the only housing associated with a reduction on LL. On the other hand, housing weaned calves grouped in a pen was the only post-weaning housing type associated with a high lifetime cumulative ECM.

Table 5.5. Accumulated local effect (ALE) of early life management practices on length of productive life (LPL), percentage of cows on third or greater lactation (3+ lactation), lifetime cumulative energy-corrected milk (ECM), and milk value. Means ± standard deviations were calculated across all 10 multiple imputed versions of the data. Ranking columns indicates how important each practice was in predicting the outcome.

Early life management practice	**LPL (year)**		**3+ Lactation (%)**		**ECM (kg)**		**Milk value ($CAD)**	
	Rank	**ALE**	**Rank**	**ALE**	**Rank**	**ALE**	**Rank**	**ALE**
Milk feeding system	1º		1º				3º	
Automatic		-0.03 ± 0.002		-0.27 ± 0.053				-11.7 ± 1.01
Bottle		-0.004 ± 0.002		-0.14 ± 0.109				-11.7 ± 1.01
Feed line (free feed)		-0.04 ± 0.004		-0.27 ± 0.053				295.6 ± 20.23
Individual bucket with teats		0.02 ± 0.001		0.43 ± 0.037				-11.7 ± 1.01
Individual bucket without teats		-0.01 ± 0.001		-0.52 ± 0.064				-11.7 ± 1.01
other		0.003 ± 0.007		-0.27 ± 0.053				-69.5 ± 1.33
Milk state	2º		4º					
Acidified		-0.04 ± 0.005		-0.1 ± 0.028				
Non-acidified		-0.03 ± 0.005		0.0001 ± 0.028				
Non-pasteurized / non-acidified		0.06 ± 0.009		-0.1 ± 0.028				
Other		-0.06 ± 0.008		-0.1 ± 0.028				
Partially acidified		-0.04 ± 0.006		0.54 ± 0.119				
Pasteurized / non-acidified		-0.06 ± 0.011		-0.1 ± 0.028				
Colostrum feeding system	3º							
Bucket with teats		0.01 ± 0.001						
Bucket without teats		0.02 ± 0.002						
Esophageal tube		-0.04 ± 0.004						
Feeding on mother		0.05 ± 0.003						
With a bottle		-0.004 ± 0.0005						
Non-weaned housing detail	4º				1º		2º	
Box		0.02 ± 0.003				-9.9 ± 3.46		-9.9 ± 2.83
Hutch		0.03 ± 0.003				419.1 ± 23.58		272.9 ± 26.94
Pen		-0.04 ± 0.004				-9.9 ± 3.46		-9.9 ± 2.83
Stall		0.03 ± 0.005				-16.1 ± 17.9		-73.5 ± 2.45
Wall		0.02 ± 0.004				-398.1 ± 20.73		-174.6 ± 13.29

Colostrum IgG assessment	5º				5º	
Colostrometer		-0.006 ± 0.002				44.0 ± 6.27
Not evaluated		0.002 ± 0.001				-25.9 ± 2.56
Other		-0.075 ± 0.010				44.0 ± 6.27
Refractometer		-0.008 ± 0.002				76.8 ± 5.94
Visual evaluation		0.005 ± 0.002				44.0 ± 6.27
Weaned housing detail			4º			
Box				-63.4 ± 9.70		
Hutch				-63.4 ± 9.70		
Pen				33.8 ± 5.17		
Stall				-63.4 ± 9.70		
Wall				-63.4 ± 9.70		

Small ALEs were also found between the quantitative early life management practices of high importance and the herd outcomes evaluated in this study. Daily feeding calves with a low milk volume as well as increasing the age when both concentrate feed and water were first offered to calves were associated with a higher 3+ lactation (Figure 5.3). On the other hand, daily feeding calves with a high milk volume was found to be associated with high lifetime cumulative ECM (Figure 5.4). Offering concentrate feed at a young age and using milk replacer with high protein percentage were both associated with high lifetime cumulative ECM (Figure 5.4) and lifetime cumulative milk value (Figure 5.5).

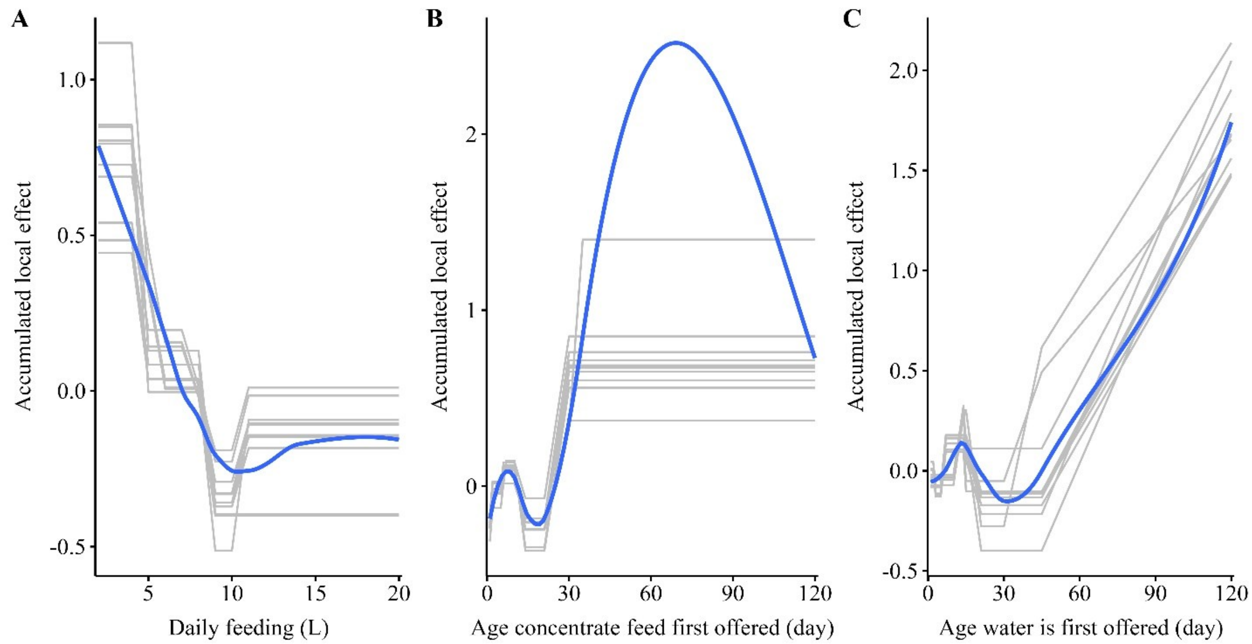


Figure 5.3. The accumulative local effect (ALE) of the early life management practices (**A**) daily feeding amount, (**B**) age when concentrate feed is first offered, and (**C**) age when water is first offered on predicting the herd percentage cows on third or greater lactation (3+lactation) across all 10 multiple imputed versions of the data (grey lines) followed by a loess trend line (blue). Practices shown were selected based on their importance.

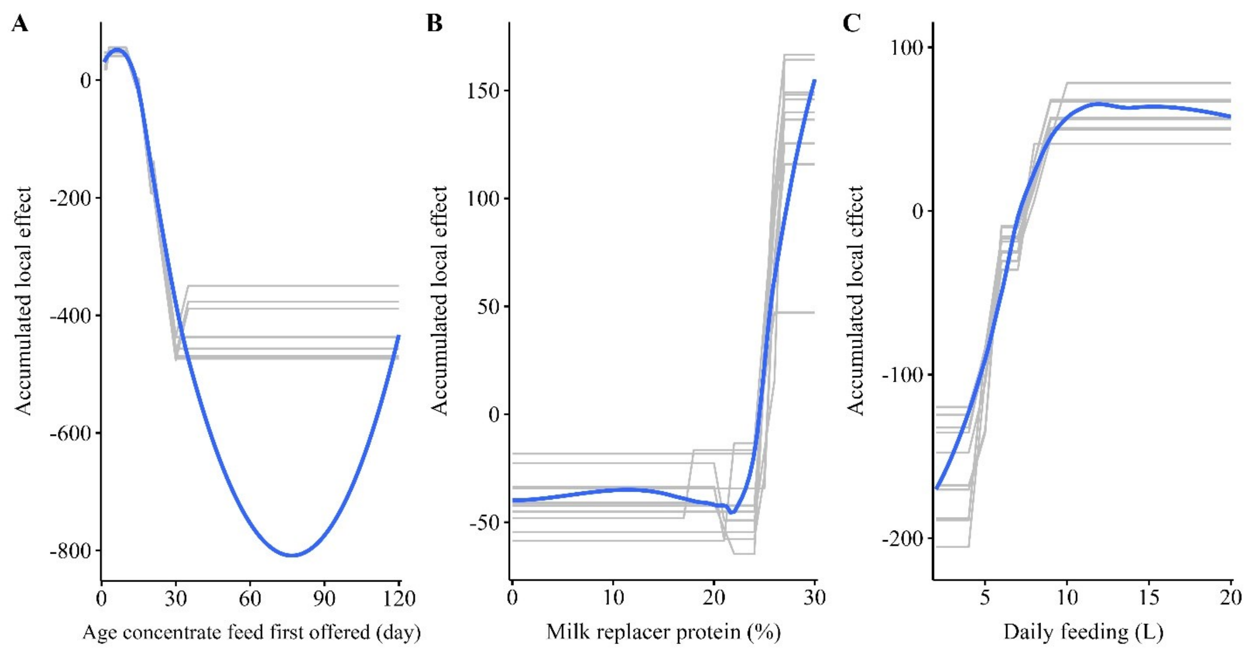


Figure 5.4. The accumulative local effect (ALE) of the early life management practices (**A**) age when concentrate feed is first offered, (**B**) protein content concentration of milk replacer, and (**C**) daily feeding amount on predicting herd cumulative energy-corrected milk (ECM) across all 10 multiple imputed versions of the data (grey lines) followed by a loess trend line (blue). Practices shown were selected based on their importance.

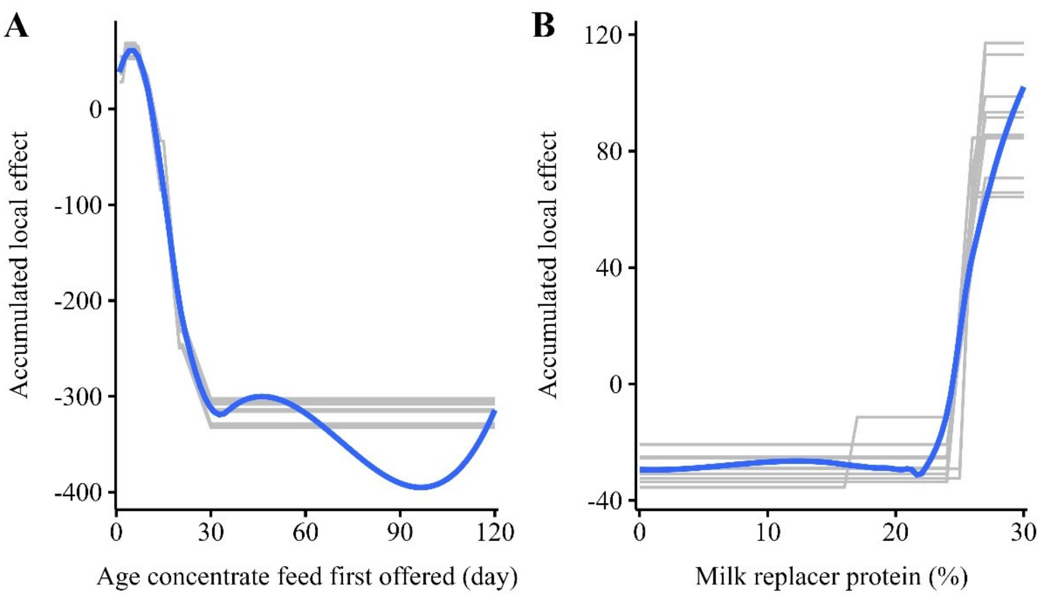


Figure 5.5. The accumulative local effect (ALE) of the early life management practices (**A**) age when concentrate feed is first offered and (**B**) protein content concentration of milk replacer on predicting herd cumulative milk value across all 10 multiple imputed versions of the data (grey lines) followed by a loess trend line (blue). Practices shown were selected based on their importance.

5.5. Discussion

Dairy farms were characterized based on their early life management practices using cluster analysis, which allowed to identify two stable clusters of farms that adopted different sets of early life management practices. Next, associations between practices that best described the cluster of farms and herd longevity (i.e., LPL and 3+ lactation), productivity (i.e., cumulative milk production), and profitability (i.e., cumulative milk value) were evaluated using machine learning models.

Early life management practices associated with feeding and housing mostly characterized the two clusters of farms identified in this study. Milk source was the main difference between them. While the cluster with most farms (64%) reported the use of powdered milk replacer, the remaining farms used either whole or waste milk. Cost, simplicity, and perceived calf performance are among the factors considered by farmers to choose between feeding practices (Vasseur et al., 2010, Palczynski et al., 2020). Event though most of the farms used powdered milk replacer, feeding whole milk has been shown to promote better growth (Godden et al., 2005), even when adjusting for the gross nutrient composition of both sources (Lee et al., 2009). The daily volume fed to calves in clusters 1 and 2 differed compared to the overall average, but they were both similar to the recommendations outlined in the Canadian Code of Practice (DFC & NFACC, 2009).

Housing conditions were also important in differentiating farms from both clusters, in which group housing was adopted by cluster 1 while individual housing was used by cluster 2. Housing animals in small groups have a beneficial effect on animal welfare, by promoting the expression of positive behaviors (Barry et al., 2020), and improve animal performance (Svensson and Liberg, 2006).

A clear pattern was found between the practices that best described the clusters and the outcomes of the farms in each respective cluster. Farms on cluster 1 were characterized as production-oriented since the management practices adopted by them were associated with an overall increase in both production and profitability but a reduction of herd longevity. The opposite was found for farms on cluster 2, which were characterized as longevity-oriented since their practices were associated with increased longevity, but reduced production and reduced profitability. Though rarely considered, the long-term effect of early life management practices on herd longevity, productivity, and profitability have also been shown in previous studies. For instance, eating high amounts of concentrate feed at an older age was associated with a greater longevity, but also higher lifetime milk, fat, and protein production (Heinrichs and Heinrichs, 2011). Other early life management practices such as heifer housing (Hultgren and Svensson, 2009) as well as age and body condition score at first breeding (Hultgren et al., 2011) were associated with the risk of culling and lifetime net milk revenue, respectively. However, culling, which dictates cow longevity as well as lifetime cumulative milk production and milk value, is a management decision often carried out by the farmer considering animal performance ranked within-herd (Essl, 1998). Such decision is often subjected to farmers' production priorities (Rilanto et al., 2022) and based on current lactation events alone (Beaudeau et al., 2000). Therefore, rather than showing causal effects, our results depict which early life management practices are often adopted by herds that are either production-oriented – higher relative production, but lower herd longevity – or longevity-oriented – lower relative production, but higher herd longevity.

Increasing dairy cow longevity is a strategy to achieve an economically sustainable dairy farm, mainly because dairy cows become profitable starting from the third lactation due to rearing costs (Horn et al., 2012, Delgado et al., 2017, Habel et al., 2021). Contrary to this, our results indicated that herds with higher animal lifetime cumulative ECM and milk value were also associated with reduced longevity. Milk production is one of the main reasons for culling (CDIC, 2021a), which in turn influences herd longevity. For instance, the occurrence of even a single case of mastitis and lameness are associated with a reduction on milk production (Puerto et al., 2021a,

b) and, consequently, milk value due to the amount of sellable milk. However, milk value does not consider costs associated with the treatment of health issues as well as reproduction failures (Delgado et al., 2017), and does not fully reveal the state of cow profitability. In addition, high yielding animals are more likely to develop health issues (Fleischer et al., 2001), which will have a negative effect on milk production (Rushen et al., 2008). The reduced production would still remain similar to cows with a lower relative yield but that were not affected by any disease (Puerto et al., 2021a, b). Therefore, the associations between management practices and herd longevity, productivity, and profitability found in our study may reflect that production-oriented farms have high yielding animals that produce more milk, resulting in a higher lifetime cumulative ECM and milk value, despite the shorter longevity compared to longevity-oriented farms.

