Measuring lameness prevalence: effects of case definition and assessment frequency

by Sahar, M.W., Beaver, A., Daros, R.R., von Keyserlingk, M.A. and Weary, D.M.

DOI link to the version of record on the publisher's site

Sahar, M.W., Beaver, A., Daros, R.R., von Keyserlingk, M.A. and Weary, D.M. (2022) 'Measuring lameness prevalence: effects of case definition and assessment frequency', *Journal of Dairy Science*, *105*(9), pp.7728-7737.

J. Dairy Sci. 105:7728–7737
https://doi.org/10.3168/jds.2021-21536

Measuring lameness prevalence: Effects of case definition and assessment frequency

Mohammad W. Sahar,[1] Annabelle Beaver,[2] Ruan R. Daros,[3] Marina A. G. von Keyserlingk,[1] and Daniel M. Weary[1]*

[1]Animal Welfare Program, Faculty of Land and Food Systems, University of British Columbia, Vancouver, British Columbia, Canada, V6T 1Z4
[2]Department of Animal Health Behaviour and Welfare, Harper Adams University, Newport, Shropshire, United Kingdom, TF10 8NB
[3]Graduate Program in Animal Science, School of Medicine and Life Sciences, Pontifícia Universidade Católica do Paraná, Curitiba, Brazil, 80215-901

ABSTRACT

Lameness assessments are commonly conducted at a single point in time, but such assessments are subject to multiple sources of error. We conducted a longitudinal study, assessing the gait of 282 lactating dairy cows weekly during the first 12 wk of lactation, with the aim of assessing how lameness prevalence changed in relation to case definition and assessment frequency. Gait was scored using a 5-point scale where scores of 1 and 2 were considered sound, 3 was clinically lame, and 4 and 5 were severely lame. We created 5 lameness definitions using increasingly stringent thresholds based upon the number of consecutive events of locomotion score ≥3. In LAME1, a cow was considered lame when locomotion score was ≥3 at any scoring event, in LAME2, LAME3, LAME4, and LAME5, a cow was considered lame when locomotion score was 3 or higher during 2, 3, 4, and 5 consecutive scoring events, respectively. We also assessed the effect of assessment frequency on measures of prevalence and incidence using weekly assessment (ASSM1), 1 assessment every 2 wk (ASSM2), 1 assessment every 3 wk (ASSM3), and 1 assessment every 4 wk (ASSM4). Using LAME1, 69.2% of cows were considered lame at some point during the trial, with an average point prevalence of 31.8% (SD: 2.8) and average incidence rate of 10.9 cases/100 cow weeks (SD: 3.7). Lameness prevalence decreased to 28.0% when using LAME5. Survival analysis was used to assess the effects of parity, using these different case definitions. Parity is a known risk for lameness, such that case definitions and prevalence estimates should be stratified by parity to inform management decisions. Using the LAME3 criterion, primiparous cows had the highest chance of reaching 12 wk without a lameness event, and fourth and higher parities had the lowest. Weighted linear and quadratic kappa values were used to assess agreement between different assessment frequencies and lameness definitions; we found substantial to excellent agreement between ASSM1 and ASSM2 using LAME1, LAME2, and LAME3 definitions. Agreement was fair to substantial between ASSM1 and ASSM3 and low to fair between ASSM1 and ASSM4. Likewise, the agreement between LAME1 and LAME2 was fair in primiparous cows, substantial in second and third parity cows, and poor to fair in fourth and greater parity cows. We conclude that lameness prevalence estimates are dependent upon case definition and that the use of more stringent case definitions results in fewer cows classified as lame. These results suggest that routine locomotion assessments be conducted at least every 2 wk, and that cows should be defined as lame on the basis of 2 consecutive assessments.

Key words: animal welfare, impaired gait, hoof lesion, accuracy

INTRODUCTION

Lameness is an important welfare concern in dairy cattle (Archer et al., 2010), resulting in pain (Rushen et al., 2007) and reduced milk production (Warnick et al., 2001), longevity (Booth et al., 2004), and reproductive performance (Bicalho et al., 2009; Hernandez et al., 2001). Unfortunately, lameness can be difficult to detect. For example, Espejo et al. (2006) found that farm managers only recognized about 30% of the lameness cases recorded by a trained observer.

Lameness prevalence (i.e., the percentage of cows lame at any one time) is commonly reported in cross-sectional studies (Barberg et al., 2007; Bicalho et al., 2009; Solano et al., 2015), based on a single scoring event. The accuracy of one-time locomotion scores remains unclear; even experienced locomotion scorers may show low reliability identifying mildly lame cows (i.e., cows with locomotion score = 3; e.g., Schlageter-Tello et al., 2015). An emerging body of research is

Received November 9, 2021.
Accepted April 25, 2022.
*Corresponding author: dan.weary@ubc.ca

now using longitudinal studies to investigate associations between lameness and individual animal factors such as BCS and milk yield (Green et al., 2014), and between hoof lesions and previous lameness cases to assess the future risk of becoming lame (Randall et al., 2016, 2018). Recent work by our research group (Daros et al., 2019; Eriksson et al., 2020) used longitudinal data to investigate the prevalence, incidence (i.e., the number of new cases of lameness) and cure rates of lameness in dairy cattle during the dry period. To our knowledge, only Randall et al. (2018) described lameness prevalence and incidence rates for lactating cows based upon regular locomotion scoring.

Parity is a known risk factor for lameness (e.g., Randall et al., 2018; Bran et al., 2019). In comparison to primiparous cows, second, third, fourth, and higher parity cows have 1.6-, 3.3-, and 4-times higher odds of being lame, respectively (Solano et al., 2015). Thus, effects of case definition and assessment frequency may vary with cow parity.

The aims of the current study were 3-fold. First, we sought to estimate the prevalence and incidence rate of lameness in lactating dairy cows using 5 increasingly stringent case definitions. The first definition required only that locomotion score (**LS**) was $\geq$3 in a single gait scoring event (**LAME1**), with other definitions requiring LS $\geq$3 for at least 2 (LAME2), 3 (LAME3), 4 (LAME4), or 5 (LAME5) consecutive scoring events. Second, we set out to assess the effect of frequency of locomotion scoring on prevalence and incidence rate estimates based on LAME1, LAME2, and LAME3, with assessment frequency ranging from 1/wk (**ASSM1**) to 1 assessment every 4 wk (**ASSM4**; with **ASSM2** and **ASSM3** based assessments every 2 and 3 wk, respectively). Our final aim was to explore the relationships between lameness definition (LAME1 to LAME5), assessment frequency (ASSM1 to ASSM4) and parity.