The present study focused on early life management practices and indicators of longevity, productivity, and profitability from dairy herds enrolled with the Canadian DHI agency. Though about 64% (3,028 farms) of Quebec (Canada) dairy herds are enrolled in milk recording programs (CDIC, 2021b, c), the surveyed farms may not be representative of the whole province. More specifically, herds with automatic milking system (AMS) are often not part of DHI programs since they may chose to rely on the AMS measurements already available. Costs associated with health events are important to fully assess the profitability of a dairy farm on top of milk value (Delgado et al., 2017). Unfortunately, these are often not quantified in DHI records as the methodology for such quantification may not be straightforward since it must consider the milk losses and disease severity as well as treatment-associated expenditures, which could come from different databases (Dolecheck and Bewley, 2018). Additionally, our study focused on herd-level outcome indicators. However, cow-level factors may account for as much as 70% of the profitability measured as net lifetime milk revenue (i.e., milk value minus estimated rearing expenses and estimated feed costs), while the remaining 30% is accounted for farm-level factors (Hultgren et al., 2011).

5.6. Conclusion

Dairy farms were characterized as production-oriented or longevity-oriented based on their early life management practices and the associations between practices and the farm outcomes of longevity (i.e., LPL and 3+ lactation), productivity (i.e., lifetime cumulative ECM), and profitability (i.e., lifetime cumulative milk value). Production-oriented farms were mainly described by using powdered milk replacer to automatically feed calves housed in groups before and after weaning. These management practices were associated with an overall increased productivity and profitability but reduced longevity. On the other hand, longevity-oriented farms did not evaluate colostrum IgG concentration, fed calves with either whole or waste non-pasteurized/non-acidified milk, and housed weaned calves individually tied up in stalls. In turn, these practices were associated with increased longevity but reduced productivity and reduced profitability. Our results highlighted that early life management practices are linked with herd outcomes of longevity, profitability, and productivity, but the adoption of the best practices is not necessarily associated with better herd outcomes.

5.7. References

Abuelo, A., F. Cullens, A. Hanes, and J. L. Brester. 2021. Impact of 2 versus 1 colostrum meals on failure of transfer of passive immunity, pre-weaning morbidity and mortality, and performance of dairy calves in a large dairy herd. Animals 11(3):782. https://doi.org/10.3390/ani11030782.

Adriaens, I., N. C. Friggens, W. Ouweltjes, H. Scott, B. Aernouts, and J. Statham. 2020. Productive life span and resilience rank can be predicted from on-farm first-parity sensor time series but not using a common equation across farms. Journal of Dairy Science 103(8):7155-7171. https://doi.org/10.3168/jds.2019-17826.

Barry, J., E. A. M. Bokkers, I. J. M. de Boer, and E. Kennedy. 2020. Pre-weaning management of calves on commercial dairy farms and its influence on calf welfare and mortality. Animal 14(12):2580-2587. https://doi.org/10.1017/S1751731120001615.

Beaudeau, F., H. Seegers, V. Ducrocq, C. Fourichon, and N. Bareille. 2000. Effect of health disorders on culling in dairy cows: a review and a critical discussion. Ann. Zootech. 49(4):293-311. https://doi.org/10.1051/animres:2000102.

Benbrook, C., C. Carman, E. A. Clark, C. Daley, W. Fulwider, M. Hansen, C. Leifert, K. Martens, L. Paine, L. Petkewitz, G. Jodarski, F. Thicke, J. Velez, and G. Wegner. 2010. A dairy farm's footprint: Evaluating the impacts of conventional and organic farming systems. Critical Issue Report. The Organic Center, Washington D.C., USA.

CDIC – Canadian Dairy Information Centre. 2021a. Culling and replacement rates in dairy herds in Canada. Accessed on 07 July 2021. https://agriculture.canada.ca/en/canadas-agriculture-sectors/animal-industry/canadian-dairy-information-centre/dairy-statistics-and-market-information/dairy-animal-genetics/culling-and-replacement-rates-dairy-herds-canada.

CDIC – Canadian Dairy Information Centre. 2021b. Dairy barns by type in Canada. Accessed on 07 December 2021. https://agriculture.canada.ca/en/canadas-agriculture-sectors/animal-industry/canadian-dairy-information-centre/dairy-statistics-and-market-information/farm-statistics/dairy-barns-type.

CDIC – Canadian Dairy Information Centre. 2021c. Number of farms with shipments of milk by province. Accessed on 07 December 2021. https://agriculture.canada.ca/en/canadas-agriculture-sectors/animal-industry/canadian-dairy-information-centre/dairy-statistics-and-market-information/farm-statistics/farms-dairy-cows-and-dairy-heifers.

Dallago, G. M., K. M. Wade, R. I. Cue, J. T. McClure, R. Lacroix, D. Pellerin, and E. Vasseur. 2021. Keeping dairy cows for longer: A critical literature review on dairy cow longevity in high milk-producing countries. Animals 11(3):808. https://doi.org/10.3390/ani11030808.

Delgado, H. A., R. I. Cue, D. Haine, A. Sewalem, R. Lacroix, D. Lefebvre, J. Dubuc, E. Bouchard, and K. M. Wade. 2017. Profitability measures as decision-making tools for Québec dairy herds. Canadian Journal of Animal Science 98(1):18-31. https://doi.org/10.1139/cjas-2016-0202.

DeVries, T. J., M. G. Aarnoudse, H. W. Barkema, K. E. Leslie, and M. A. G. von Keyserlingk. 2012. Associations of dairy cow behavior, barn hygiene, cow hygiene, and risk of elevated somatic cell count. Journal of Dairy Science 95(10):5730-5739. https://doi.org/10.3168/jds.2012-5375.

DFC – Dairy Farmers of Canada. 2021. proAction®: On-farm excellence. Accessed on 14 June 2021. https://www.dairyfarmers.ca/proaction.

DFC & NFACC – Dairy Farmers of Canada & National Farm Animal Care Council. 2009. Code of Practice: For the care and handling of dairy cattle. Accessed on 17 September 2021. https://www.nfacc.ca/codes-of-practice/dairy-cattle#current.

Dolecheck, K. and J. Bewley. 2018. Animal board invited review: Dairy cow lameness expenditures, losses and total cost. animal 12(7):1462-1474. https://doi.org/10.1017/S1751731118000575.

Essl, A. 1998. Longevity in dairy cattle breeding: A review. Livestock Production Science 57(1):79-89. https://doi.org/10.1016/S0301-6226(98)00160-2.

Fleischer, P., M. Metzner, M. Beyerbach, M. Hoedemaker, and W. Klee. 2001. The relationship between milk yield and the incidence of some diseases in dairy cows. Journal of Dairy Science 84(9):2025-2035. https://doi.org/10.3168/jds.S0022-0302(01)74646-2.

Godden, S. M., J. P. Fetrow, J. M. Feirtag, L. R. Green, and S. J. Wells. 2005. Economic analysis of feeding pasteurized nonsaleable milk versus conventional milk replacer to dairy calves. Journal of the American Veterinary Medical Association 226(9):1547-1554. https://doi.org/10.2460/javma.2005.226.1547.

Gower, J. C. 1971. A general coefficient of similarity and some of its properties. Biometrics 27(4):857-871. https://doi.org/10.2307/2528823.

Grandl, F., M. Furger, M. Kreuzer, and M. Zehetmeier. 2019. Impact of longevity on greenhouse gas emissions and profitability of individual dairy cows analysed with different system boundaries. Animal 13(1):198-208. https://doi.org/10.1017/S175173111800112X.

Habel, J., V. Uhlig, S. Hoischen-Tauber, E.-M. Schwabenbauer, T. Rumphorst, L. Ebert, D. Möller, and A. Sundrum. 2021. Income over service life cost – Estimation of individual profitability of dairy cows at time of death reveals farm-specific economic trade-offs. Livestock Science 254:104765. https://doi.org/10.1016/j.livsci.2021.104765.

Heinrichs, A. J. and B. S. Heinrichs. 2011. A prospective study of calf factors affecting first-lactation and lifetime milk production and age of cows when removed from the herd. Journal of Dairy Science 94(1):336-341. https://doi.org/10.3168/jds.2010-3170.

Hennig, C. 2020. fpc: Flexible procedures for clustering.Version 2.2-9. https://CRAN.R-project.org/package=fpc.

Horn, M., W. Knaus, L. Kirner, and A. Steinwidder. 2012. Economic evaluation of longevity in organic dairy cows. Organic Agriculture 2(2):127-143. https://doi.org/10.1007/s13165-012-0027-6.

Hultgren, J. and C. Svensson. 2009. Heifer rearing conditions affect length of productive life in Swedish dairy cows. Preventive Veterinary Medicine 89(3):255-264. https://doi.org/10.1016/j.prevetmed.2009.02.012.

Hultgren, J., C. Svensson, and M. Pehrsson. 2011. Rearing conditions and lifetime milk revenues in Swedish dairy cows. Livestock Science 137(1):108-115. https://doi.org/10.1016/j.livsci.2010.10.005.

Husson, F., S. Le, and J. Pagès. 2017. Exploratory multivariate analysis by example using R. 2nd ed. Chapman and Hall/CRC, New York. https://doi.org/10.1201/b21874.

Kim, S. and H. Kim. 2016. A new metric of absolute percentage error for intermittent demand forecasts. International Journal of Forecasting 32(3):669-679. https://doi.org/10.1016/j.ijforecast.2015.12.003.

Kuhn, M. 2014. Futility analysis in the cross-validation of machine learning models. arXiv:1405.6974v1.

Kuhn, M. 2020. caret: Classification and regression training.Version 6.0-86. https://CRAN.R-project.org/package=caret.

Lee, H. J., M. A. Khan, W. S. Lee, S. H. Yang, S. B. Kim, K. S. Ki, H. S. Kim, J. K. Ha, and Y. J. Choi. 2009. Influence of equalizing the gross composition of milk replacer to that of whole milk on the performance of Holstein calves. Journal of Animal Science 87(3):1129-1137. https://doi.org/10.2527/jas.2008-1110.

Maechler, M., P. Rousseeuw, A. Struyf, M. Hubert, and K. Hornik. 2021. cluster: Cluster analysis basics and extensions.Version 2.1.2. Accessed November 25, 2021.

McPherson, S. E. and E. Vasseur. 2020. Graduate student literature review: The effects of bedding, stall length, and manger wall height on common outcome measures of dairy cow welfare in stall-based housing systems. Journal of Dairy Science 103(11):10940-10950. https://doi.org/10.3168/jds.2020-18332.

Molnar, C. 2019. Interpretable Machine Learning. Christoph Molnar.

Molnar, C., G. Casalicchio, and B. Bischl. 2018. iml: An R package for interpretable machine learning. Journal of Open Source Software 3(26):786. https://doi.org/10.21105/joss.00786.

Palczynski, L. J., E. C. L. Bleach, M. L. Brennan, and P. A. Robinson. 2020. Appropriate dairy calf feeding from birth to weaning: “It’s an investment for the future”. Animals 10(1):116. https://doi.org/10.3390/ani10010116.

Pinedo, P. J., A. De Vries, and D. W. Webb. 2010. Dynamics of culling risk with disposal codes reported by Dairy Herd Improvement dairy herds. Journal of Dairy Science 93(5):2250-2261. https://doi.org/10.3168/jds.2009-2572.

Puerto, M. A., E. Shepley, R. I. Cue, D. Warner, J. Dubuc, and E. Vasseur. 2021a. The hidden cost of disease: I. Impact of the first incidence of mastitis on production and economic indicators of primiparous dairy cows. Journal of Dairy Science 104(7):7932-7943. https://doi.org/10.3168/jds.2020-19584.

Puerto, M. A., E. Shepley, R. I. Cue, D. Warner, J. Dubuc, and E. Vasseur. 2021b. The hidden cost of disease: II. Impact of the first incidence of lameness on production and economic indicators of primiparous dairy cows. Journal of Dairy Science 104(7):7944-7955. https://doi.org/10.3168/jds.2020-19585.

Reneau, J. K., A. J. Seykora, B. J. Heins, M. I. Endres, R. J. Farnsworth, and R. F. Bey. 2005. Association between hygiene scores and somatic cell scores in dairy cattle. Journal of the American Veterinary Medical Association 227(8):1297-1301. https://doi.org/10.2460/javma.2005.227.1297.

Rilanto, T., D.-A. Viidu, T. Kaart, T. Orro, A. Viltrop, U. Emanuelson, E. Ferguson, and K. Mõtus. 2022. Attitudes and personality of farm managers and association with cow culling rates and longevity in large-scale commercial dairy farms. Research in Veterinary Science 142:31-42. https://doi.org/10.1016/j.rvsc.2021.11.006.

Rushen, J., A. M. Passillé, M. A. G. von Keyserlingk, and D. M. Weary. 2008. The welfare of cattle. Vol. 5. Animal Welfare Series. Springer, Dordrecht, Netherlands.

Santman-Berends, I. M. G. A., J. M. Swinkels, T. J. G. M. Lam, J. Keurentjes, and G. van Schaik. 2016. Evaluation of udder health parameters and risk factors for clinical mastitis in Dutch dairy herds in the context of a restricted antimicrobial usage policy. Journal of Dairy Science 99(4):2930-2939. https://doi.org/10.3168/jds.2015-10398.

Schreiner, D. A. and P. L. Ruegg. 2003. Relationship between udder and leg hygiene scores and subclinical mastitis. Journal of Dairy Science 86(11):3460-3465. https://doi.org/10.3168/jds.S0022-0302(03)73950-2.

Sebastien Le, Julie Josse, and F. Husson. 2008. FactoMineR: An R Package for Multivariate Analysis. Journal of Statistical Software 25(1):1-18. https://doi.org/10.18637/jss.v025.i01.

Somers, J. G. C. J., K. Frankena, E. N. Noordhuizen-Stassen, and J. H. M. Metz. 2003. Prevalence of claw disorders in Dutch dairy cows exposed to several floor systems. Journal of Dairy Science 86(6):2082-2093. https://doi.org/10.3168/jds.S0022-0302(03)73797-7.

Svensson, C. and P. Liberg. 2006. The effect of group size on health and growth rate of Swedish dairy calves housed in pens with automatic milk-feeders. Preventive Veterinary Medicine 73(1):43-53. https://doi.org/10.1016/j.prevetmed.2005.08.021.

Tang, F. and H. Ishwaran. 2017. Random forest missing data algorithms. Stat Anal Data Min 10(6):363 – 377. https://doi.org/10.1002/sam.11348.

van Buuren, S. 2019. Flexible imputation of missing data. 2 ed. Chapman & Hall/CRC, New York. https://doi.org/10.1201/9780429492259.

van Buuren, S. and K. Groothuis-Oudshoorn. 2011. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software 45(3):1 - 67. https://doi.org/10.18637/jss.v045.i03.

Vanegas, J., M. Overton, S. L. Berry, and W. M. Sischo. 2006. Effect of rubber flooring on claw health in lactating dairy cows housed in free-stall barns. Journal of Dairy Science 89(11):4251-4258. https://doi.org/10.3168/jds.S0022-0302(06)72471-7.

Vasseur, E., F. Borderas, R. I. Cue, D. Lefebvre, D. Pellerin, J. Rushen, K. M. Wade, and A. M. de Passillé. 2010. A survey of dairy calf management practices in Canada that affect animal welfare. Journal of Dairy Science 93(3):1307-1316. https://doi.org/10.3168/jds.2009-2429.

5.8. Supplementary material

Supplementary Table 5.1. Bootstrap stability assessment of up to 7 clusters obtained after 100 resampling runs of the clustering algorithms hierarchical with Ward's minimum variance method, partitioning around medoids (PAM), and normal mixture model-based. Overall and cluster-wise stable results are bolded.

Clustering algorithm	**Jaccard bootstrap mean distance**[1]							
	Overall[2]	**Cluster**						
		1	**2**	**3**	**4**	**5**	**6**	**7**
Hierarchical	**0.999**	**1.00**	**1.00**					
	0.922	**0.87**	**1.00**	**0.90**				
	0.807	**0.87**	**0.91**	0.71	**0.75**			
	0.659	**0.87**	0.51	**0.80**	**0.81**	0.31		
	0.761	**0.87**	**0.78**	**0.75**	**0.80**	0.62	0.74	
	0.754	**0.88**	**0.83**	0.72	**0.79**	0.60	0.73	0.74
PAM	**0.978**	**0.98**	**0.97**					
	0.853	**0.81**	**0.98**	**0.77**				
	0.693	0.62	**0.90**	0.72	0.53			
	0.638	0.59	0.73	0.72	0.61	0.54		
	0.622	0.57	0.71	0.63	0.67	0.61	0.54	
	0.583	0.56	0.66	0.64	0.61	0.55	0.54	0.52
Normal mixture model-based	**1.000**	**1.00**	**1.00**					
	0.956	**0.94**	**1.00**	**0.93**				
	0.877	**0.83**	0.74	**0.99**	**0.95**			
	0.645	**0.82**	0.71	0.60	0.61	0.49		
	0.706	**0.80**	0.71	0.60	**0.75**	0.52	**0.87**	
	0.633	0.69	0.64	0.43	0.67	0.66	0.54	0.79

[1] Below 0.60: cluster should not be trusted; between 0.60 and 0.75: indication of patterns in the data, but cluster membership is doubtful; 0.75 or more: cluster is stable; 0.85 and above: cluster is highly stable (Hennig, 2020).
[2] Calculated as the average of the Jaccard bootstrap mean distances.