MATERIALS AND METHODS

This study was conducted at the Dairy Education and Research Centre of the University of British Columbia (**UBC**; Vancouver, BC, Canada), approved under UBC animal-care protocol A19-0299. We enrolled all cows that calved and entered the lactating herd from December 2019 to September 2020, resulting in a final sample of 282 Holstein cows (197 multiparous; mean parity = 3.3, SD = 1.3; range of 2–7) with mean DIM of 45.1 (SD = 25.9); no a priori power analysis was performed. All cows were locomotion scored weekly for 12 consecutive wk, resulting in 3,384 scoring events. The total duration of the study (and hence also the sample of cows followed) was set by the availability of the PhD student (WS) leading data collection.

Within 24 h of calving, fresh cows were moved to a pen containing early lactation cows (mean DIM = 85, SD = 43.8; pen size = 24 to 36 cows). Stocking density never exceeded 100% of stall capacity. All cows were housed in pens containing 3 rows of freestalls. The stall dimensions were 2.6 × 1.2 m with a neck rail 1.2 m above the stall surface, separated by Y2K partitions (Artex). Sand bedding was approximately 20 cm deep and replenished as needed. Stalls were raked twice per day. Manure was removed from alleys using an automatic scraper (Houle) 6 times/d.

Cows were fed fresh TMR twice per day at approximately 0800 and 1600 h. The TMR was balanced based on National Research Council (NRC, 2001) recommendations for a hypothetical Holstein cow at 65 DIM, producing 45 L of milk/d with 4.25% butter fat and 3% protein, and weighing 720 kg with an ADG of 0.1 kg/d. Cows were milked twice per day at approximately 0500 to 0900 and 1500 to 1900 h in a parallel double-12 milking parlor. Cows were never kept more than 1h in the holding area before milking.

Locomotion Assessment

Locomotion scoring was conducted weekly immediately after milking as the cows exited the parlor, using a 5-point scale (Flower and Weary, 2006). Scores of 1 and 2 were considered sound, 3 as clinically lame, and 4 and 5 as severely lame. Locomotion assessment was done by 1 observer at a time (a total of 2 observers) as cows walked individually down a 12 × 1 m alley. Observers were trained to score locomotion using prerecorded videos of 37 cows walking in a straight line. Following video-based training observers were trained with live scoring. Interobserver reliability of live scoring (n = 93) indicated good to substantial agreement (linear weighting $\kappa_w = 0.55$, 95% CI = 0.38–0.72; quadratic weighting $\kappa_w = 0.84$, 95% CI = 0.81–0.86). The kappa values for video scoring (n = 37) indicated substantial to excellent agreement (linear weighting $\kappa_w = 0.73$, 95% CI = 0.52–0.94; quadratic weighting $\kappa_w = 0.88$, 95% CI = 0.81–0.94; Cohen, 1968). Bias index was 0.05 for live scoring and 0 for video scoring, and the prevalence index was 0.02 and −0.08 for live and video scoring, respectively (Byrt et al., 1993). Intra-observer reliability was only assessed using video scoring; kappa values indicated substantial to excellent agreement (Cohen, 1968; linear weighting $\kappa_w = 0.69$, 95% CI = 0.46–0.92; quadratic weighting $\kappa_w = 0.93$, 95% CI = 0.90–0.96; Supplemental Tables S1–S4; https://doi.org/10.5683/SP3/HNN9HH; Sahar et al., 2022).

Cows were categorized into 4 parity groups: first, second, third, and fourth and higher. As per standard farm practice, no footbath was used and all cows were

Table 1. Percent of lactating cows classified as lame at some point of the study based on the 5 case definitions of lameness; cows (n = 282) were locomotion scored for 12 consecutive wk

Definition	Total lame	Severely lame[1]	% Total lame	% Severely lame	% Severely lame as % of total lame
LAME1[2]	195	49	69.2	17.4	25.1
LAME2[3]	138	45	48.9	16.0	32.6
LAME3[4]	112	45	39.7	16.0	40.2
LAME4[5]	97	44	34.4	15.6	45.4
LAME5[6]	79	43	28.0	15.2	54.4

[1]Cows that locomotion scored ≥4 at least once.
[2]Cows that were locomotion scored ≥3 at least once.
[3]Cows that scored ≥3 at least 2 consecutive times.
[4]Cows that scored ≥3 at least 3 consecutive times.
[5]Cows that scored ≥3 at least 4 consecutive times.
[6]Cows that scored ≥3 at least 5 consecutive times.

hoof trimmed at the end of lactation. In addition, some cows were trimmed when identified as lame by farm staff; during the 12-wk observational period, only 21 of the enrolled cows were trimmed. Because of the low number of cows we did not include hoof trimming as a covariate in our final models. Preliminary models with and without these 21 cows did not meaningfully change agreement or significance of any result reported below.

Data Preparation and Statistical Analyses

Lameness Definitions. Building upon Eriksson et al. (2020), we created 5 lameness definitions (LAME1, LAME2, LAME3, LAME4, and LAME5) to assess how an increasingly stringent case definition affected the identification of lameness cases. There are likely errors in locomotion assessments especially when identifying mildly lame (LS = 3) cows; increasingly stringent definitions, requiring that cows score as lame (LS ≥3) on multiple consecutive occasions, reduce the risk of false positives.

Assessment Frequency. To determine the effect of assessment frequency on identification of new lameness cases we considered the results of weekly assessments (i.e., ASSM1) as the reference. We first calculated the percentage of cows classified as lame based on ASSM1, and then calculated the percentage of lame cows missed when moving from ASSM1 to ASSM2, ASSM3, and ASSM4 (i.e., assessments every 2, 3, and 4 wk); using the following formula:

$$\%\text{ of missed lameness cases} = 100 - \frac{\%\text{ of lameness cases based on ASSM2, 3, or 4,} \times 100}{\%\text{ of lameness cases based on ASSM1}}.$$

This assessment was done separately using the LAME1, LAME2, and LAME3 criteria, assessing linear and quadratic weighted kappa values following Eriksson et al. (2020). To gauge reliability of one-time scoring, we used weighted linear and quadratic kappa to assess the agreement of LAME1 with LAME2 and LAME3 using weekly LS assessment (ASSM1) and 1 assessment every 2 wk (ASSM2).

We used R version 4.0.3 (https://r-project.org) for statistical analysis. Data were prepared using the tidyverse package (Wickham et al., 2019). Figures were created using the ggplot2 package (Wickham, 2011). Using the survival package, we performed Kaplan-Meier survival analysis and log-rank tests to assess lameness survival probability across parities and definitions, and used Cox proportional hazard models to compare lameness hazard according to each lameness definition. Weighted linear and quadratic kappa values, as well bias and prevalence indices (see Byrt et al., 1993), were used to assess interobserver reliability. We calculated the incidence rate as follows:

$$\frac{\text{Number newly diagnosed in the current week}}{\text{Number of sound cows in the previous week}} \times 100.$$

The data sets and accompanying R scripts are available at https://doi.org/10.5683/SP3/HNN9HH.

RESULTS

The number and percentage of cows classified as lame and severely lame, based on the LAME1 to LAME5 case definitions, are shown in Table 1. To describe changes in the prevalence and incidence rate in relation to parity, we used the LAME1 case definition as presented in Table 2.

The hazard of being classified as lame increased with time since calving, and decreased with increasing stringency of the case definition (Figure 1). The category LAME1 classified more animals as lame than all other categories ($P < 0.01$); LAME2 classified more cows as

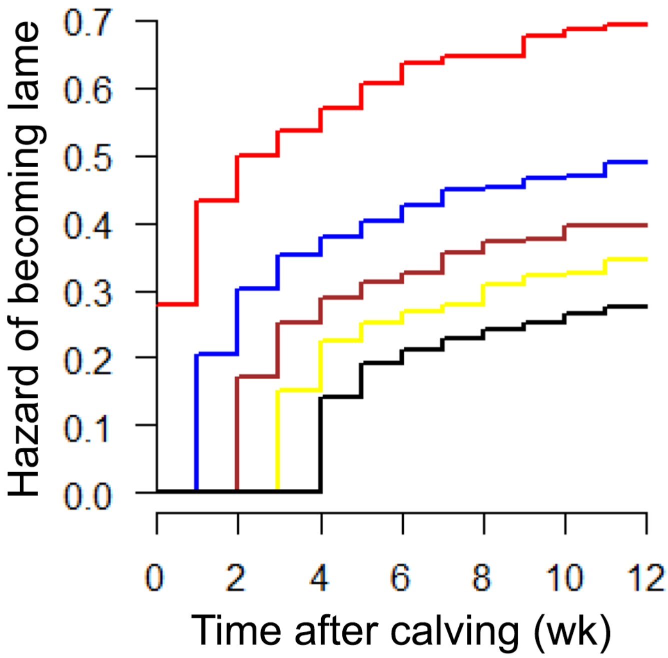

Figure 1. Hazard of lactating Holstein cows becoming lame by lameness definitions. LAME1 is when a cow is scored ≥3 at least once; LAME2, LAME3, LAME4, and LAME5 are when a cow is scored ≥3 at least 2, 3, 4, and 5 consecutive times, respectively. The red, blue, brown, yellow, and black lines represent LAME1, LAME2, LAME3, LAME4, and LAME5, respectively.

lame than LAME3, LAME4, and LAME5 ($P < 0.01$); LAME3 tended to classify more cows as lame than LAME4 ($P = 0.06$) and classified more cows as lame than LAME5 ($P < 0.01$); and LAME4 classified more cows as lame than LAME5 ($P = 0.04$).

Survival analysis revealed that the number of cows reaching the end of 12-wk period without becoming lame increased with increasing stringency of lameness case definition for all parities. As expected, higher parity cows were more likely to become lame regardless of lameness definition applied (Figure 2). However, we did note an interaction between lameness definition and parity. When using the LAME1 criterion, similar proportions of primiparous and second parity cows became lame ($P = 0.36$), and third, fourth, and higher parity cows were also comparable ($P = 0.36$), but these 2 parity clusters differed from one another ($P < 0.01$). Conversely, when using LAME2 criterion, primiparous cows differed from all older cows ($P < 0.02$), and second parity cows also differed from older cows ($P < 0.02$), and cows of parity 3 and higher remained similar ($P = 0.23$). Finally, when using LAME3, LAME4, and LAME5 criteria, a smaller proportion of primiparous cows became lame relative to older cows ($P < 0.01$), and second and third parity cows were similar ($P = 0.20$) and had lower proportions of lameness compared with older cows ($P = 0.03$).

First, second, third, and fourth and greater parity cows had a 52, 56, 86, and 90% probability of being classified as lame for at least 1 wk (i.e., LAME1), respectively. Those that were classified lame in one weekly scoring event, had a 48, 74, 76, and 81% probability of remaining lame the following week. Those that remained lame for 2 consecutive weekly scoring events continued to be lame for a third week in 67, 86, 73, and 90% of cases, respectively.

The duration of lameness cases (as defined in this study) averaged 1.7 (SD = 2.6) wk lame for primiparous cows versus 3.1 (SD = 3.9), 4.6 (SD = 3.4), and 6.5 (SD = 3.2) wk lame, for second, third, and fourth or greater parity cows. Figure 3 shows the probability of a lameness episode continuing into the next consecutive week by parity. The probability of a cow remaining lame for 2 consecutive wk never exceeded 70% for primiparous cows. Second, third and fourth and above parity cows met or exceeded this 70% threshold of remaining lame during the consecutive week, after they had a lameness episode lasting 5 wk (for second and third parity cows) or 3 wk (for cows fourth parity and greater).