Supplementary Table 5.2. Prediction performance of recursive partitioning and regression tree (RPART), gradient boosting machine (GBM), random forest (RF), and support vector machine (SVM) with a radial basis kernel models on each of the outcomes evaluated. Performance is shown both on training and validation data sets. Means ± standard deviations were calculated across all 10 multiple imputed versions of the data. Best results within outcome, data set, and metric are bolded.

Outcome[1]	Model	Training[2]				Validation[2]			
		R^2	RMSE	MAE	MAAPE	R^2	RMSE	MAE	MAAPE
LPL	RPART	---	0.76 ± 0.0001	0.59 ± 0.0001	18.0 ± 0.0001	---	0.75 ± 0.0001	0.58 ± 0.0001	17.5 ± 0.0001
	GBM	0.04 ± 0.002	0.74 ± 0.0004	0.57 ± 0.0004	17.6 ± 0.01	0.02 ± 0.002	0.75 ± 0.0007	0.58 ± 0.001	17.4 ± 0.04
	RF	**0.48 ± 0.073**	**0.67 ± 0.0128**	**0.52 ± 0.0098**	**16.0 ± 0.28**	**0.03 ± 0.003**	**0.74 ± 0.001**	**0.57 ± 0.001**	17.3 ± 0.03
	SVM	0.05 ± 0.001	0.74 ± 0.0005	0.56 ± 0.0002	16.7 ± 0.01	0.01 ± 0.001	0.76 ± 0.0004	**0.57 ± 0.001**	**16.8 ± 0.02**
3+ Lactation	RPART	---	8.16 ± 0.057	6.30 ± 0.041	15.8 ± 0.09	---	**7.88 ± 0.07**	6.21 ± 0.04	**15.4 ± 0.09**
	GBM	0.03 ± 0.001	8.08 ± 0.003	6.22 ± 0.005	15.6 ± 0.01	**0.0009 ± 0.0008**	7.90 ± 0.01	**6.19 ± 0.01**	**15.4 ± 0.02**
	RF	**0.60 ± 0.051**	**7.23 ± 0.104**	**5.58 ± 0.082**	**14.1 ± 0.19**	0.0003 ± 0.0003	7.89 ± 0.013	6.21 ± 0.01	**15.4 ± 0.03**
	SVM	0.06 ± 0.001	8.00 ± 0.003	6.08 ± 0.003	15.2 ± 0.01	0.0007 ± 0.0007	7.89 ± 0.005	6.22 ± 0.01	**15.4 ± 0.01**
ECM	RPART	---	1,548.0 ± 7.11	1,185.7 ± 5.59	10.7 ± 0.05	---	1,591.6 ± 7.30	1,206.4 ± 3.69	10.8 ± 0.03
	GBM	0.07 ± 0.001	1,511.1 ± 1.13	1,156.3 ± 1.34	10.5 ± 0.01	**0.09 ± 0.005**	1,549.2 ± 2.73	**1,173.4 ± 3.09**	**10.5 ± 0.03**
	RF	**0.47 ± 0.004**	**1,354.7 ± 1.73**	**1,035.6 ± 1.54**	**9.4 ± 0.01**	**0.09 ± 0.005**	1,550.8 ± 2.09	1,174.0 ± 2.50	**10.5 ± 0.02**
	SVM	0.09 ± 0.001	1,496.5 ± 0.84	1,123.9 ± 0.64	10.2 ± 0.01	0.08 ± 0.002	**1,546.9 ± 1.46**	1,175.4 ± 2.67	**10.5 ± 0.02**
MV	RPART	---	1,194.9 ± 6.84	879.3 ± 3.48	11.5 ± 0.05	---	1,111.4 ± 9.03	856.8 ± 10.00	10.9 ± 0.14
	GBM	0.07 ± 0.001	1,159.5 ± 0.60	853.5 ± 0.47	11.1 ± 0.01	**0.07 ± 0.004**	**1,080.5 ± 1.73**	**827.8 ± 1.90**	**10.5 ± 0.02**
	RF	**0.46 ± 0.048**	**1,041.2 ± 15.96**	**765.7 ± 10.75**	**10.1 ± 0.13**	0.05 ± 0.004	1,092.9 ± 1.82	842.4 ± 1.33	10.7 ± 0.02
	SVM	0.10 ± 0.001	1,147.3 ± 0.62	828.0 ± 0.44	10.8 ± 0.01	0.05 ± 0.002	1,087.5 ± 1.11	834.8 ± 1.26	10.6 ± 0.01

[1] LPL = Herd average length of productive life (year); 3+ lactation = Herd average percentage of cows on third lactation or greater (%); ECM = Herd average lifetime cumulative energy-corrected milk (kg); MV = Herd average lifetime cumulative milk value ($CAD).

[2] R^2 = Coefficient of determination; RMSE = Root mean squared error; MAE = Mean absolute error; MAAPE = Mean arctangent absolute percentage error (Kim and Kim, 2016).

5.9. Annex

Annex 5.1. Early life management practices questionnaire.

Colostrum

NAME OF HERD

For each question, consider the majority of female calves (75%). Do not describe exceptions
Colostrum: milk from the first milking

1. After how many hours is the calf separated from it's dam after his birth?

- ○ Less than 1hr
- ○ 1 to 6hrs
- ○ 7 to 12hrs
- ○ 13 to 24hrs
- ○ More than 24hrs

2. How many hours after birth does the calf receive it's first meal?

- ○ Less than 1h after birth
- ○ From 1 to 6hrs
- ○ From 7 to 12hrs
- ○ More than 12hrs

3. How is colostrum fed?

- ○ Feeding on mother
- ○ With a bottle
- ○ Bucket with teats
- ○ Bucket without teats
- ○ Esophageal tube
- ○ Other
- ○ No colostrum offered

3.1 If cow colostrum is offered, how is the IgG concentration evaluated?

- ○ Colostrometer
- ○ Refractometer
- ○ Visual evaluation
- ○ Not evaluated
- ○ Other

3.2 What is the main source of colostrum?

- ○ Individual cows
- ○ Pool of cows
- ○ Powdered colostrum
- ○ Other

3.2.1 What is the form of the colostrum? <u>(FOR INDIVIDUAL COWS / POOL OF COWS)</u>

- ○ Fresh pasteurized
- ○ Fresh non pasteurized
- ○ Frozen pasteurized
- ○ Frozen non pasteurized

3.3 How many colostrum meals are fed on the 1st day?

- ○ 1
- ○ 2
- ○ 3

3.4 Indicate the average quantity of colostrum <u>fed</u> during these specific meals:

3.4.1 First meal

- ○ 0 liter
- ○ 1 liter
- ○ 2 liters
- ○ 3 liters
- ○ 4 liters or more

3.4.2 Second meal

- ◯ 0 liter
- ◯ 1 liter
- ◯ 2 liters
- ◯ 3 liters
- ◯ 4 liters or more

---

4. Are male calves managed in the same way as female calves?

- ◯ Yes
- ◯ No

4.1 In what ways are they managed differently?

- Quantity of colostrum ☐
- Quality of colostrum ☐
- Quantity of milk ☐
- Source of milk ☐
- Housing ☐

Milk

5. What system is mainly used to feed milk to calves?

- ◯ Bottle
- ◯ Individual bucket with teats
- ◯ Individual bucket without teats
- ◯ Group bucket
- ◯ Group bucket with teats
- ◯ Feed line (free feed)
- ◯ Automatic feeding system
- ◯ Floating teat
- ◯ Nursing cow

---

6. What is the main source of milk for calves?

- ◯ Whole milk
- ◯ Waste milk
- ◯ Powdered milk replacer

---

7. What is the state of the milk? (FOR WHOLE MILK OR WASTE MILK)

- ◯ Pasturized / Non acidified
- ◯ Pasteurized / Acidified
- ◯ Non pasturized / Non acidified
- ◯ Non pasturized / Acidified

---

7. What is the state of the milk? (FOR POWDERED MILK REPLACER)

- ◯ Non acidified
- ◯ Acidified
- ◯ Partially acidified

7.1. Please enter information on the milk replacer:

7.1.a % fat: (From 15 to 30)

7.1.b % protein: (From 12 to 30)

7.1.c Is it medicated?

☐

---

8. What is the total quantity of milk fed per day in liters? (From 1 to 20 liters)

---

9. Normally how many meals are fed per day?

- ○ 1
- ○ 2
- ○ 3
- ○ Free-choice (Automated feeder, acidified milk, nursing cow)

Solid feed and Weaning

10. At what age are the following feed offered for the first time?

10.a Water (Days) (Max 120 days)

10.b Starter feed or grains (Days) (Max 120 days)

10.c Hay or other forages (Days) (Max 120 days)

10.d Fermented feeds (silage, TMR, wet bales) (Months)

11. What is the type of starter offered?

- Commercial meal mix
- Own grain and concentrate mix
- Dry TMR
- Other

11.1 What is the protein percentage?

- 16
- 18
- 20
- 22
- 24
- 26
- 28
- 30

12. At what age are heifers weaned (days)?

- Less than 50
- From 51 to 70
- From 71 to 90
- More than 90

13. What is the main factor that dictates weaning?

- Consumes at least 1 kg of starter feed for at least 3 consecutive days
- Consumes at least 2 kg of starter feed for at least 3 consecutive days
- Reaches the target age for weaning
- Reaches the target weight for weaning
- Space required for new calves
- Other

13.1 The amount is: <u>(FOR CONSUMES AT LEAST 1 KG / 2 KG)</u>

- Estimated
- Weighed

13.1 The weight is: <u>(FOR THE TARGET WEIGHT)</u>

- Estimated
- Measured

---

14 What is the principal factor which dictates breeding?

- Reaches the target age for breeding
- Reaches the target weight for breeding
- Reaches the target height for breeding
- The ideal time to fill quota
- Other

14.1 The weight is: (FOR THE TARGET WEIGHT)

- Estimated
- Measured

14.1 The height is: (FOR THE TARGET HEIGHT)

- Estimated
- Measured

Housing

15. Are housing walls cleaned and disinfected between every calf or lots of calves?

- ◯ Yes
- ◯ No

---

16. What is the main housing type for non-weaned calves?

- ◯ Individual loose housing
- ◯ Individual tied up
- ◯ Group(s)

16.1 Specify: (FOR INDIVIDUAL LOOSE HOUSING)

- ◯ Hutch
- ◯ Box

16.1 Specify: (FOR INDIVIDUAL TIED UP)

- ◯ Hutch
- ◯ Box
- ◯ Stall
- ◯ Wall

16.1 Specify: (FOR GROUP(S))

- ◯ Hutch
- ◯ Pen

---

17. What type of bedding do you use for non-weaned calves?

- ☐ Cereal Straw
- ☐ Peat moss
- ☐ Wood (sawdust/shavings)
- ☐ Other

---

18. At what frequency is bedding added?

Number of times per:

○ day
○ week
○ month

---

19. What is the main type of housing used immediately after weaning?

○ Individual loose housing
○ Individual tied up
○ Group(s)

19.1 Specify: (FOR INDIVIDUAL LOOSE HOUSING)

○ Hutch
○ Box

19.1 Specify: (FOR INDIVIDUAL TIED UP)

○ Hutch
○ Box
○ Stall
○ Wall

19.1 Specify: (FOR GROUP(S))

○ Hutch
○ Pen

19.2 How long the calf stay in that housing following weaning? {0} weeks

weeks

Connecting statement 4

The relationships between early life and longevity, productivity, and profitability have been analysed up to this point in this thesis from different perspectives. From an animal centric perspective, however, longevity, and both productivity and profitability could be considered components of dairy farming environmental and economic sustainability, respectively. Animal welfare would make up the remaining social sustainability component. The sustainability of milk production is often questioned, especially when considering the increased availability of alternative substitutes, which have a similar nutritional composition but are relatively more sustainable. Therefore, indicators of longevity, productivity, profitability, and animal welfare could be used to assess the sustainability of dairy farming.

CHAPTER 6 – Mapping the sustainability of dairy farming through an animal centric approach

Gabriel M. Dallago[1,*], Daniel Warner[2], Kevin M. Wade[1], Roger I. Cue[1], René Lacroix[2], Abdoulaye Baniré Diallo[3], and Elsa Vasseur[4]

[1] Department of Animal Science, McGill University, Sainte-Anne-de-Bellevue, Quebec, H9X 3V9, Canada.

[2] Lactanet, Sainte-Anne-de-Bellevue, Quebec, H9X 3R4, Canada.

[3] Department of Computer Science, Université du Québec à Montréal, Montreal, Quebec, H3C 3P8 Canada.

* Corresponding author: gabriel.dallago@mail.mcgill.ca

Manuscript in preparation to be submitted Nature Sustainability

6.1. Abstract

Animal-sourced foods are criticized for their negative environmental impact, but they play an important role in achieving food security. Therefore, promoting their sustainability is imperative. From an animal-centric approach, social, economic, and environmental aspects of sustainability are based on animal welfare, production, and performance, respectively. Here we identified different welfare profiles based on animal welfare outcomes, which were used to describe social sustainability of dairy farms. Relationships between welfare profiles and production and performance indicators were used to analyze the interdependency between social sustainability and economic and environmental sustainability. Three welfare profiles were found and most of the herds were in the best welfare profile. Weak relationships between welfare profiles and performance and production indicators were found, highlighting that a balance must be achieved between the three profiles to sustainably meet the future demand for milk and better indicators are needed to fully map the sustainability of dairy farming.

6.2. Introduction

A sustainable food production system makes use of renewable resources at a rate that does not surpass Earth's potential to replenish them (Godfray et al., 2010). Such definition can be broken down into the environmental, economic, and social pillars (Boogaard et al., 2011) and a system is only fully sustainable if sustainability is achieved in all of them. Animal-sourced foods are criticized for their negative environmental impact as they contribute to most of the greenhouse gas emissions when compared to substitute vegetable protein sources (Poore and Nemecek, 2018, Springmann et al., 2018). Though reducing its amount in human diets has the greatest potential to ease the environmental pressure of food production systems (Stehfest et al., 2009, Tilman and Clark, 2014, Springmann et al., 2018), dietary change alone would not be enough to keep the impact of food systems within planetary boundaries (Springmann et al., 2018). Animal-sourced food also play an important role on food security (Tricarico et al., 2020). Dairy farming, specially, provides milk, which is a non-luxury rich source of nutrients to infants and adults while being the main source of income for many farmers globally (Herrero et al., 2013, Segerkvist et al., 2020). Therefore it is imperative to support the sustainability of dairy farming since there is potential to reduce its environmental impact (Poore and Nemecek, 2018).

When studying dairy farming sustainability, the environmental and economic aspects have received the greatest attention. (Lebacq et al., 2015, Galloway et al., 2018, Herzog et al., 2018, Segerkvist et al., 2020), while the social aspects (Lebacq et al., 2013) as well as relationships between all three pillar have been overlooked (Herzog et al., 2018). Animal welfare is a core concept of dairy farming sustainability as good welfare status is necessary for animals to be productive and profitable (Villettaz Robichaud et al., 2019a, Villettaz Robichaud et al., 2019b). For example, the occurrence of diseases such as mastitis and lameness, which are among the most prevalent, not only negatively impact the welfare status of cows, but also impact both animal production and profitability (Puerto et al., 2021a, b). In addition, the occurrence of health problems greatly affect the ability of farmer to keep animals in the herd (Dallago et al., 2021). Therefore, sustainability evaluations should be centered on animals rather than the herd, which allow for uncovering synergies between sustainability aspects of dairy farming (Hoischen-Taubner et al., 2021).

Animal indicators, covering all three pillars of sustainability, should be therefore considered. Production and performance metrics are regularly collected and maintained in Dairy Herd Improvement (**DHI**) databases, making it possible to not only obtain both milk production and milk values, but cow longevity as well. Milk production and milk value make up the economic pillar (Lovarelli et al., 2020, Segerkvist et al., 2020). Cow longevity, which can be measured as the length of productive life and the percentage of cows on third or greater lactation (Dallago et al., 2021), is an animal centric indicator of environmental sustainability as high cow longevity is associated with a reduction on the emission of methane with each kilogram of milk (Grandl et al., 2019) as well as total methane emission of the herd (Knapp et al., 2014). Animal welfare outcome measures, which are direct indicators of animal welfare (Broom, 1991, Knierim and Winckler, 2009) and newly routinely collected (DFC, 2019, 2021), comprises the social component of sustainability because they not only address ethical concerns raised by the society but they also translate the ability of farmers to fulfill the needs of the animals by categorizing management and housing styles (Lovarelli et al., 2020, Segerkvist et al., 2020). Therefore, making it possible to map the sustainability of dairy farming using an animal centric approach. In this paper, we describe the social sustainability profile of dairy farms and analyse its relationship with the economic and environmental profiles.

6.3. Results

Description of the social profile based on animal welfare outcome measures are reported first, followed by its association with the economic and environmental profiles. Descriptive statistics of the data analysed in our study is shown in Table 6.1.

Table 6.1. Descriptive statistics of welfare and dairy herd improvement indicators as well as supplementary variables considered in this study from 2,980 Quebec – Canada dairy farms.

Measure	Mean	Standard deviation	Minimum	Maximum
Numeric indicators				
BCS ≤ 2 (%)[1]	1.8	4.10	0.0	39.0
Hock lesion (%)	17.9	14.91	0.0	91.0
Knee lesion (%)	7.4	9.06	0.0	71.0
Neck lesion (%)	4.3	7.31	0.0	61.0
Lameness (%)	10.4	10.01	0.0	99.0
Herd status index[2]	0.52	0.10	0.17	0.80
ECM (kg)[3]	11,271.4	1,714.35	4,244.0	23,490.0
Milk value ($CAD)[4]	7,950.8	1,256.31	1,393.1	16,685.2
Length of productive life (years)	3.3	0.77	0.46	7.9
3+ lactation (%)[5]	41.3	8.22	7.33	75.8
Supplementary variables	N		%	
Barn type				
Tie-stall	2,412		80.9	
Free-stall	568		19.1	
Season				
Winter	736		24.7	
Spring	727		24.4	
Summer	649		21.8	
Fall	868		29.1	
Year				
2016	2		0.1	
2017	961		32.2	
2018	1,781		59.8	
2019	236		7.9	

[1] Prevalence of cows with body condition score lower than or equal to two.
[2] Herd composite indicator for remote assessment of herd welfare status (Warner et al., 2020).
[3] Herd average animal lifetime cumulative energy-corrected milk.
[4] Herd average animal lifetime cumulative milk value.
[5] Herd average percentage of cows on third or greater lactations.

By analyzing animal welfare outcomes, which highlight how the animals respond to their environment, we revealed the existence of different animal welfare profiles among dairy herds. Even though we identified herds in which the prevalence of all welfare issues was low, there were no herds in which all welfare issues had a high prevalence. Instead, each two or three of the welfare outcome measures were shown to be problematic for some specific group of farms (Figure 6.1). This influenced the number and the type of the clusters identified in our study. Herds were grouped into three stable and homogeneous welfare clusters (Figure 6.2 and Table 6.2). Herds on cluster 3 had the best overall welfare status since it had the lowest prevalence of most welfare issues and the highest Herd Status Index (HSI). Cluster 1 was described by a high percentage of cows with body conditions score lower than or equal to two, hock lesions, and lameness, whereas cluster 2 was mostly described by a high percentage of both neck and knee lesions (Figure 6.2 and Table 6.2).

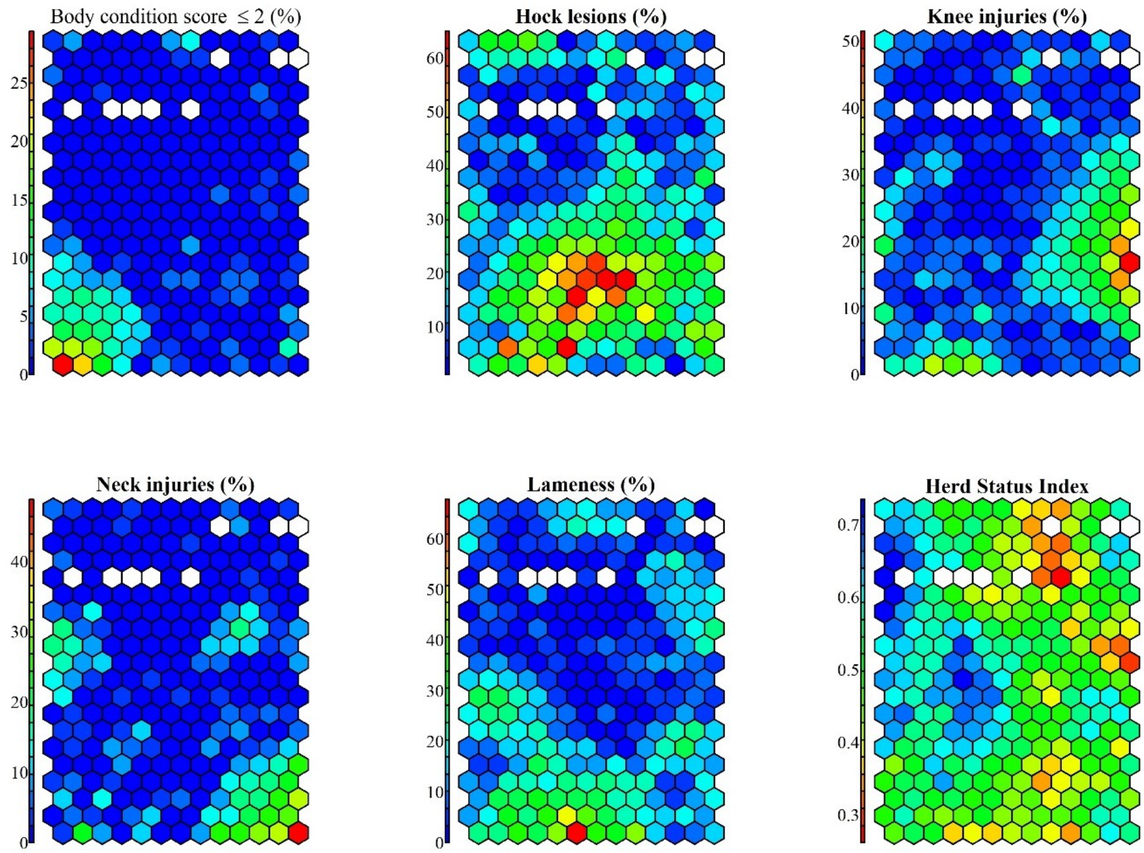


Figure 6.1. Heat maps produced by the self-organizing map algorithm showing the distribution of the six welfare indicators collected on 2,982 dairy farms from Quebec – Canada. The map was built in a hexagonal topology with 13 × 20 units. Using average values of farms assigned to each unit, blue color indicates units with low values (desirable), red indicates units with high values (undesirable), and white represent empty units (i.e., units in which no farms were assigned). The map provides a visual qualification of relationships between the indicators, allowing to explore the value distribution of each welfare outcome in relationship with the other outcomes. For example, farms in the unit having the highest prevalence of cows with body condition score ≤ 2 (i.e., bottom left of the heat map) also had a low prevalence of neck injuries.

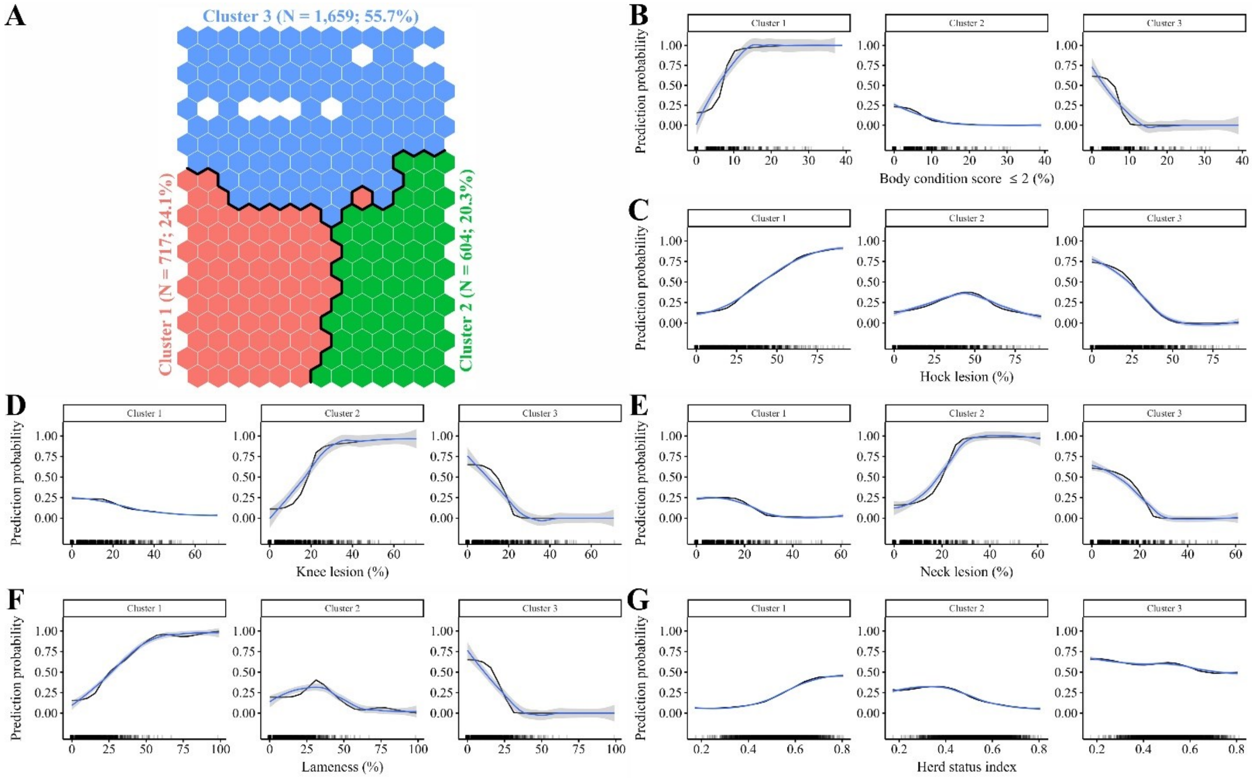


Figure 6.2. Self-organizing map with cluster divisions (**A**) and partial dependency plots (PDP) depicting the marginal effect of the herd welfare outcome measures of percentage of cows with body conditions score lower than or equal to two (**B**), hock lesion (**C**), knee lesion (**D**), neck lesion (**E**), lameness (**F**), and Herd Status Index (**G**) in each cluster. Cluster were created using partitioning around medoids algorithm since it produced clusters more stable and more homogeneous. Different classification machine learning models were trained to predict cluster labels based on herd welfare outcome measures which in turn were used to create the clusters. The model with the highest accuracy was used to generate PDPs to describe the clusters.

Table 6.2. Mean and standard deviation (SD) by clusters for the welfare and dairy herd improvement indicators of production and performance from 2,980 Quebec – Canada dairy farms.

Measure	**Cluster 1 (N = 717)**		**Cluster 2 (N = 604)**		**Cluster 3 (N = 1,659)**	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**
BCS ≤ 2 (%)[1]	5.2	6.58	1.3	2.51	0.5	1.57
Hock lesion (%)	28.2	16.85	25.5	14.07	10.7	9.28
Knee lesion (%)	7.8	8.47	16.2	12.19	4.0	4.74
Neck lesion (%)	4.1	5.90	9.9	44.59	2.3	4.09
Lameness (%)	17.3	11.67	13.9	10.01	6.1	6.40
Herd status index[2]	0.56	0.09	0.47	0.09	0.52	0.10
ECM (kg)[3]	11,459.2	1,637.84	11,047.9	1,650.56	11,271.6	1,760.09
Milk value ($CAD)[4]	8,081.3	1,197.99	7,782.4	1,212.72	7,955.7	1,289.75
Length of productive life (years)	3.4	0.77	3.3	0.76	3.3	0.77
3+ lactation (%)[5]	42.3	7.99	40.5	8.23	41.3	8.29

[1] Prevalence of cows with body condition score lower than or equal to two on a 5-point scale.
[2] Herd composite indicator for remote assessment of herd welfare status (Warner et al., 2020).
[3] Herd average animal lifetime cumulative energy-corrected milk.
[4] Herd average animal lifetime cumulative milk value.
[5] Herd average percentage of cows on third or greater lactations.

We found a weak relationship between the social and both economic and environmental profiles. This was observed based on the overall low accuracy of the models in predicting welfare clusters according to the production and performance indicators (Supplementary Table 6.4). Cluster 3 was mostly described by high energy-corrected milk (**ECM**) production, low milk value, and higher herd longevity (i.e., percentage of cows on third or greater lactation; **3+ lactation**; Figure 6.3 and Table 6.2). The opposite was observed for cluster 2, which was described by high milk value and low herd longevity (i.e., 3+ lactation; Figure 6.3 and Table 6.2).

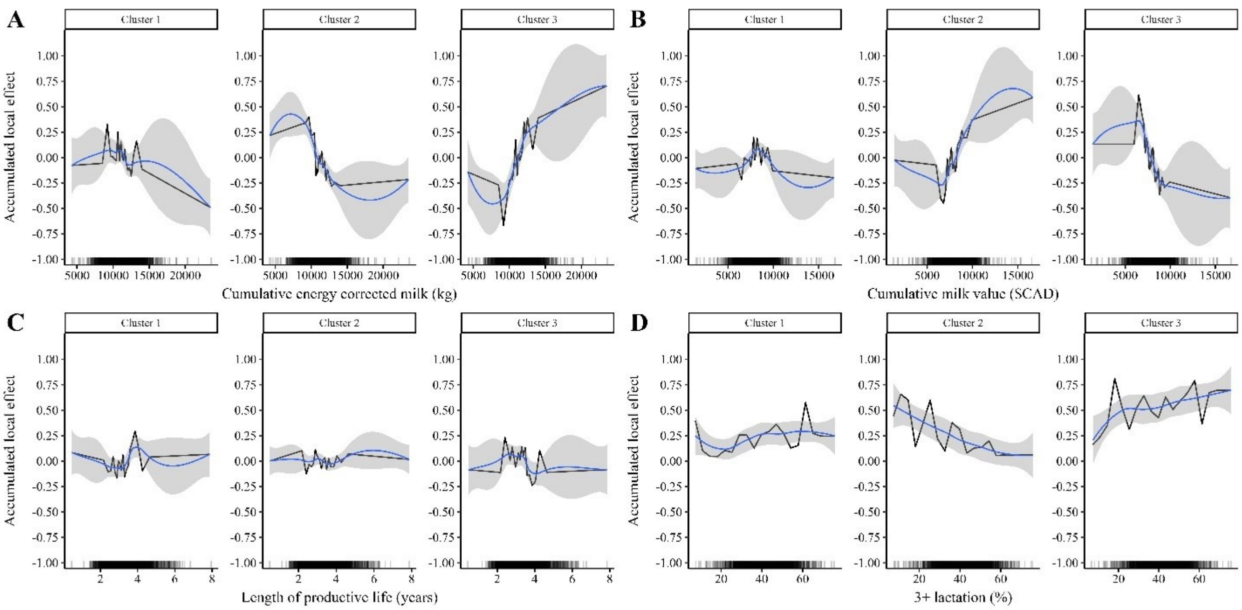


Figure 6.3. Accumulated local effect (ALE) plots depicting the average influence of production and performance indicators in predicting clusters with different welfare profiles. Different classification machine learning models were trained to predict cluster labels based on herd averages of cumulative energy corrected milk (kg), cumulative milk value ($CAD), length of productive life (years), and percentage of cows on third or greater lactation (3+ lactation; %). The model with the highest accuracy was used to generate ALE plots to analyse the relationship between herd welfare and both production and performance indicators.

6.4. Discussion

Animal welfare is composed of the biological functioning, affective state, and natural behavior of the animals (Fraser, 2008). Such definition led to the development of indicators covering each one of these aspects. Even though all aspects of welfare are correlated, only metrics of biological functioning are now routinely measured (DFC, 2019, FARM, 2020, DFC, 2021). Body condition score provides an indication of nutrition practices of the farms (Roche et al., 2009) and is associated with animal performance and health. Animals being both overly thin or overly conditioned are prone to health issues and compromised performance (Roche et al., 2009), which have a negative impacting on their welfare. Health issues and body lesions are a source of distress and pain for animals. Lameness, for example, is a major health welfare problem that occur in dairy cows (Whay et al., 1998, Rutherford et al., 2009, Popescu et al., 2014). Animals affected by this disease have reduced nociceptive threshold (Whay et al., 1998) and redistribute their body weight when they are walking, thus reducing the weight placed on the affected limb (Rushen et al., 2007). Both situations indicate that these animals are suffering from physical pain and therefore their welfare is compromised.