Table 3 shows the percentage of cows classified as lame based on LAME1, LAME2, and LAME3 criteria, using locomotion scoring frequencies of ASSM1,

Table 2. Weekly lameness prevalence and incidence per 100 cows, shown separately by parity

Parity	n	Type[1]	Wk 1	Wk 2	Wk 3	Wk 4	Wk 5	Wk 6	Wk 7	Wk 8	Wk 9	Wk 10	Wk 11	Wk 12	Mean
1	85	Prev.	10.6	17.6	14.1	10.6	14.1	12.9	21.2	14.1	12.9	11.8	14.1	14.1	14.0
		Inc.		13.2	7.1	2.7	9.2	4.1	10.8	1.5	5.5	6.8	8.0	5.5	6.8
2	70	Prev.	24.3	25.7	30.0	35.7	35.7	35.7	21.4	18.6	24.3	24.3	17.1	21.4	26.2
		Inc.		13.2	11.5	10.2	6.7	6.7	4.4	3.6	8.8	5.7	1.9	10.3	7.5
3	57	Prev.	42.1	43.9	31.6	40.4	33.3	33.3	43.9	42.1	49.1	31.6	31.6	38.6	38.5
		Inc.		24.2	6.3	15.4	8.8	13.2	26.3	15.6	30.3	13.8	12.8	15.4	16.6
≥4	70	Prev.	41.4	57.1	61.4	55.7	55.7	55.7	52.9	51.4	60.0	48.6	50.0	54.3	53.7
		Inc.		34.1	20.0	11.1	19.4	29.0	16.1	9.1	23.5	7.1	11.1	28.6	19.0
All	282	Prev.	28.0	34.8	33.3	34.0	33.7	33.3	33.7	30.1	34.8	28.0	27.3	30.9	31.8
		Inc.		19.2	10.3	8.5	10.2	10.7	13.3	5.9	13.7	7.6	7.9	12.7	10.9

[1]Prev. = prevalence. Inc. = incidence.

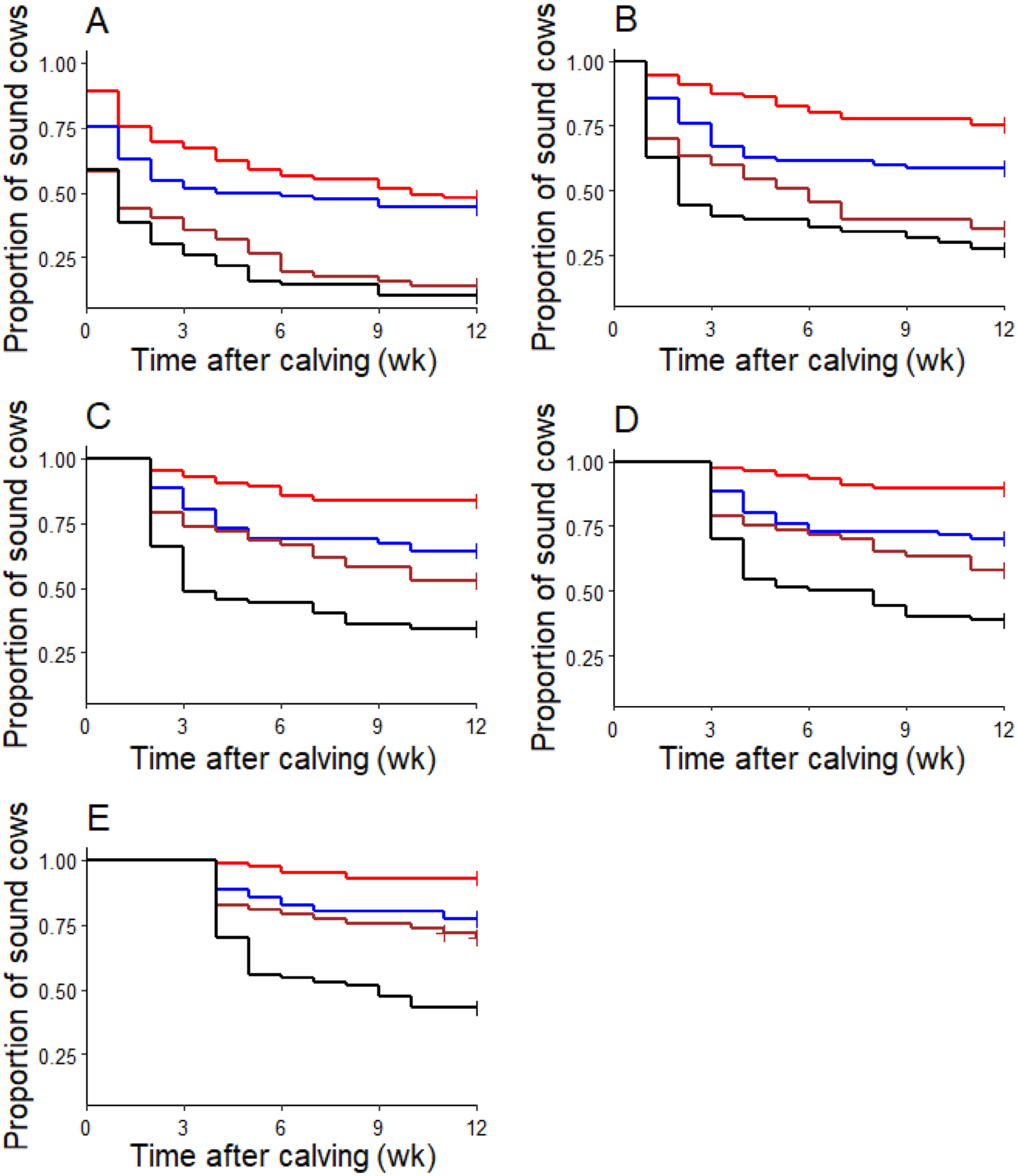

Figure 2. Survival analysis of lactating Holstein cows by parity, shown separately for each the 5 different case definitions of lameness (where, LAME1 is when a cow scores ≥3 at least once, LAME2, LAME3, LAME4, and LAME5 are when a cows scores ≥3 at least 2, 3, 4, and 5 consecutive times, respectively). Panel titles A, B, C, D, and E represent definitions of lameness as LAME1, LAME2, LAME3, LAME4, and LAME5, respectively. The red, blue, brown, and black lines represent primiparous cows and cows in second, third, and fourth and greater parities, respectively.