Animal welfare outcome measures reflect how different housings and management aspects affect dairy cows, which was highlighted in our analysis as different welfare profiles were found. A high prevalence of both lameness and lean animals were found to best describe one of the farm clusters. In fact, low body condition was shown to be one of the risk factors for the development of lameness (Jewell et al., 2019). The hoof of overly lean cows has less adipose tissue to support it (Bicalho et al., 2009), making for a high prevalence of lameness on lean cows (Espejo et al., 2006, Randall et al., 2015, Solano et al., 2015). However, this does not mean that a cow with low body condition will necessarily become lame, since there are different risk factors that can impact animals at the farm level. For instance, hard walking surface (Somers et al., 2003, Vanegas et al., 2006) and resting area comfort (McPherson and Vasseur, 2020) are direct risk factors associated with lameness while inappropriate nutritional status is a risk factor linked with low body condition. In addition, lameness could also be a risk factor for low body condition score. Given the painful nature of this disease, inflicted animals may have a reduced feed intake and willingness to visit feeding bunk (Norring et al., 2014), which will lead to a reduction in their body condition score. Our results indicate that there exist similar risk factors associated with both body condition and lameness, and that management practices should be adapted accordingly.

Absence of skin lesions give an indication on how adequate the housing conditions are for the animals. Cows spend about 12 to 13 hours a day laid down (Jensen et al., 2005, Fregonesi et al., 2007), but since they have little or no ability to change their lying down movement (Österman and Redbo, 2001) in the environment they are provided with, it is no surprise that the frequent contact with housing structures while lying down lead to an increase of skin lesions. Resting area characteristics such as base material (Nash et al., 2016) and bedding (Kielland et al., 2009) are management factors associated with the incidence of hock and knee lesions while neck injuries are more commonly associated with the dimensions of the resting area (Bouffard et al., 2017, Jewell et al., 2019). The prevalence of different welfare issues on distinct farm clusters indicates that targeted management practices and housing conditions must be implemented to further improve animal welfare and meet society ethical concerns.

We found a weak association between the welfare and both the economic (i.e., lifetime cumulative ECM and milk value) and environmental (i.e., LPL and 3+ lactation) profiles. Identifying and quantifying interdependencies among sustainability pillars is fundamental to seek ways to promote it (Segerkvist et al., 2020). Though the link between economic and environmental profiles are more commonly explored (Dallago et al., 2021), the link between the social, which is represented by animal welfare in an animal centric approach, and the other sustainability pillars remains mostly unexplored. The economic and environmental profiles are characterized by the productivity and resilience [i.e., high adaptability to challenges and cumulative good health and fertility, resulting in a longer longevity (Adriaens et al., 2020)] of an animal, respectively. In turn, our results imply that high welfare profile would lead to neither an increase nor reduction in production and resilience. Improving the welfare by reducing the occurrence of health issues and improving the nutritional status of the animals could actually promote an increase in production. However, the intense genetic selection for high milk production, which has occurred, has hindered animal reproduction (Bedere et al., 2018), welfare (Oltenacu and Broom, 2010), and overall health (Berry et al., 2011). Therefore, high milk production alone does not guarantee good animal welfare nor farm profitability, and a balance must be achieved between the three profiles to achieve the sustainability of dairy farming.

The sustainability mapping carried out in this study was as complete as possible given data availability. Not all aspects of social sustainability were covered, but also there remain a lack of additional indicators that represent both the economic and environmental sustainability in an animal centric approach. Milk production and milk value are both important indicators of economic sustainability. However, other factors such as feed and reproduction costs as well as costs associated with the treatment of diseases are important factors that make up the true profitability of dairy cows (Delgado et al., 2017, Puerto et al., 2021a, b). Though the social aspect was evaluated here using two of the most common metrics of longevity, different definitions of cow longevity exist and there is no single comprehensive metric available (Dallago et al., 2021). These could have been the reason for finding only weak links between the sustainability pillars in our study.

Farmers’ and animals’ interests should be aligned to meet the increasing demand for animal food products while ensuring the sustainability of the activity. Historically, increase in production has been achieved by intensification of production, driven by an increased use of resources and an intense genetic selection for high milk yield. This resulted in animals being more susceptible to challenging environments (Blanc et al., 2006) and highly dependent on optimum environmental conditions to express their genetic potential (VandeHaar et al., 2016). However, intensifying production is limited to resource availability which is becoming scarce and unpredictable (Godde et al., 2021). Animal food production efficiency should be achieved by simultaneously pursuing a broad range of options, such as improving breeding goals (Oltenacu and Algers, 2005) and management practices (Pretty and Bharucha, 2019), to establish a synergy between sustainability profiles. Resilient animals remain in the herd for longer, making them more efficient, profitable, and more likely to have a better welfare status. The pressing issues associated with climate change and competition for resources indicates that a change should occur soon and should no longer focus on one sustainability aspect alone but search for an optimum solution across its social, environmental, and economic aspects (Godfray et al., 2010).

6.5. Conclusion

We found different animal welfare profiles among dairy herds with currently routinely collected data. From an animal centric perspective, this not only highlights how the animals respond to their surrounding environment, but it also shows that a set of different solutions and innovations are needed to adapt management practices and housing conditions to enhance the social acceptability of dairy farming. Additionally, more comprehensive indicators of social, economic, and environmental profiles are necessary to fully map the sustainability of dairy farming. Only with those indicators, it will be possible to fully identify and quantify interdependencies among sustainability pillars, making possible to propose concrete strategies to sustainably produce milk to meet future demands.

6.6. Materials and methods

A cross-sectional study was conducted using welfare outcome measures from Quebec – Canada dairy herds that were collected by the Dairy Farmers of Canada as part of the Animal Care module of the proAction® Quality Assurance Program (DFC, 2021). Additionally, test day performance and production indicators were provided by Lactanet (Sainte-Anne-de-Bellevue, Quebec, Canada). Data handling, cleaning, and modelling were done using the R statistical software (R Core Team, 2021) and its specific packages. Self-organizing maps (**SOM**) were used for data dimensionality reduction to describe the herds in a two-dimensional space (Kohonen, 2001). A hierarchical clustering algorithm was then applied to create groups of herds with similar welfare status based on the measured proAction animal response records. Production and performance indicators were compared between welfare clusters.

6.6.1. Dataset

Five animal-based welfare outcome measures were assessed on 4,730 Quebec dairy herds between February 2015 and December 2019 by 31 independent technicians as part of the Animal Care module of the proAction® Quality Assurance Program (DFC, 2021). A detailed description of the assessment protocol is described in the proAction® reference manual (DFC, 2019). In short, a random sample of animals (from 8 to 55 cows) from the lactating herd was drawn on each herd based on the average number of lactating cows. Next, body condition score (Vasseur et al., 2013), hock, knee, and neck lesions (Gibbons et al., 2012), and lameness (Flower and Weary, 2006, Gibbons et al., 2014) were assessed on the sampled animals. Animals were then classified into "Acceptable" or requiring "Corrective Action" following a grading scale specific for each welfare outcome (DFC, 2019). The prevalence of animals classified as requiring "Corrective Action" in the sample was used as a proxy to the herd prevalence of each welfare outcome.

The dairy herd improvement (**DHI**) indicators of milk production (kg), fat production (kg), protein production (kg), milk value ($CAD), and length of productive life (**LPL**; years) were provided by Lactanet (Sainte-Anne-de-Bellevue, Quebec, Canada). Production indicators and milk value were provided for 209,749 animals while LPL was provided for 64,041 animals both from 3,242 herds. In addition, Lactanet provided the 13 pre-recorded DHI indicators of longevity, nutrition, production, profitability, young stock, and reproduction required to calculate the herd status index (**HSI**), which is a composite index to remotely assess the herd welfare status (Warner et al., 2020). These indicators were provided for 32,943 animals from 3,240 herds All data were extracted from a 12-month period prior to the assessment of the welfare outcome measures (between December 2015 and December 2019).

6.6.2. Data handling and cleaning

Herds with two proAction® assessments done on the same day (i.e., duplicated observations) were removed (N = 26) as were the data from herds involved in a pilot study (N = 27). Next, herds with negative values in the prevalence on any of the welfare outcomes were excluded (N = 5). Lastly, herds with tie-stall and free-stall housing types were kept for further analyses, as they are the most prevalent in the study region, excluding those with different housing types (N = 433).

Cleaning the DHI and calculation of the HSI was done as outlined by Warner et al. (2020). In short, the indicators were standardized to percentile ranks and transformed to their additive inverse, except for percentage of cows at third or greater lactation, herd management index, and the transition cow index, in which higher values are desirable. Lastly, the index was calculated by summing all indicators with non-missing values and dividing it by the total number of indicators with non-missing values. Therefore, HSI values could ranges from 0 to 1, in which the closer to 1, the better the overall herd status(Warner et al., 2020).

Herd average LPL as well as cumulative production and milk value were calculated based on animal level data. Animals with negative LPL were removed (N = 33). Next, animals with either missing (N = 12,170) or zero-value (N = 345) observations on production and milk value indicators were excluded. For animals that moved between herds, both productions and milk value were associated with the herd in which the cow ended each of her lactations for the calculation of the cumulative milk, fat, and protein productions and cumulative milk value. Cumulative energy-corrected milk (**ECM**; kg) was then calculated as ECM (cumulative kg) = 12.55 × fat (cumulative kg) + 7.39 × protein (cumulative kg) + 0.2595 × milk yield (cumulative kg). Lastly, herd cumulative averages were calculated for LPL, ECM, and milk value. The percentage of animals on third or greater lactations (**3+ lactation**), which was one of the indicators used in the calculation of the HSI, was also used separately as a measure of longevity in this study.

The final data submitted for analysis was created by merging both welfare outcome measures and DHI indicators based on herd number. It contained data measured on 2,980 herds which were assessed for welfare outcomes between December 2016 and December 2019.

6.6.3. Data analysis

6.6.3.1. Self-organizing map

A two-layered SOM was created based on descriptive variables and welfare indicators using the function *supersom* from the R package *kohonen* (Wehrens and Buydens, 2007, Wehrens and Kruisselbrink, 2018). The first layer comprised the barn type (tie-stall or free-stall), year (2016, 2017, 2018, or 2019), and season (winter, spring, summer, or fall) in which the welfare outcome measures were collected, while the second layer comprised the five welfare outcome measures and the HSI. This was done to account for the additional sources of variation other than the welfare outcomes in the construction of the SOM vector codebook later used for clustering analysis. Barn type was one-hot encoded, and season was cyclic encoded to account for its cyclic pattern. Next, variables were scaled to mean = 0 and standard deviation = 1. The number of units in the map was set to 260 according to the equation $5\times\sqrt{n}$, in which *n* represents the number of observations, and were organized in a rectangular grid (Kohonen, 2001).

The Euclidean distance measure was used, and different weights were assigned to each SOM layer during the training to ensure equal contribution of all layers to the final map. Even though the *supersom* function applies internal weights by default to avoid some layers to overwhelm other, it was not sufficient in this case. A Bayesian optimization, with expected improvement acquisition function, was used to establish the optimum weights given to each layer of the SOM using the function *bayesOpt* from the R package *ParBayesianOptimization* (Wilson, 2021). The following metric was used in the optimization to be maximized: $-1 \times \sqrt{\frac{\sum_{i=1}^{n}(\bar{L} - L_i)^2}{n}}$, in which $\bar{L}$ represents the average quantization error between both layers; Li represents the quantization error in the ith layer, and n represents the number of layers. A combination of 50 weight values ranging from one to 20 were drawn from a uniform distribution for evaluation and used as the initial sampling. In this step, 10,000 iterations were allowed for a fast learning of the SOMs (Kohonen, 2001). The optimization algorithm ran for 50 iterations, which was sufficient since the utility values of the points approached zero (Wilson, 2021). Once the optimum set of weights were identified, the final SOM was trained for 130,000 iterations according to the equation $500\times\sqrt{unit}$, in which *unit* represents the number of units in the map (Kohonen, 2001).

6.6.3.2. Clustering

Cluster analysis was performed to identify groups of herds with similar welfare profiles. The codebook vector (i.e., the vector of outcome averages that was mapped to each unit in the map) from the SOM layer with the welfare indicators was extracted and used to calculate a dissimilarity matrix using the Euclidean distance. The matrix was multiplied by the distances between the grid units in the SOM to account for its topographic structure. Cluster quality was evaluated using the partitioning around medoids (**PAM**), hierarchical with Ward's minimum variance linkage, and normal mixture model-based clustering procedures. First, cluster stability was assessed based on bootstrapped Jaccard mean distance obtained after 100 resamples and with cluster number varying from two to seven (Supplementary Table 6.1). This was done using the function *clusterboot* from the R Package *fpc* (Hennig, 2020). Next, overall stable clusters were subjected to internal quality assessment analysis (Supplementary Table 6.2) using the functions *cluster.stats* and *clusterbenchstats* both from the package *fpc* as well (Hennig, 2020). The identified best cluster algorithm and cluster number were used for further analysis.

6.6.3.3. Inferential analysis

The welfare clusters were described through a machine learning approach. The machine-learning algorithms recursive partitioning and regression tree (**RPART**), gradient boosting machine (**GBM**), extreme gradient boosting machine (**XGBM**), random forest (**RF**), and support vector machine (**SVM**) with a radial basis kernel were trained in this study. The welfare indicators data (i.e., the five welfare outcome measures and the HSI) were randomly split into training and validation data sets following an 80:20 ratio, respectively, stratified such to ensure an equal split among clusters. All models were trained with 10-fold cross-validation on the training data set using the *caret* package (Kuhn, 2020) by specifying the methods *rpart*, *gbm*, *xgbTree*, *ranger*, and *svmRadial* respectively for the algorithms RPART, GBM, XGBM, RF and SVM. Hyperparameters for these models were tuned using adaptive resampling, which resample the hyperparameter tuning grid by concentrating on values closer to the identified optimal settings (Kuhn, 2014, 2020). The best model was selected based on overall accuracy calculated on the validation data set (Supplementary Table 6.3) and used to describe the welfare clusters.

The influence of each welfare indicator on the different welfare clusters was analysed using partial dependence plots (**PDP**), which indicates the marginal effect of the welfare indicator on each cluster prediction by the model (Friedman, 2001). It depicts the probability shape (i.e., linear, monotonic, or more complex) of predicting a given cluster depending on the welfare indicator value (Molnar, 2019). The PDPs were calculated with the function *FeatureEffect* from the R package *iml* (Molnar et al., 2018) and using the complete welfare indicators data set (i.e., training and validation data set combined).

The relationship between welfare clusters and herd production and performance indicators was analyzed using a similar methodology as used to describe the clusters. The production and performance data (i.e., ECM, milk value, LPL, and 3+ lactation) were split into training and validation data sets following an 80:20 ratio, respectively, stratified such to ensure an equal split among clusters. The RPART, GBM, XGBM, RF, and SVM models were trained and evaluated using a similar methodology as to describe earlier. The accuracy of the models was relatively low, which could had been caused by the imbalanced distribution of cluster labels. Therefore, weighted classification, synthetic minority over-sampling technique (SMOTE) (Chawla et al., 2002), and up-sampling were also tested. The multiplicative inverse of cluster frequencies was used as weights. The best model was selected based on overall accuracy after excluding models in which some cluster labels were not predicted (Supplementary Table 6.4). This model was used further to analyse the relationship between welfare clusters and both production and performance indicators. Accumulated local effects (**ALE**) plots were used to explain the model predictions using the complete production and performance data set (i.e., training and validation data set combined) and were also calculated with the function *FeatureEffect* from the R package *iml* (Molnar et al., 2018). The ALE plots describe the average influence of the variables in the cluster prediction and its use is more appropriate when variables are highly correlated (Molnar, 2019), which was observed in this study between the production and performance indicators (Supplementary Table 6.5).

6.7. Data availability

Dairy producers consented for the use of their data for the research purposes of this study, but the data may not be shared without their consent.

6.8. Code availability

The R code that supports the findings of this study can be found in the following public GitHub repository: https://github.com/CowLifeMcGill/proAction_Sustainability.

6.9. References

Adriaens, I., N. C. Friggens, W. Ouweltjes, H. Scott, B. Aernouts, and J. Statham. 2020. Productive life span and resilience rank can be predicted from on-farm first-parity sensor time series but not using a common equation across farms. Journal of Dairy Science 103(8):7155-7171. https://doi.org/10.3168/jds.2019-17826.

Akhanli, S. E. and C. Hennig. 2020. Comparing clusterings and numbers of clusters by aggregation of calibrated clustering validity indexes. Statistics and Computing 30(5):1523-1544. https://doi.org/10.1007/s11222-020-09958-2.

Bedere, N., E. Cutullic, L. Delaby, F. Garcia-Launay, and C. Disenhaus. 2018. Meta-analysis of the relationships between reproduction, milk yield and body condition score in dairy cows. Livestock Science 210:73-84. https://doi.org/10.1016/j.livsci.2018.01.017.

Berry, D. P., M. L. Bermingham, M. Good, and S. J. More. 2011. Genetics of animal health and disease in cattle. Irish Veterinary Journal 64(1):5. https://doi.org/10.1186/2046-0481-64-5.

Bicalho, R. C., V. S. Machado, and L. S. Caixeta. 2009. Lameness in dairy cattle: A debilitating disease or a disease of debilitated cattle? A cross-sectional study of lameness prevalence and thickness of the digital cushion. Journal of Dairy Science 92(7):3175-3184. https://doi.org/10.3168/jds.2008-1827.

Blanc, F., F. Bocquier, J. Agabriel, P. D'hour, and Y. Chilliard. 2006. Adaptive abilities of the females and sustainability of ruminant livestock systems. A review. Animal Research 55(6):489-510. https://doi.org/10.1051 / animres: 2006040.

Boogaard, B. K., S. J. Oosting, B. B. Bock, and J. S. C. Wiskerke. 2011. The sociocultural sustainability of livestock farming: An inquiry into social perceptions of dairy farming. Animal 5(9):1458-1466. https://doi.org/10.1017/S1751731111000371.

Bouffard, V., A. M. de Passillé, J. Rushen, E. Vasseur, C. G. R. Nash, D. B. Haley, and D. Pellerin. 2017. Effect of following recommendations for tiestall configuration on neck and leg lesions, lameness, cleanliness, and lying time in dairy cows. Journal of Dairy Science 100(4):2935-2943. https://doi.org/10.3168/jds.2016-11842.

Broom, D. M. 1991. Animal welfare: concepts and measurement. Journal of Animal Science 69(10):4167-4175. https://doi.org/10.2527/1991.69104167x.

Chawla, N. V., K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. 2002. SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research 16:321–357. https://doi.org/10.1613/jair.953.

Dallago, G. M., K. M. Wade, R. I. Cue, J. T. McClure, R. Lacroix, D. Pellerin, and E. Vasseur. 2021. Keeping dairy cows for longer: A critical literature review on dairy cow longevity in high milk-producing countries. Animals 11(3):808. https://doi.org/10.3390/ani11030808.

Delgado, H. A., R. I. Cue, D. Haine, A. Sewalem, R. Lacroix, D. Lefebvre, J. Dubuc, E. Bouchard, and K. M. Wade. 2017. Profitability measures as decision-making tools for Québec dairy herds. Canadian Journal of Animal Science 98(1):18-31. https://doi.org/10.1139/cjas-2016-0202.

DFC – Dairy Farmers of Canada. 2019. proAction®: Reference manual. Accessed on 13 September 2021. https://www.dairyfarmers.ca/proaction/resources/overview.

DFC – Dairy Farmers of Canada. 2021. proAction®: On-farm excellence. Accessed on 14 June 2021. https://www.dairyfarmers.ca/proaction.

Espejo, L. A., M. I. Endres, and J. A. Salfer. 2006. Prevalence of lameness in high-producing Holstein cows housed in freestall barns in Minnesota. Journal of Dairy Science 89(8):3052-3058. https://doi.org/10.3168/jds.S0022-0302(06)72579-6.

FARM – Farmers Assuring Responsible Management. 2020. Animal care reference manual version 4. Accessed on 05 November 2021. https://nationaldairyfarm.com/farm-animal-care-version-4-0/.

Flower, F. C. and D. M. Weary. 2006. Effect of hoof pathologies on subjective assessments of dairy cow gait. Journal of Dairy Science 89(1):139-146. https://doi.org/10.3168/jds.S0022-0302(06)72077-X.

Fraser, D. 2008. Understanding animal welfare. Acta Veterinaria Scandinavica 50(1):S1. https://doi.org/10.1186/1751-0147-50-S1-S1.

Fregonesi, J. A., C. B. Tucker, and D. M. Weary. 2007. Overstocking reduces lying time in dairy cows. Journal of Dairy Science 90(7):3349-3354. https://doi.org/10.3168/jds.2006-794.

Friedman, J. H. 2001. Greedy function pproximation: A gradient boosting machine. The Annals of Statistics 29(5):1189-1232. https://doi.org/10.1214/aos/1013203451.

Galloway, C., B. Conradie, H. Prozesky, and K. Esler. 2018. Opportunities to improve sustainability on commercial pasture-based dairy farms by assessing environmental impact. Agricultural Systems 166:1-9. https://doi.org/10.1016/j.agsy.2018.07.008.

Gibbons, J., D. B. Haley, J. Higginson Cutler, C. Nash, J. Zaffino Heyerhoff, D. Pellerin, S. Adam, A. Fournier, A. M. de Passillé, J. Rushen, and E. Vasseur. 2014. Technical note: A comparison of 2 methods of assessing lameness prevalence in tiestall herds. Journal of Dairy Science 97(1):350-353. https://doi.org/10.3168/jds.2013-6783.

Gibbons, J., E. Vasseur, J. Rushen, and A. M. de Passillé. 2012. A training programme to ensure high repeatability of injury scoring of dairy cows. Animal Welfare 21(3):379-388. https://doi.org/10.7120/09627286.21.3.379.

Godde, C. M., D. Mason-D'Croz, D. E. Mayberry, P. K. Thornton, and M. Herrero. 2021. Impacts of climate change on the livestock food supply chain; a review of the evidence. Global Food Security 28:100488. https://doi.org/10.1016/j.gfs.2020.100488.

Godfray, H. C. J., J. R. Beddington, I. R. Crute, L. Haddad, D. Lawrence, J. F. Muir, J. Pretty, S. Robinson, S. M. Thomas, and C. Toulmin. 2010. Food security: The challenge of feeding 9 billion people. Science 327(5967):812-818. https://doi.org/10.1126/science.1185383.

Grandl, F., M. Furger, M. Kreuzer, and M. Zehetmeier. 2019. Impact of longevity on greenhouse gas emissions and profitability of individual dairy cows analysed with different system boundaries. Animal 13(1):198-208. https://doi.org/10.1017/S175173111800112X.

Hennig, C. 2020. fpc: Flexible procedures for clustering.Version 2.2-9. https://CRAN.R-project.org/package=fpc.

Herrero, M., D. Grace, J. Njuki, N. Johnson, D. Enahoro, S. Silvestri, and M. C. Rufino. 2013. The roles of livestock in developing countries. Animal 7:3-18. https://doi.org/10.1017/S1751731112001954.

Herzog, A., C. Winckler, and W. Zollitsch. 2018. In pursuit of sustainability in dairy farming: A review of interdependent effects of animal welfare improvement and environmental impact mitigation. Agriculture, Ecosystems & Environment 267:174-187. https://doi.org/10.1016/j.agee.2018.07.029.

Hoischen-Taubner, S., J. Habel, V. Uhlig, E.-M. Schwabenbauer, T. Rumphorst, L. Ebert, D. Möller, and A. Sundrum. 2021. The whole and the parts – A new perspective on production diseases and economic sustainability in dairy farming. Sustainability 13(16):9044. https://doi.org/10.3390/su13169044.

Jensen, M. B., L. J. Pedersen, and L. Munksgaard. 2005. The effect of reward duration on demand functions for rest in dairy heifers and lying requirements as measured by demand functions. Applied Animal Behaviour Science 90(3–4):207-217. https://doi.org/10.1016/j.applanim.2004.08.006.

Jewell, M. T., M. Cameron, J. Spears, S. L. McKenna, M. S. Cockram, J. Sanchez, and G. P. Keefe. 2019. Prevalence of lameness and associated risk factors on dairy farms in the Maritime Provinces of Canada. Journal of Dairy Science 102(4):3392-3405. https://doi.org/10.3168/jds.2018-15349.

Kielland, C., L. E. Ruud, A. J. Zanella, and O. Østerås. 2009. Prevalence and risk factors for skin lesions on legs of dairy cattle housed in freestalls in Norway. Journal of Dairy Science 92(11):5487-5496. https://doi.org/10.3168/jds.2009-2293.

Knapp, J. R., G. L. Laur, P. A. Vadas, W. P. Weiss, and J. M. Tricarico. 2014. Invited review: Enteric methane in dairy cattle production: Quantifying the opportunities and impact of reducing emissions. Journal of Dairy Science 97(6):3231-3261. https://doi.org/10.3168/jds.2013-7234.

Knierim, U. and C. Winckler. 2009. On-farm welfare assessment in cattle: Validity, reliability and feasibility issues and future perspectives with special regard to the Welfare Quality® approach. Animal Welfare 18(4):451-458.

Kohonen, T. 2001. Self-Organizing Maps. Vol. 30. 3rd ed. Springer Series in Information Sciences. Springer, Berlin/Heidelberg, Germany. https://doi.org/10.1007/978-3-642-56927-2.

Kuhn, M. 2014. Futility analysis in the cross-validation of machine learning models. arXiv:1405.6974v1.

Kuhn, M. 2020. caret: Classification and regression training.Version 6.0-86. https://CRAN.R-project.org/package=caret.

Lebacq, T., P. V. Baret, and D. Stilmant. 2013. Sustainability indicators for livestock farming: A review. Agronomy for Sustainable Development 33(2):311-327. https://doi.org/10.1007/s13593-012-0121-x.

Lebacq, T., P. V. Baret, and D. Stilmant. 2015. Role of input self-sufficiency in the economic and environmental sustainability of specialised dairy farms. Animal 9(3):544-552. https://doi.org/10.1017/S1751731114002845.

Lovarelli, D., J. Bacenetti, and M. Guarino. 2020. A review on dairy cattle farming: Is precision livestock farming the compromise for an environmental, economic and social sustainable production? Journal of Cleaner Production 262:121409. https://doi.org/10.1016/j.jclepro.2020.121409.

McPherson, S. E. and E. Vasseur. 2020. Graduate student literature review: The effects of bedding, stall length, and manger wall height on common outcome measures of dairy cow welfare in stall-based housing systems. Journal of Dairy Science 103(11):10940-10950. https://doi.org/10.3168/jds.2020-18332.

Molnar, C. 2019. Interpretable Machine Learning. Christoph Molnar.

Molnar, C., G. Casalicchio, and B. Bischl. 2018. iml: An R package for interpretable machine learning. Journal of Open Source Software 3(26):786. https://doi.org/10.21105/joss.00786.

Nash, C. G. R., D. F. Kelton, T. J. DeVries, E. Vasseur, J. Coe, J. C. Z. Heyerhoff, V. Bouffard, D. Pellerin, J. Rushen, A. M. de Passillé, and D. B. Haley. 2016. Prevalence of and risk factors for hock and knee injuries on dairy cows in tiestall housing in Canada. Journal of Dairy Science 99(8):6494-6506. https://doi.org/10.3168/jds.2015-10676.

Norring, M., J. Häggman, H. Simojoki, P. Tamminen, C. Winckler, and M. Pastell. 2014. Short communication: Lameness impairs feeding behavior of dairy cows. Journal of Dairy Science 97(7):4317-4321. https://doi.org/10.3168/jds.2013-7512.

Oltenacu, P. A. and B. Algers. 2005. Selection for increased production and the welfare of dairy cows: Are new breeding goals needed? AMBIO: A Journal of the Human Environment 34(4):311-315. https://doi.org/10.1579/0044-7447-34.4.311.

Oltenacu, P. A. and D. M. Broom. 2010. The impact of genetic selection for increased milk yield on the welfare of dairy cows. Animal Welfare 19(2):39-49.

Österman, S. and I. Redbo. 2001. Effects of milking frequency on lying down and getting up behaviour in dairy cows. Applied Animal Behaviour Science 70(3):167-176. https://doi.org/10.1016/S0168-1591(00)00159-3.

Poore, J. and T. Nemecek. 2018. Reducing food's environmental impacts through producers and consumers. Science 360(6392):987. https://doi.org/10.1126/science.aaq0216.

Popescu, S., C. Borda, E. A. Diugan, M. Niculae, R. Stefan, and C. D. Sandru. 2014. The effect of the housing system on the welfare quality of dairy cows. Italian Journal of Animal Science 13(1):2940. https://doi.org/10.4081/ijas.2014.2940.

Pretty, J. and Z. P. Bharucha. 2019. Sustainable livestock and animal-sourced food. Pages 225-232 in Sustainable food and agriculture: An integrated approach. C. Campanhola and S. Pandey, ed. The Food and Agriculture Organization of the Uniteed Nations & Elsevier Inc. https://doi.org/10.1016/B978-0-12-812134-4.00022-4.

Puerto, M. A., E. Shepley, R. I. Cue, D. Warner, J. Dubuc, and E. Vasseur. 2021a. The hidden cost of disease: I. Impact of the first incidence of mastitis on production and economic indicators of primiparous dairy cows. Journal of Dairy Science 104(7):7932-7943. https://doi.org/10.3168/jds.2020-19584.

Puerto, M. A., E. Shepley, R. I. Cue, D. Warner, J. Dubuc, and E. Vasseur. 2021b. The hidden cost of disease: II. Impact of the first incidence of lameness on production and economic indicators of primiparous dairy cows. Journal of Dairy Science 104(7):7944-7955. https://doi.org/10.3168/jds.2020-19585.

R Core Team. 2021. R: A language and environment for statistical computing.Version 4.1.0 "Camp Pontanezen". Accessed 2021-05-18. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/.

Randall, L. V., M. J. Green, M. G. G. Chagunda, C. Mason, S. C. Archer, L. E. Green, and J. N. Huxley. 2015. Low body condition predisposes cattle to lameness: An 8-year study of one dairy herd. Journal of Dairy Science 98(6):3766-3777. https://doi.org/10.3168/jds.2014-8863.

Roche, J. R., N. C. Friggens, J. K. Kay, M. W. Fisher, K. J. Stafford, and D. P. Berry. 2009. Invited review: Body condition score and its association with dairy cow productivity, health, and welfare. Journal of Dairy Science 92(12):5769-5801. https://doi.org/10.3168/jds.2009-2431.

Rushen, J., E. Pombourcq, and A. M. d. Passillé. 2007. Validation of two measures of lameness in dairy cows. Applied Animal Behaviour Science 106(1–3):173-177. https://doi.org/10.1016/j.applanim.2006.07.001.

Rutherford, K. M. D., F. M. Langford, M. C. Jack, L. Sherwood, A. B. Lawrence, and M. J. Haskell. 2009. Lameness prevalence and risk factors in organic and non-organic dairy herds in the United Kingdom. The Veterinary Journal 180(1):95-105. https://doi.org/10.1016/j.tvjl.2008.03.015.

Segerkvist, K. A., H. Hansson, U. Sonesson, and S. Gunnarsson. 2020. Research on environmental, economic, and social sustainability in dairy farming: A systematic mapping of current literature. Sustainability 12(14):5502. https://doi.org/10.3390/su12145502.

Solano, L., H. W. Barkema, E. A. Pajor, S. Mason, S. J. LeBlanc, J. C. Zaffino Heyerhoff, C. G. R. Nash, D. B. Haley, E. Vasseur, D. Pellerin, J. Rushen, A. M. de Passillé, and K. Orsel. 2015. Prevalence of lameness and associated risk factors in Canadian Holstein-Friesian cows housed in freestall barns. Journal of Dairy Science 98(10):6978-6991. https://doi.org/10.3168/jds.2015-9652.

Somers, J. G. C. J., K. Frankena, E. N. Noordhuizen-Stassen, and J. H. M. Metz. 2003. Prevalence of claw disorders in Dutch dairy cows exposed to several floor systems. Journal of Dairy Science 86(6):2082-2093. https://doi.org/10.3168/jds.S0022-0302(03)73797-7.

Springmann, M., M. Clark, D. Mason-D'Croz, K. Wiebe, B. L. Bodirsky, L. Lassaletta, W. de Vries, S. J. Vermeulen, M. Herrero, K. M. Carlson, M. Jonell, M. Troell, F. DeClerck, L. J. Gordon, R. Zurayk, P. Scarborough, M. Rayner, B. Loken, J. Fanzo, H. C. J. Godfray, D. Tilman, J. Rockström, and W. Willett. 2018. Options for keeping the food system within environmental limits. Nature 562(7728):519-525. https://doi.org/10.1038/s41586-018-0594-0.

Stehfest, E., L. Bouwman, D. P. van Vuuren, M. G. J. den Elzen, B. Eickhout, and P. Kabat. 2009. Climate benefits of changing diet. Climatic Change 95(1):83-102. https://doi.org/10.1007/s10584-008-9534-6.

Tilman, D. and M. Clark. 2014. Global diets link environmental sustainability and human health. Nature 515(7528):518-522. https://doi.org/10.1038/nature13959.