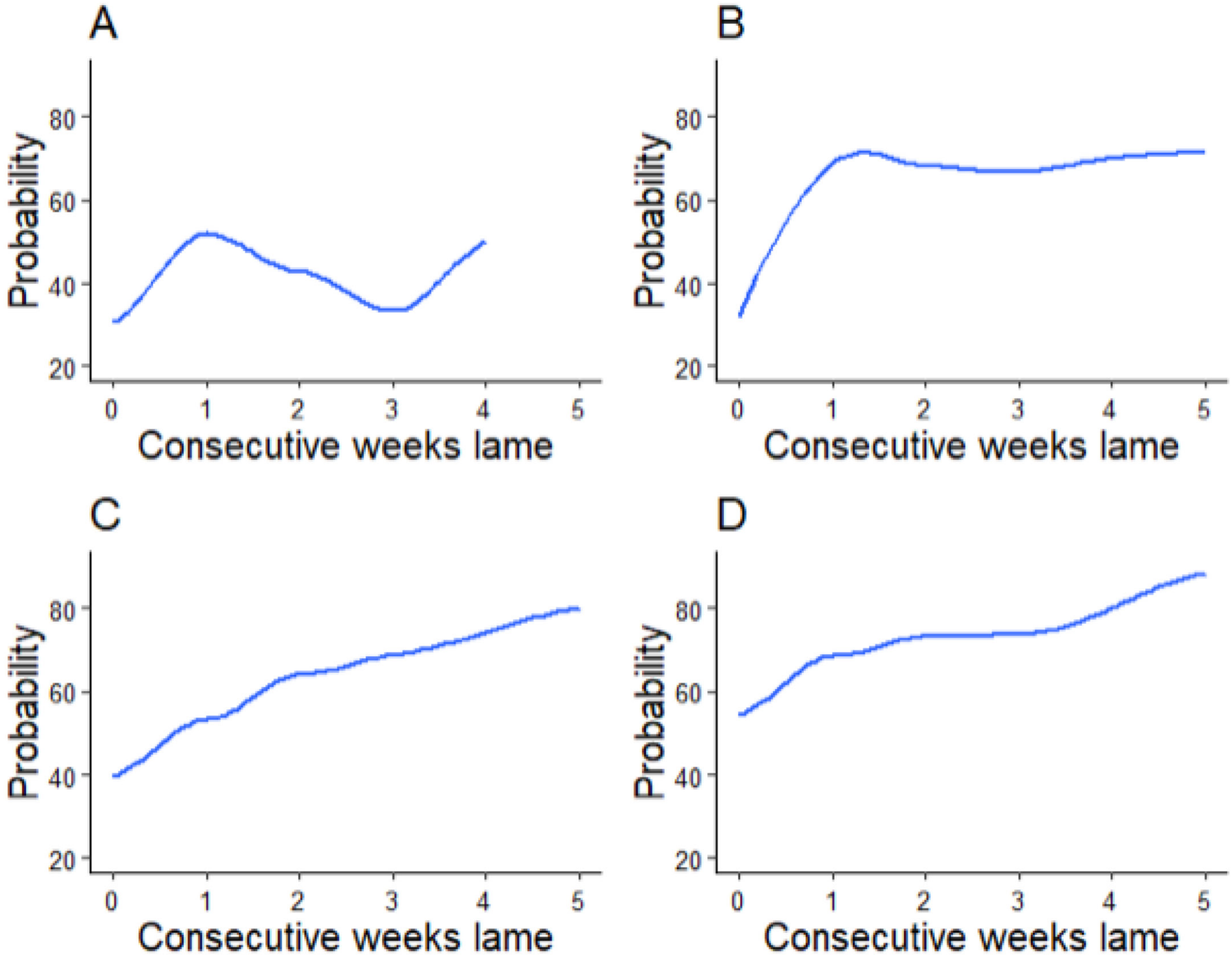

Figure 3. Probability of lameness episodes continuing the next consecutive week, stratified by parity for (A) primiparous cows, (B) second parity cows, (C) third parity cows, and (D) cows in fourth parity and greater. The probability of remaining lame is plotted on the y-axis, and the number of consecutive weeks for which a lameness episode continues is plotted on the x-axis. The probability of lameness episodes continuing the next consecutive week was calculated based on the number of episodes that continued to the current week divided by the number of episodes in the previous week multiplied by 100. The curves were fitted using the geom_smooth function in R version 4.0.3 (https://r-project.org).

ASSM2, ASSM3, and ASSM4. As expected, ASSM1 classified the largest number of cows as lame, followed by ASSM2, ASSM3, and ASSM4. Less regular assessment meant that some of the cows identified as lame based upon more regular assessments were missed, and this percentage of missed lameness cases increased with decreasing assessment frequency.

Agreement between lameness prevalence estimates generated using different assessment frequencies is shown in Table 4. Weighted linear kappa values indicate excellent agreement between ASSM1 and ASSM2 when using LAME1, and substantial agreement when using LAME2 and LAME3 criteria. Weighted quadratic kappa indicate excellent agreement between ASSM1 and ASSM2 using LAME1, LAME2, or LAME3 criteria. Agreement was fair to poor when ASSM1 was assessed against ASSM3 and ASSM4 using weighted linear kappa and the LAME2 or LAME3 criteria.

Agreement between LAME1 and LAME2 based on ASSM1 is shown in Supplemental Table S5 (https://doi.org/10.5683/SP3/HNN9HH; Sahar et al., 2022); agreement was fair in primiparous, third and fourth and greater parity cows (linear weighting κ_w = 46.8, 95% CI = 31.1–62.5; quadratic weighting κ_w = 51.4, 95% CI = 48.2–54.6). The agreement between LAME1 and LAME2 in second parity cows was substantial to excellent (linear weighting κ_w = 71.9, 95% CI = 56.5–87.4; quadratic weighting κ_w = 89.5, 95% CI = 87.5–91.5).

Table 3. The percentage of lactating cows classified as lame and the percentage of missed lame cows, using LAME1, LAME2, and LAME3 criteria[1] based on scoring assessment undertaken once a week (ASSM1), once every 2 wk (ASSM2), once every 3 wk (ASSM3), and once every 4 wk (ASSM4)

Definition	Assessment	% Lame cows	% Missed lame cows
LAME1[1]	ASSM1	69.2	—
	ASSM2	62.4	9.8
	ASSM3	57.1	17.5
	ASSM4	50.7	26.7
LAME2[2]	ASSM1	48.9	—
	ASSM2	39.4	19.4
	ASSM3	29.8	39.1
	ASSM4	27.7	43.4
LAME3[3]	ASSM1	39.7	—
	ASSM2	25.5	35.8
	ASSM3	19.1	51.9
	ASSM4	10.3	74.1

[1]LAME1 = cows that were locomotion scored ≥3 at least once; LAME2 = cows that were scored ≥3 at least 2 consecutive times; and LAME3 = cows that were scored ≥3 at least 3 consecutive times. LAME4 and LAME5 are not shown, as these definitions required a minimum of 16 and 20 consecutive wk of locomotion scoring, respectively, and the current study followed cows for just 12 wk.

The agreement of ASSM1 with ASSM2, ASSM3, and ASSM4 varied with parity (Supplemental Table S6; https://doi.org/10.5683/SP3/HNN9HH; Sahar et al., 2022). For primiparous cows, agreement between ASSM1 and ASSM2 was substantial to excellent (linear weighting κ_w = 65.0, 95% CI =45.4–84.6; quadratic weighting κ_w = 82.4, 95% CI = 78.3–86.5) and agreement between ASSM1 and ASSM3 was fair to substantial (linear weighting κ_w = 57.8, 95% CI = 36.7–78.9; quadratic weighting κ_w = 61.1, 95% CI = 54.8–67.4). In second and third parity cows, agreement between ASSM1 and ASSM2 was substantial to excellent (linear weighting κ_w = 70.3, 95% CI = 58.3–82.3; quadratic weighting κ_w = 90.2, 95% CI = 89.1–91.3), and the agreement between ASSM1 and ASSM3 was fair (linear weighting κ_w = 44.3, 95% CI = 30.6–57.9; quadratic weighting κ_w = 54.5, 95% CI = 52.4–56.7). Likewise, in fourth and higher parity cows, the agreement between ASSM1 and ASSM2 was excellent (linear weighting κ_w = 80.3, 95% CI = 65.5–95.1; quadratic weighting κ_w = 94.4, 95% CI = 92.6–96.2), and the agreement between ASSM1 and ASSM3 was substantial (linear weighting κ_w = 61.3, 95% CI = 43.9–78.8; quadratic weighting κ_w = 75.8, 95% CI = 72.4–97.1).

DISCUSSION

Previous studies have reported lameness prevalence at 37% in England (Barker et al., 2010), 34% in Germany and Austria (Dippel et al., 2009), 58% in California, and 35% in British Columbia (von Keyserlingk et al., 2012), and 19% in Alberta, 22% in Ontario, and 24% in Quebec (Solano et al., 2015). These values correspond well with the average point prevalence in the current study of 32% (SD = 2.8%) across 12 wk of LS. When lameness prevalence was assessed longitudinally, 69% of cows experienced lameness on at least one instance over 12 wk of observation. Similarly, Eriksson et al. (2020) reported 82% lameness incidence based upon weekly assessment of dry cows over a 9-wk period.

In the current study, increased parity was associated with higher lameness prevalence, a result consistent with previous literature (e.g., Solano et al., 2015). This relationship varied somewhat with lameness definition. The effects of parity were less pronounced when applying the LAME1 case definition, perhaps because this definition overestimates prevalence especially in first parity cows. Based on the more stringent definitions of lameness there was a clear separation of 3 parity groups: primiparous, second and third parity cows, and

Table 4. Agreement between ASSM1 and ASSM2, ASSM1 and ASSM3, and ASSM1 and ASSM4 using LAME1, LAME2, and LAME3 criteria[1]

Assessment	Criteria	Weighted linear kappa (95% CI)	Weighted quadratic kappa (95% CI)
ASSM1 vs. ASSM2	LAME1	85.1 (78.7, 91.5)	97.1 (96.8, 97.4)
	LAME2	75.1 (67.5, 82.7)	93.5 (93.1, 93.9)
	LAME3	68.5 (59.9, 77.0)	84.0 (83.2, 84.8)
ASSM1 vs. ASSM3	LAME1	74.5 (66.7, 82.3)	90.7 (90.2, 91.3)
	LAME2	57.1 (48.2, 66.0)	72.9 (72.1, 73.8)
	LAME3	51.3 (41.6, 60.9)	56.9 (55.7, 58.2)
ASSM1 vs. ASSM4	LAME1	62.9 (54.5, 71.4)	78.8 (78.1, 79.6)
	LAME2	55.6 (46.8, 64.4)	68.4 (67.5, 69.3)
	LAME3	27.9 (18.9, 37.0)	14.8 (13.9, 15.8)

[1]The agreements are calculated using weighted linear kappa and weighted quadratic kappa, both with 95% CI. ASSM1 = locomotion scoring once a week; ASSM2 = locomotion scoring once every 2 wk; ASSM3 = locomotion scoring once every 3 wk; ASSM4 = locomotion scoring once every 4 wk. LAME1 = cows that scored ≥3 at least once; LAME2 = cows that scored ≥3 at least 2 consecutive times; and LAME3 = cows that scored ≥3 at least 3 consecutive times. LAME4 and LAME5 are not shown, as these definitions required a minimum of 16 and 20 consecutive wk of locomotion scoring, respectively, and the current study followed cows for just 12 wk.

fourth and greater parity cows. That more stringent definitions of lameness were better able to differentiate between parities suggests that these definitions are more biologically meaningful.

The effect of parity may be driven in part by increased likelihood of re-occurrence, as previous cases put cattle at risk for future cases; for example, Randall et al. (2018) found that the odds of cows becoming lame was greatest for those that had previously experienced 2 consecutive weeks of severe lameness (LS = 4 or 5). Overall, 79 to 83% of lameness cases were associated with a previous episode (Randall et al., 2018).

As expected, fewer cows were classified as lame when more stringent case definitions were used (from LAME1 to LAME5). This result agrees with Eriksson et al. (2020), who assessed multiple lameness definitions for dry cows. The case definitions used in the current study were slightly different than those used by Eriksson et al. (2020), who included lameness severity into their classification; for example, defining LAME2 as LS = 3 for 2 consecutive scoring events, or LS $\geq$4 at least once. The case definition applied in the current study considered only the number of weeks in which LS $\geq$ 3 so as to specifically track the duration of lameness episodes.