Tricarico, J. M., E. Kebreab, and M. A. Wattiaux. 2020. MILK Symposium review: Sustainability of dairy production and consumption in low-income countries with emphasis on productivity and environmental impact*. Journal of Dairy Science 103(11):9791-9802. https://doi.org/10.3168/jds.2020-18269.

VandeHaar, M. J., L. E. Armentano, K. Weigel, D. M. Spurlock, R. J. Tempelman, and R. Veerkamp. 2016. Harnessing the genetics of the modern dairy cow to continue improvements in feed efficiency1. Journal of Dairy Science 99(6):4941-4954. https://doi.org/10.3168/jds.2015-10352.

Vanegas, J., M. Overton, S. L. Berry, and W. M. Sischo. 2006. Effect of rubber flooring on claw health in lactating dairy cows housed in free-stall barns. Journal of Dairy Science 89(11):4251-4258. https://doi.org/10.3168/jds.S0022-0302(06)72471-7.

Vasseur, E., J. Gibbons, J. Rushen, and A. M. de Passillé. 2013. Development and implementation of a training program to ensure high repeatability of body condition scoring of dairy cows. Journal of Dairy Science 96(7):4725-4737. https://doi.org/10.3168/jds.2012-6359.

Villettaz Robichaud, M., J. Rushen, A. M. de Passillé, E. Vasseur, D. Haley, and D. Pellerin. 2019a. Associations between on-farm cow welfare indicators and productivity and profitability on Canadian dairies: II. On tiestall farms. Journal of Dairy Science 102(5):4352-4363. https://doi.org/10.3168/jds.2018-14818.

Villettaz Robichaud, M., J. Rushen, A. M. de Passillé, E. Vasseur, K. Orsel, and D. Pellerin. 2019b. Associations between on-farm animal welfare indicators and productivity and profitability on Canadian dairies: I. On freestall farms. Journal of Dairy Science 102(5):4341-4351. https://doi.org/10.3168/jds.2018-14817.

Warner, D., E. Vasseur, M. Villettaz Robichaud, S. Adam, D. Pellerin, D. M. Lefebvre, and R. Lacroix. 2020. Development of a benchmarking tool for dairy herd management using routinely collected herd records. Animals 10(9):1689. https://doi.org/10.3390/ani10091689.

Wehrens, R. and L. M. C. Buydens. 2007. The kohonen package. Journal of Statistical Software 21(5):1-19. https://doi.org/10.18637/jss.v021.i05.

Wehrens, R. and J. Kruisselbrink. 2018. Flexible self-organizing maps in kohonen 3.0. Journal of Statistical Software 87(7):1-18. https://doi.org/10.18637/jss.v087.i07.

Whay, H. R., A. E. Waterman, A. J. F. Webster, and J. K. O'Brien. 1998. The influence of lesion type on the duration ofhyperalgesia associated with hindlimb lameness in dairy cattle. The Veterinary Journal 156(1):23-29. https://doi.org/10.1016/S1090-0233(98)80058-0.

Wilson, S. 2021. ParBayesianOptimization: Parallel Bayesian Optimization of Hyperparameters.Version 1.2.5. https://github.com/AnotherSamWilson/ParBayesianOptimization.

6.10. Supplementary material

Supplementary Table 6.1. Bootstrap stability assessment of up to 7 clusters obtained after 100 resampling runs of the clustering algorithms partitioning around medoids (PAM), hierarchical with Ward's minimum variance method, and normal mixture model-based. Overall and cluster-wise stable results are bolded.

Clustering algorithm	Jaccard bootstrap mean distance[1]							
	Overall	Cluster						
		1	2	3	4	5	6	7
PAM	**0.88**	**0.84**	**0.91**					
	0.84	**0.80**	**0.84**	**0.88**				
	0.70	**0.78**	**0.75**	0.66	0.61			
	0.59	0.66	0.70	0.47	0.60	0.51		
	0.58	0.62	0.58	0.49	0.52	0.60	0.65	
	0.54	0.54	0.63	0.48	0.44	0.52	0.57	0.63
Hierarchical	**0.80**	**0.78**	**0.81**					
	0.67	0.63	0.58	**0.80**				
	0.58	0.62	0.56	0.56	0.59			
	0.58	0.56	0.57	0.52	0.60	0.63		
	0.60	0.66	0.45	0.70	0.55	0.60	0.65	
	0.57	0.67	0.40	0.74	0.36	0.54	0.59	0.66
Normal mixture model-based	0.59	0.50	0.68					
	0.49	0.55	0.48	0.45				
	0.44	0.53	0.48	0.40	0.33			
	0.35	0.43	0.39	0.23	0.32	0.36		
	0.43	0.52	0.33	0.43	0.51	0.33	0.45	
	0.43	0.54	0.38	0.38	0.39	0.45	0.33	0.52

[1] Below 0.60: cluster should not be trusted; between 0.60 and 0.75: indication of patterns in the data, but cluster membership is doubtful; 0.75 or more: cluster is stable; 0.85 an above: cluster is highly stable (Hennig, 2020)

Supplementary Table 6.2. Internal quality assessment indexes of different clustering methods considered stable based on overall mean bootstrapped Jaccard distance. Best results are bolded within rows.

Index	Clustering method[1]		
	PAM2	PAM3	Hierarchical2
Calinski–Harabasz index	356	360	**284**
Average silhouette width[2]	**1.34**	1.05	0.81
Maximum cluster diameter [2]	0.57	**2.0**	0.38
Within cluster sum of squares[2]	0.68	**1.4**	0.45
Average distance to cluster centroid[2]	0.55	**1.7**	0.09
Average within-cluster dissimilarities[2]	0.50	**1.4**	0.35
PearsonΓ[2]	**1.32**	1.03	0.57
Bootstab index[2]	1.90	**2.10**	**2.10**
Separation index[2]	**-0.09**	-0.40	-0.26
Widest within-cluster gap[2]	**0.48**	0.44	**0.48**
A_1 [2,3]	1.2	**1.5**	1.0
A_2 [2,4]	**0.76**	0.72	0.77

[1] PAM2 = Two clusters based on partitioning around medoids clustering algorithm; PAM3 = Three clusters based on partitioning around medoids clustering algorithm; and Hierarchical2 = Two cluster based on hierarchical clustering algorithm with Ward's minimum variance method.

[2] Index values are calibrated relative to a set of random clusterings to allow comparisons between different number and clustering methods (Akhanli and Hennig, 2020).

[3] A_1 = Calibrated composite index indicating cluster homogeneity by combining the average within-cluster dissimilarities, the PearsonΓ, and the Bootstab indexes (Akhanli and Hennig, 2020).

[4] A_2 = Calibrated composite index indicating cluster separation by combining the separation index, the widest within-cluster gap, and the Bootstab indexes (Akhanli and Hennig, 2020).

Supplementary Table 6.3. Overall and by cluster performance results of recursive partitioning and regression tree (RPART) gradient boosting machine (GBM), extreme gradient boosting machine (XGBM), random forest (RF), and support vector machine (SVM) models in predicting cluster labels on the validation data set of welfare indicators. Best results are bolded within columns for overall statistics and within columns by cluster label for statistics by cluster.

Model	Overall Statistics		Statistics by cluster				
	Accuracy	**95% CI**	**Cluster**	**Sensitivity**	**Specificity**	**F1**	**Balanced accuracy**
RPART	0.78	0.74 – 0.81	1	0.71	0.91	0.71	0.81
			2	0.67	0.93	0.69	0.80
			3	0.85	0.78	0.84	0.82
GBM	0.83	0.80 – 0.86	1	0.76	0.94	**0.79**	**0.85**
			2	0.74	**0.95**	0.76	0.84
			3	0.90	0.81	0.88	0.85
XGBM	0.84	0.81 – 0.87	1	0.74	**0.96**	**0.79**	**0.85**
			2	0.73	**0.95**	0.76	0.84
			3	**0.92**	0.79	0.88	0.86
RF	0.84	0.81 – 0.87	1	0.72	0.95	0.77	0.84
			2	0.80	0.94	0.79	0.87
			3	0.91	**0.83**	**0.89**	**0.87**
SVM	**0.85**	0.81 – 0.87	1	**0.75**	0.95	0.78	**0.85**
			2	**0.80**	**0.95**	**0.81**	**0.88**
			3	0.90	**0.83**	0.88	**0.87**

Supplementary Table 6.4. Overall and by cluster performance results of recursive partitioning and regression tree (RPART) gradient boosting machine (GBM), extreme gradient boosting machine (XGBM), random forest (RF), and support vector machine (SVM) models in predicting cluster labels on the validation data set of production and performance indicators. Weighted classification, synthetic minority over-sampling technique (SMOTE), and up-sampling strategies were also employed because of the imbalanced distribution of cluster labels. The model with the overall best accuracy is bolded.

Model	Overall Statistics		Statistics by cluster				
	Accuracy	95% CI	Cluster	Sensitivity	Specificity	F1	Balanced accuracy
			Regular data				
RPART	0.56	0.52 – 0.60	1	0.00	1.00	---	0.50
			2	0.00	1.00	---	0.50
			3	1.00	0.00	0.71	0.50
GBM	0.56	0.52 – 0.60	1	0.01	1.00	0.01	0.50
			2	0.00	1.00	---	0.50
			3	1.00	0.01	0.72	0.50
XGBM	0.56	0.52 – 0.60	1	0.01	1.00	0.01	0.50
			2	0.00	1.00	---	0.50
			3	1.00	0.01	0.72	0.50
RF	0.52	0.48 – 0.56	1	0.14	0.92	0.20	0.53
			2	0.05	0.93	0.07	0.49
			3	0.85	0.18	0.68	0.51
SVM	0.56	0.52 – 0.60	1	0.00	1.00	---	0.50
			2	0.00	1.00	---	0.50
			3	1.00	0.00	0.71	0.50
			Weighted				
RPART	0.32	0.28 – 0.36	1	0.42	0.69	0.35	0.56
			2	0.61	0.52	0.35	0.57
			3	0.17	0.87	0.27	0.52
GBM	0.37	0.33 – 0.41	1	0.30	0.73	0.28	0.52
			2	0.28	0.71	0.23	0.50
			3	0.43	0.55	0.48	0.49
XGBM	**0.44**	**0.40 – 0.48**	1	0.26	0.78	0.27	0.52
			2	0.26	0.81	0.25	0.53
			3	0.58	0.46	0.58	0.52
RF	0.42	0.38 – 0.46	1	0.27	0.77	0.27	0.52
			2	0.18	0.80	0.19	0.49
			3	0.57	0.43	0.56	0.50
SVM	---	---	1	---	---	---	---
			2	---	---	---	---
			3	---	---	---	---
			Synthetic minority over sampling technique (SMOTE)				
RPART	0.26	0.23 – 0.30	1	0.58	0.46	0.35	0.52
			2	0.41	0.63	0.29	0.52
			3	0.08	0.93	0.13	0.50
GBM	0.35	0.31 – 0.39	1	0.35	0.70	0.30	0.52
			2	0.34	0.68	0.26	0.51
			3	0.35	0.63	0.43	0.49
XGBM	0.35	0.31 – 0.38	1	0.26	0.72	0.24	0.49
			2	0.30	0.67	0.23	0.49
			3	0.40	0.60	0.47	0.50
RF	0.36	0.32 – 0.40	1	0.33	0.68	0.29	0.51
			2	0.34	0.68	0.26	0.51
			3	0.38	0.67	0.47	0.53
SVM	0.36	0.32 – 0.39	1	0.42	0.66	0.34	0.54
			2	0.44	0.66	0.31	0.55

			3	0.30	0.74	0.40	0.52
			Up Sampling				
RPART	0.23	0.20 – 0.27	1	0.26	0.81	0.28	0.53
			2	0.84	0.22	0.34	0.53
			3	0.00	1.00	---	0.50
GBM	0.39	0.35 – 0.43	1	0.27	0.74	0.26	0.50
			2	0.29	0.74	0.25	0.51
			3	0.48	0.55	0.52	0.52
XGBM	0.40	0.36 – 0.44	1	0.23	0.76	0.23	0.50
			2	0.22	0.76	0.21	0.49
			3	0.53	0.47	0.54	0.50
RF	0.45	0.41 – 0.49	1	0.21	0.83	0.24	0.52
			2	0.12	0.85	0.14	0.48
			3	0.68	0.33	0.61	0.50
SVM	0.34	0.30 – 0.38	1	0.38	0.67	0.32	0.53
			2	0.43	0.63	0.30	0.53
			3	0.28	0.73	0.38	0.51

Supplementary Table 6.5. Correlation matrix of welfare, production, and performance indicators from 2, 980 Quebec – Canada dairy farms.

	BCS ≤ 2 (%)[1]	Hock lesion (%)	Knee lesion (%)	Neck lesion (%)	Lameness (%)	HSI[2]	ECM (kg)[3]	Milk value ($CAD)[4]	LPL (years)[5]
Hock lesion (%)	0.17	---	---	---	---	---	---	---	---
Knee lesion (%)	0.16	0.24	---	---	---	---	---	---	---
Neck lesion (%)	0.08	0.11	0.20	---	---	---	---	---	---
Lameness (%)	0.33	0.32	0.26	0.18	---	---	---	---	---
HSI[2]	-0.01	-0.03	-0.15	-0.03	-0.12	---	---	---	---
ECM (kg)[3]	-0.09	0.06	-0.08	-0.06	-0.01	0.20	---	---	---
Milk value ($CAD)[4]	-0.07	0.06	-0.08	-0.05	-0.02	0.19	0.93	---	---
LPL (years)[5]	0.03	-0.05	-0.01	0.04	-0.01	0.18	0.01	0.03	---
3+ lactation (%)[6]	0.03	-0.08	-0.05	0.07	-0.04	0.33	-0.15	-0.13	0.55

[1] Prevalence of cows with body condition score lower than or equal to two.
[2] Herd composite indicator for remote assessment of herd welfare status (Warner et al., 2020).
[3] Animal cumulative energy-corrected milk.
[4] Lifetime cumulative milk value.
[5] Length of productive life.
[6] Percentage of cows at third or greater lactations.

CHAPTER 7 – General discussion

Throughout this thesis, the intertwined relationships between dairy cow longevity, productivity, and profitability were analyzed not only based on a set of diverse indicators from different stages of an animal life, but also using various machine learning algorithms and analytical techniques. Lastly, an attempt was made to map the sustainability of dairy farming through an innovative animal-centered approach.

Cow longevity was a central focal point throughout this thesis. However, its very definition as an animal trait rather than a biological characteristic remains as a complex unanswered question. Contrary to milk, for example, which has a relatively straightforward definition, cow longevity is very often the consequence of a culling decision made by farmers, as discussed in Chapter 2. Unquestionably, the concept of time must be incorporated in the definition, but animal performance influences culling decision and should also be incorporated. Existing attempts going further from only considering time relied on animal stayability, i.e., the ability of the animal to avoid culling due the reproductive problems and sickness (Ducrocq et al., 1988, Ducrocq, 1994), or functional longevity, adjusting for the effect of milk yield (Sewalem et al., 2008). Both imply that cows are to blame for their own longevity (or lack thereof) when it might not be the case. For instance, the intense artificial genetic selection for milk production have had a negative effect on reproductive performance (Nebel and McGilliard, 1993, Pryce et al., 2004), though such relationship is not the consensus for longevity (Dallago et al., 2021). Therefore, the definition of dairy cow longevity should be focused on the animal outcomes as suggested in Chapter 2 but, most importantly, it should consider that, as an animal trait, it is the consequence of a decision rather than a biological event. Thus, longevity metrics should incorporate both aspects.

Possibly as a consequence of not having a standard definition, there is a disconnect between the idea of longevity and the metrics available to measure it, making it appear as though culling decisions are not objective (Beaudeau et al., 2000, Adriaens et al., 2020). This was highlighted in Chapters 4 and 5, where it was found that farmers, as decision making agents, hold an important role in cow longevity as well as in herd productivity and profitability. Production-oriented farmers were associated with shorter herd longevity but produced more sellable milk compared to longevity-oriented farmer, which in turn had a greater longevity but lower production of sellable milk. It tempts us to take correlation with causation and conclude that a herd should have shorter longevity to be profitable. If that was the case, why should we try to increase cow longevity? (would have this thesis been an unjustifiable venture?) Perhaps (hopefully) there is more to it. For instance, there is a consensus that, depending on the amount of milk a cow produces, she would not be profitable until her third lactation (Horn et al., 2012, Pellerin et al., 2014, Delgado et al., 2017). In addition, some herds in Chapter 4 were indeed associated with high longevity, productivity, and profitability. However, the results of both Chapters 4 and 5 alone do not allow us to reach a final answer about how longevity, productivity, and profitability are intertwined. What they do allow us to safely conclude, though, is on how important the role played by farmers on these outcomes is, especially on longevity.

A promising conclusion, as it indicates the feasibility of conceiving intervention strategies to improve longevity and the sustainability of the industry. In fact, the long-term effects of birth conditions on offspring longevity found in Chapter 3 could be incorporated into such strategies to identify animals more likely to reach their potential. Additional measures to maintain animals healthy and comfortable could also reduce the need for farmers to remove animals early on because of these reasons. However, other reasons such as space constraint (De Vries and Marcondes, 2020) and quota (Van Doormaal, 2009) could pressure farmers to cull cows. If that was the case, the welfare status of these animals is likely not compromised due to sickness at the very least, contributing to the social acceptability and sustainability of dairy farming. Though welfare is a core component of social sustainability from an animal perspective, large-scale collections of data on animal welfare outcomes are still scarce. Perhaps was it taken for granted, as production did not seem to be largely compromised or its importance was outweighed compared to other factors (e.g., milk yield) with a more straightforward connection with economic returns. However, the pressure for sustainable milk production has rightfully increased, and sustainability aspects should be met. As highlighted in Chapter 6, the social component seems disconnected from both the economic and the environmental components, but more precise indicators are still needed to thoroughly evaluate dairy farming sustainability from an animal centric perspective.

Machine learning algorithms were another core theme in this thesis as they were used to analyze the data in all studies (Chapter 3 to 6). Models obtained though this methodology naturally handles non-linearity and high order interactions, all while making no assumptions about the data distribution to model the relationship between input and output variables. Though they are often complex, the interpretation of the results requires shifting the focus from “statistically significant” to “statistically relevant”. As done throughout this thesis, the overall importance of the inputs can be determined based on the accuracy of the model, which can be calculated using relative metrics (i.e., how good is one algorithm compared to another?) and absolute metrics with known lower and upper limits (i.e., percentage). Additionally, we can used them to not only gather insights about the overall meaningfulness of input variables, but also to obtain the effect of each one of them without any pre-defined assumption about their shape. This analytical framework may provide an extra protection layer against statistical flukes such as misinterpretation of p-values (Gelman and Stern, 2006, Betensky, 2019), but most importantly, it forces us to interpret the overall practical implications of the results. Such a shift becomes even more relevant in the big data era. The high volume of data makes it more likely to find statistically significant results given the relationship between p-value calculations and sample size, even though they might be of little practical significance (Lantz, 2013). As demonstrated throughout this thesis, machine learning algorithms are a flexible alternative to overcome such limitations.

However, the use of machine learning models to obtain insights that are biologically meaningful is still on its infant ages. Survival analysis is one of the areas still requiring significant development. For instance, data from animals not yet culled, which accounted for about 15 to 20% of the data, were removed from the analyses carried out on Chapter 4. Animals in that study were born around the same period and such procedure may be a source of bias, since these animals may not have been culled yet because they are resilient cows, have a good production level, and/or have no reproductive problems. Survival analysis methodologies could be one alternative in cases where events of interest (i.e., longevity as well as lifetime cumulative production and profitability) have not yet been observed in all animals, which is similar to the methodology adopted in Chapter 3. However, the number of machine learning algorithms able to handle such conditions is limited as most of them requires data to be observed. In most cases, simple imputations, such as overall mean, or data exclusion are used by algorithms said to handle missing data automatically.

Regardless of analytical limitations, data on dairy production remains a significant constraint worldwide. The need for systems to centralize data storage and purposely sharing is not recent, and efforts have been made to achieve that (e.g., Dairy Brain; https://dairybrain.wisc.edu/). However, those remain as isolated examples and have limited coverage. Data collection is another significant limitation, which is also not news and was stated as a limitation in Chapter 3. Farmers are often incentivized to collect data based on the assertion that farm management can be improved once data is available. However, including additional tasks to the already overbooked daily workload of running of a dairy farm is not simple. Producers' willingness to keep data records is often determined by their perceived short-term benefits such as the case with antibiotic treatments. The milk from cows receiving antibiotics cannot the sold while she is under treatment and even for a given number of days after the treatment is finished due to residual contamination. Therefore, recording such data is clearly beneficial in the short-term. Culling reasons are another example. Usually only the last event is recorded, even though culling is most likely due to a combination of multiple factors. Reproductive status and milk yield are protective factors against culling (Hultgren and Svensson, 2009, Pinedo et al., 2010, Stojkov et al., 2020). A pregnant high producing cow is likely to be kept in the herd if she develops mastitis, whereas a low producing cow that is not pregnant has a higher chance of being culled in case of a mastitis infection. The reported reason for culling of the latter cow will be mastitis, which is an arguably incomplete indication of the true culling reason.

In the (near) future, precision livestock tools are promising to not only automate most of the data collection but, most importantly, to broaden the variables measured (Tullo et al., 2019). For instance, automated systems are necessary to monitor animal behaviors realistically and continuously in farm animals. Once developed and implemented on farms, automated systems would not only reduce costs of animal welfare assessments, but it would also allow for increasing its frequency and optimizing behavior-based health monitoring. This could be a solution to the need for better indicators identified in Chapter 6, making it possible to conceive concrete and comprehensive strategies to improve the sustainability of dairy farming while supporting farmers to optimize production efficiency.

Master reference list

Adriaens, I., N. C. Friggens, W. Ouweltjes, H. Scott, B. Aernouts, and J. Statham. 2020. Productive life span and resilience rank can be predicted from on-farm first-parity sensor time series but not using a common equation across farms. Journal of Dairy Science 103(8):7155-7171. https://doi.org/10.3168/jds.2019-17826.

Aghakeshmiri, F., M. Azizzadeh, N. Farzaneh, and M. Gorjidooz. 2017. Effects of neonatal diarrhea and other conditions on subsequent productive and reproductive performance of heifer calves. Veterinary Research Communications 41(2):107-112. https://doi.org/10.1007/s11259-017-9678-9.

Beaudeau, F., H. Seegers, V. Ducrocq, C. Fourichon, and N. Bareille. 2000. Effect of health disorders on culling in dairy cows: a review and a critical discussion. Ann. Zootech. 49(4):293-311. https://doi.org/10.1051/animres:2000102.

Betensky, R. A. 2019. The p-value requires context, not a threshold. The American Statistician 73(sup1):115-117. https://doi.org/10.1080/00031305.2018.1529624.

Cardoso, C. S., M. J. Hötzel, D. M. Weary, J. A. Robbins, and M. A. G. von Keyserlingk. 2016. Imagining the ideal dairy farm. Journal of Dairy Science 99(2):1663-1671. https://doi.org/10.3168/jds.2015-9925.

CDIC – Canadian Dairy Information Centre. 2021. Culling and replacement rates in dairy herds in Canada. Accessed on 07 July 2021. https://agriculture.canada.ca/en/canadas-agriculture-sectors/animal-industry/canadian-dairy-information-centre/dairy-statistics-and-market-information/dairy-animal-genetics/culling-and-replacement-rates-dairy-herds-canada.

Compton, C. W. R., C. Heuer, P. T. Thomsen, T. E. Carpenter, C. V. C. Phyn, and S. McDougall. 2017. Invited review: A systematic literature review and meta-analysis of mortality and culling in dairy cattle. Journal of Dairy Science 100(1):1-16. https://doi.org/10.3168/jds.2016-11302.

Dallago, G. M., K. M. Wade, R. I. Cue, J. T. McClure, R. Lacroix, D. Pellerin, and E. Vasseur. 2021. Keeping dairy cows for longer: A critical literature review on dairy cow longevity in high milk-producing countries. Animals 11(3):808. https://doi.org/10.3390/ani11030808.

De Vries, A. and M. I. Marcondes. 2020. Review: Overview of factors affecting productive lifespan of dairy cows. Animal 14(S1):s155-s164. https://doi.org/10.1017/S1751731119003264.

Delgado, H. A., R. I. Cue, D. Haine, A. Sewalem, R. Lacroix, D. Lefebvre, J. Dubuc, E. Bouchard, and K. M. Wade. 2017. Profitability measures as decision-making tools for Québec dairy herds. Canadian Journal of Animal Science 98(1):18-31. https://doi.org/10.1139/cjas-2016-0202.

DeVries, T. J., M. G. Aarnoudse, H. W. Barkema, K. E. Leslie, and M. A. G. von Keyserlingk. 2012. Associations of dairy cow behavior, barn hygiene, cow hygiene, and risk of elevated somatic cell count. Journal of Dairy Science 95(10):5730-5739. https://doi.org/10.3168/jds.2012-5375.

Ducrocq, V. 1994. Statistical analysis of length of productive life for dairy cows of the Normande breed. Journal of Dairy Science 77(3):855-866. https://doi.org/10.3168/jds.S0022-0302(94)77020-X.

Ducrocq, V., R. L. Quaas, E. J. Pollak, and G. Casella. 1988. Length of productive life of dairy cows. 1. Justification of a Weibull model. Journal of Dairy Science 71(11):3061-3070. https://doi.org/10.3168/jds.S0022-0302(88)79906-3.

Fetrow, J., K. V. Nordlund, and H. D. Norman. 2006. Invited Review: Culling: Nomenclature, definitions, and recommendations. Journal of Dairy Science 89(6):1896-1905. https://doi.org/10.3168/jds.S0022-0302(06)72257-3.

Gelman, A. and H. Stern. 2006. The difference between “significant” and “not significant” is not itself statistically significant. The American Statistician 60(4):328-331. https://doi.org/10.1198/000313006X152649.

Grandl, F., M. Furger, M. Kreuzer, and M. Zehetmeier. 2019. Impact of longevity on greenhouse gas emissions and profitability of individual dairy cows analysed with different system boundaries. Animal 13(1):198-208. https://doi.org/10.1017/S175173111800112X.

Grandl, F., S. L. Amelchanka, M. Furger, M. Clauss, J. O. Zeitz, M. Kreuzer, and A. Schwarm. 2016. Biological implications of longevity in dairy cows: 2. Changes in methane emissions and efficiency with age. Journal of Dairy Science 99(5):3472-3485. https://doi.org/10.3168/jds.2015-10262.

Habel, J., V. Uhlig, S. Hoischen-Tauber, E.-M. Schwabenbauer, T. Rumphorst, L. Ebert, D. Möller, and A. Sundrum. 2021. Income over service life cost – Estimation of individual profitability of dairy cows at time of death reveals farm-specific economic trade-offs. Livestock Science 254:104765. https://doi.org/10.1016/j.livsci.2021.104765.

Heise, J., Z. Liu, K. F. Stock, S. Rensing, F. Reinhardt, and H. Simianer. 2016. The genetic structure of longevity in dairy cows. Journal of Dairy Science 99(2):1253-1265. https://doi.org/10.3168/jds.2015-10163.

Horn, M., W. Knaus, L. Kirner, and A. Steinwidder. 2012. Economic evaluation of longevity in organic dairy cows. Organic Agriculture 2(2):127-143. https://doi.org/10.1007/s13165-012-0027-6.

Hultgren, J. and C. Svensson. 2009. Heifer rearing conditions affect length of productive life in Swedish dairy cows. Preventive Veterinary Medicine 89(3):255-264. https://doi.org/10.1016/j.prevetmed.2009.02.012.

Lactanet. 2020. The evolution of Lactanet Atlantic dairy production: 2019 Stats & Tips. Part 2. Accessed on 29 October 2020. https://www.valacta.com/en-CA/library/the-2019-evolution-lactanet-atlantic-dairy-production-part-2.

Lantz, B. 2013. The large sample size fallacy. Scandinavian Journal of Caring Sciences 27(2):487-492. https://doi.org/10.1111/j.1471-6712.2012.01052.x.

McPherson, S. E. and E. Vasseur. 2020. Graduate student literature review: The effects of bedding, stall length, and manger wall height on common outcome measures of dairy cow welfare in stall-based housing systems. Journal of Dairy Science 103(11):10940-10950. https://doi.org/10.3168/jds.2020-18332.

Nebel, R. L. and M. L. McGilliard. 1993. Interactions of High Milk Yield and Reproductive Performance in Dairy Cows. Journal of Dairy Science 76(10):3257-3268. https://doi.org/https://doi.org/10.3168/jds.S0022-0302(93)77662-6.

Pellerin, D., S. Adams, F. Bécotte, R. Cue, R. Moore, and R. Roy. 2014. Pour une vache, l'âge d'or c'est la 4e lactation! [For a cow, the golden age is the fourth lactation!]. Pages 133-147 in Proc. Symposium sur les bovins laitiers: Choix d'aujourd'hui pour les défis de demain. CRAAQ – Le Centre de référence en agriculture et agroalimentaire du Québec, Saint-Hyacinthe, QC, Canada.

Pinedo, P. J., A. De Vries, and D. W. Webb. 2010. Dynamics of culling risk with disposal codes reported by Dairy Herd Improvement dairy herds. Journal of Dairy Science 93(5):2250-2261. https://doi.org/10.3168/jds.2009-2572.

Pryce, J. E., M. D. Royal, P. C. Garnsworthy, and I. L. Mao. 2004. Fertility in the high-producing dairy cow. Livestock Production Science 86(1):125-135. https://doi.org/10.1016/S0301-6226(03)00145-3.

Reneau, J. K., A. J. Seykora, B. J. Heins, M. I. Endres, R. J. Farnsworth, and R. F. Bey. 2005. Association between hygiene scores and somatic cell scores in dairy cattle. Journal of the American Veterinary Medical Association 227(8):1297-1301. https://doi.org/10.2460/javma.2005.227.1297.

Rilanto, T., D.-A. Viidu, T. Kaart, T. Orro, A. Viltrop, U. Emanuelson, E. Ferguson, and K. Mõtus. 2022. Attitudes and personality of farm managers and association with cow culling rates and longevity in large-scale commercial dairy farms. Research in Veterinary Science 142:31-42. https://doi.org/10.1016/j.rvsc.2021.11.006.

Santman-Berends, I. M. G. A., J. M. Swinkels, T. J. G. M. Lam, J. Keurentjes, and G. van Schaik. 2016. Evaluation of udder health parameters and risk factors for clinical mastitis in Dutch dairy herds in the context of a restricted antimicrobial usage policy. Journal of Dairy Science 99(4):2930-2939. https://doi.org/10.3168/jds.2015-10398.

Schreiner, D. A. and P. L. Ruegg. 2003. Relationship between udder and leg hygiene scores and subclinical mastitis. Journal of Dairy Science 86(11):3460-3465. https://doi.org/10.3168/jds.S0022-0302(03)73950-2.

Sewalem, A., F. Miglior, G. J. Kistemaker, P. Sullivan, and B. J. Van Doormaal. 2008. Relationship between reproduction traits and functional longevity in Canadian dairy cattle. Journal of Dairy Science 91(4):1660-1668. https://doi.org/10.3168/jds.2007-0178.

Somers, J. G. C. J., K. Frankena, E. N. Noordhuizen-Stassen, and J. H. M. Metz. 2003. Prevalence of claw disorders in Dutch dairy cows exposed to several floor systems. Journal of Dairy Science 86(6):2082-2093. https://doi.org/10.3168/jds.S0022-0302(03)73797-7.

Stojkov, J., M. A. G. von Keyserlingk, T. Duffield, and D. Fraser. 2020. Management of cull dairy cows: Culling decisions, duration of transport, and effect on cow condition. Journal of Dairy Science 103(3):2636-2649. https://doi.org/10.3168/jds.2019-17435.

Svensson, C. and J. Hultgren. 2008. Associations between housing, management, and morbidity during rearing and subsequent first-lactation milk production of dairy cows in southwest Sweden. Journal of Dairy Science 91(4):1510-1518. https://doi.org/10.3168/jds.2007-0235.

Tullo, E., A. Finzi, and M. Guarino. 2019. Review: Environmental impact of livestock farming and Precision Livestock Farming as a mitigation strategy. Science of the Total Environment 650:2751-2760. https://doi.org/10.1016/j.scitotenv.2018.10.018.

Van Doormaal, B. 2009. A closer look at longevity. Accessed on 28 October 2020. https://www.cdn.ca/document.php?id=162.

Vanegas, J., M. Overton, S. L. Berry, and W. M. Sischo. 2006. Effect of rubber flooring on claw health in lactating dairy cows housed in free-stall barns. Journal of Dairy Science 89(11):4251-4258. https://doi.org/10.3168/jds.S0022-0302(06)72471-7.

Vasseur, E., F. Borderas, R. I. Cue, D. Lefebvre, D. Pellerin, J. Rushen, K. M. Wade, and A. M. de Passillé. 2010. A survey of dairy calf management practices in Canada that affect animal welfare. Journal of Dairy Science 93(3):1307-1316. https://doi.org/10.3168/jds.2009-2429.