Much research on cattle lameness has relied upon a single observation of gait (Barker et al., 2010; Solano et al., 2015; Costa et al., 2018), with the notable exception of Randall et al. (2018) who classified cows as lame based on 2 consecutive scores. The results of current study suggest lameness assessments based upon a single observation are error prone and likely to inflate prevalence estimates. For example, we found that agreement between LAME1 and LAME2 was only fair in primiparous cows, substantial in second and third parity cows, and poor to fair in fourth and greater parity cows.

This lack of agreement may be driven by the ability of some cows to self-cure; for instance, Leach et al. (2012) found that 30% of cows in a control group self-cured. We calculated the probability of lame cows remaining lame the following week. Approximately 52% of primiparous cows were classified as lame at some point during the study compared with 90% of fourth or greater parity cows. Of these lame cows, approximately 48% of primiparous and 81% of fourth or greater parity cows continued to be lame the week following initial diagnosis. In other words, 52% of primiparous and 19% of fourth or greater parity cows changed from lame to sound in the week after an initial diagnosis. However, some changes in how cows were classified across assessments were likely due to misclassification of sound cows as lame (and vice-versa). Schlageter-Tello et al. (2014) reported that identifying clinically lame cows (e.g., distinguishing between LS 2 and 3 in the current study) is difficult even for experienced observers. Based on the current results, we suggest that the estimates of prevalence based on the LAME1 definition be treated with caution. More stringent case definitions, such as LAME2, may result in better estimates of prevalence, incidence, cure rate, and lameness bout duration.

Locomotion assessment is time consuming. The results of the current study can help better identify the optimal frequency of locomotion assessments. We found that the percentage of cows classified as lame gradually reduced as we reduced the frequency of scoring. When we assessed the level of agreement between assessment frequencies, we found that ASSM1 has excellent agreement with ASSM2 using LAME1 and substantial agreement using LAME2 or LAME3 criteria. These results agree with Eriksson et al. (2020) comparing ASSM1 with ASSM2 using the LAME2 criterion and together suggest that ASSM2 can be used to effectively identify new lameness cases.

Importantly, we are not recommending against more frequent locomotion scoring, or against the detection and treatment of lameness cases outside of routine assessments and those caused by events such as falls; rather, we recommend that routine scoring should be done *at least* every 2 wk to better track new and existing lameness cases. We also note that reassessments are likely less beneficial in cases of more severe lameness (i.e., scores of 4 or higher); for those cows we advocate for immediate treatment.

A potential limitation of our study is that some cows were selected for treatment by the farm staff; we had no control of which cows were selected or how they were treated, although it is important to note that our assessments were not provided to farm staff and we saw no evidence that the inclusion of these cows affected our results. A controlled experiment would be beneficial to assess the effects of different treatment strategies.

Lameness is a multifactorial condition associated with a variety of ailments including digital dermatitis, foot rot, and injuries. The current study did not collect data on specific causes of lameness, which may be considered a limitation of our work. However, our approach reflects the reality on many commercial farms; when farmers assess lameness, they may not be aware of underlying causes. That said, future studies could seek to disentangle the multifactorial causes of lameness, perhaps aided by the definitions and assessment frequencies recommended in the current study.

CONCLUSIONS

Lameness prevalence estimates are reduced when using more stringent lameness definitions and less frequent assessments; estimates based upon one-time assessments of locomotion should be viewed with caution and we recommend routine assessments at least every 2 wk. Regardless of lameness definition, primiparous cows are less likely to become lame than older cows.

ACKNOWLEDGMENTS

General funding for the Animal Welfare Program [University of British Columbia (UBC), Vancouver, BC, Canada] is provided by Canada's Natural Sciences and Engineering Research Council (NSERC) via the Industrial Research Chair Program with industry contributions from Dairy Farmers of Canada (Ottawa, ON, Canada), Alberta Milk (Edmonton, AB, Canada), Saputo (Montreal, QC, Canada), British Columbia Dairy Association (Burnaby, BC, Canada), Merck (Kirkland, QC, Canada), British Columbia Cattle Industry Development Fund (Kamloops, BC, Canada), Boehringer Ingelheim (Burlington, ON, Canada), Semex Alliance (Guelph, ON, Canada), Lactanet (Sainte-Anne-de-Bellevue, QC, Canada), Dairy Farmers of Manitoba (Winnipeg, MB, Canada), and SaskMilk (Regina, SK, Canada). We thank the research assistants in the UBC Animal Welfare Program for helping in data collection, and the farmers at the UBC Dairy Education and Research Centre (Agassiz, BC, Canada) for their help and support throughout the study. The authors have not stated any conflicts of interest.

REFERENCES

Archer, S. C., M. J. Green, and J. N. Huxley. 2010. Association between milk yield and serial locomotion score assessments in UK dairy cows. J. Dairy Sci. 93:4045–4053. https://doi.org/10.3168/jds.2010-3062.

Barberg, A. E., M. I. Endres, J. A. Salfer, and J. K. Reneau. 2007. Performance and welfare of dairy cows in an alternative housing system in Minnesota. J. Dairy Sci. 90:1575–1583. https://doi.org/10.3168/jds.S0022-0302(07)71643-0.

Barker, Z. E., K. A. Leach, H. R. Whay, N. J. Bell, and D. C. J. Main. 2010. Assessment of lameness prevalence and associated risk factors in dairy herds in England and Wales. J. Dairy Sci. 93:932–941. https://doi.org/10.3168/jds.2009-2309.

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. J. Dairy Sci. 92:3175–3184. https://doi.org/10.3168/jds.2008-1827.

Booth, C. J., L. D. Warnick, Y. T. Gröhn, D. O. Maizon, C. L. Guard, and D. Janssen. 2004. Effect of lameness on culling in dairy cows. J. Dairy Sci. 87:4115–4122. https://doi.org/10.3168/jds.S0022-0302(04)73554-7.

Bran, J. A., J. H. Costa, M. A. von Keyserlingk, and M. J. Hötzel. 2019. Factors associated with lameness prevalence in lactating cows housed in freestall and compost-bedded pack dairy farms in southern Brazil. Prev. Vet. Med. 172:104773. https://doi.org/10.1016/j.prevetmed.2019.104773.

Byrt, T., J. Bishop, and J. B. Carlin. 1993. Bias, prevalence and kappa. J. Clin. Epidemiol. 46:423–429.

Cohen, J. 1968. Nominal scale agreement with provision for scaled disagreement or partial credit. Psychol. Bull. 70:213–220.

Costa, J. H. C., T. A. Burnett, M. A. G. von Keyserlingk, and M. J. Hötzel. 2018. Prevalence of lameness and leg lesions of lactating dairy cows housed in southern Brazil: Effects of housing systems. J. Dairy Sci. 101:2395–2405. https://doi.org/10.3168/jds.2017-13462.

Daros, R. R., H. K. Eriksson, D. M. Weary, and M. A. G. von Keyserlingk. 2019. Lameness during the dry period: Epidemiology and associated factors. J. Dairy Sci. 102:11414–11427. https://doi.org/10.3168/jds.2019-16741.

Dippel, S., M. Dolezal, C. Brenninkmeyer, J. Brinkmann, S. March, U. Knierim, and C. Winckler. 2009. Risk factors for lameness in freestall-housed dairy cows across two breeds, farming systems, and countries. J. Dairy Sci. 92:5476–5486. https://doi.org/10.3168/jds.2009-2288.

Eriksson, H. K., R. R. Daros, M. A. G. von Keyserlingk, and D. M. Weary. 2020. Effects of case definition and assessment frequency on lameness incidence estimates. J. Dairy Sci. 103:638–648. https://doi.org/10.3168/jds.2019-16426.

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. J. Dairy Sci. 89:3052–3058. https://doi.org/10.3168/jds.S0022-0302(06)72579-6.

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

Green, L. E., J. N. Huxley, C. Banks, and M. J. Green. 2014. Temporal associations between low body condition, lameness and milk yield in a UK dairy herd. Prev. Vet. Med. 113:63–71. https://doi.org/10.1016/j.prevetmed.2013.10.009.

Hernandez, J., J. K. Shearer, and D. W. Webb. 2001. Effect of lameness on the calving-to-conception interval in dairy cows. J. Am. Vet. Med. 218:1611–1614.

Leach, K. A., D. A. Tisdall, N. J. Bell, D. C. J. Main, and L. E. Green. 2012. The effects of early treatment for hindlimb lameness in dairy cows on four commercial UK farms. Vet. J. 193:626–632. https://doi.org/10.1016/j.tvjl.2012.06.043.

NRC. 2001. Nutrient Requirements of Dairy Cattle. 7th rev. ed. Natl. Acad. Press.

Randall, L. V., M. J. Green, M. G. G. Chagunda, C. Mason, L. E. Green, and J. N. Huxley. 2016. Lameness in dairy heifers; impacts of hoof lesions present around first calving on future lameness, milk yield and culling risk. Prev. Vet. Med. 133:52–63. https://doi.org/10.1016/j.prevetmed.2016.09.006.

Randall, L. V., M. J. Green, L. E. Green, M. G. G. Chagunda, C. Mason, S. C. Archer, and J. N. Huxley. 2018. The contribution of previous lameness events and body condition score to the occurrence of lameness in dairy herds: A study of 2 herds. J. Dairy Sci. 101:1311–1324. https://doi.org/10.3168/jds.2017-13439.

Rushen, J., E. Pombourcq, and A. M. de Passillé. 2007. Validation of two measures of lameness in dairy cows. Appl. Anim. Behav. Sci. 106:173–177. https://doi.org/10.1016/j.applanim.2006.07.001.

Sahar, M. W., A. Beaver, R. Daros, M. von Keyserlingk, and D. Weary. 2022. Measuring lameness prevalence: Effects of case definition and assessment frequency. Borealis V1. https://doi.org/10.5683/SP3/HNN9HH.

Schlageter-Tello, A., E. A. M. Bokkers, P. W. G. Groot Koerkamp, T. Van Hertem, S. Viazzi, C. E. B. Romanini, I. Halachmi, C. Bahr, D. Berckmans, and K. Lokhorst. 2014. Effect of merging levels of locomotion scores for dairy cows on intra- and interrater reliability and agreement. J. Dairy Sci. 97:5533–5542. https://doi.org/10.3168/jds.2014-8129.

Schlageter-Tello, A., E. A. M. Bokkers, P. W. G. Groot Koerkamp, T. Van Hertem, S. Viazzi, C. E. B. Romanini, I. Halachmi, C. Bahr, D. Berckmans, and K. Lokhorst. 2015. Relation between observed locomotion traits and locomotion score in dairy cows. J. Dairy Sci. 98:8623–8633. https://doi.org/10.3168/jds.2014-9059.

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. J. Dairy Sci. 98:6978–6991. https://doi.org/10.3168/jds.2015-9652.
von Keyserlingk, M. A. G., A. Barrientos, K. Ito, E. Galo, and D. M. Weary. 2012. Benchmarking cow comfort on North American freestall dairies: Lameness, leg injuries, lying time, facility design, and management for high-producing Holstein dairy cows. J. Dairy Sci. 95:7399–7408. https://doi.org/10.3168/jds.2012-5807.
Warnick, L. D., D. Janssen, C. L. Guard, and Y. T. Gröhn. 2001. The effect of lameness on milk production in dairy cows. J. Dairy Sci. 84:1988–1997. https://doi.org/10.3168/jds.S0022-0302(01)74642-5.
Wickham, H. 2011. Ggplot2. Wiley Interdiscip. Rev. Comput. Stat. 3:180–185. https://doi.org/10.1002/wics.147.
Wickham, H., M. Averick, J. Bryan, W. Chang, L. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. Pedersen, E. Miller, S. Bache, K. Müller, J. Ooms, D. Robinson, D. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, and H. Yutani. 2019. Welcome to the Tidyverse. J. Open Source Softw. 4:1686. https://doi.org/10.21105/joss.01686.

ORCIDS

Mohammad W. Sahar https://orcid.org/0000-0001-6069-7119
Annabelle Beaver https://orcid.org/0000-0002-2953-9574
Ruan R. Daros https://orcid.org/0000-0003-2331-1648
Marina A. G. von Keyserlingk https://orcid.org/0000-0002-1427-3152
Daniel M. Weary https://orcid.org/0000-0002-0917-3982