Comparison of modeled sampling strategies for estimation of dairy herd lameness prevalence and cow-level variables associated with lameness

A. C. Hoffman,* D. A. Moore,*¹ J. R. Wenz,* and J. Vanegast†

*Department of Veterinary Clinical Sciences, College of Veterinary Medicine, Washington State University, Pullman 99164

†Department of Clinical Sciences, College of Veterinary Medicine, Oregon State University, Corvallis 97331

ABSTRACT

Monitoring herd lameness prevalence has utility for dairy producers and veterinarians in their efforts to reduce lameness, for animal welfare assessment programs, and for researchers. Locomotion scoring is a method used to quantify lameness and calculate prevalence. Because of the time necessary to locomotion score each cow in large dairy herds, a sampling strategy to determine herd lameness prevalence that allows scoring of fewer cows would be useful. Such a sampling strategy must be validated for accuracy compared with the lameness prevalence when all cows in a herd are locomotion scored. The purpose of this study was to assess 3 previously suggested methods of estimating lameness prevalence by strategic sampling of dairy herds. Sampling strategies tested included (1) sampling a calculated number of cows in the middle third of the milking parlor exit order for each pen, (2) sampling a calculated number of cows weighted across pens and distributed evenly within each pen, and (3) sampling all cows in the high production, low production, and hospital pens. Lactating cows on 5 dairy farms in Washington and Oregon ($n = 4,422$) were locomotion scored using a 5-point scale to determine herd-level lameness prevalence (percentage with locomotion score ≥ 3). Milking parlor exit order, order in headlocks at the feed bunk within each pen, and breed were recorded for each cow. The number of days in lactation, milk production, and parity were collected from farm computer records. Pen grouping strategy for each farm was obtained by interview with farm management. Sampling strategies were modeled using the locomotion score data set for each herd. Estimates of lameness prevalence obtained from the milking parlor exit order sample and the sample distributed across pens were within 5 percentage points of the whole herd prevalence. The third strategy estimated the lameness prevalence within 5 percentage points on 4 farms, but overestimated prevalence on

1 farm. Pen-level prevalence obtained by locomotion score of all cows in the pen was variable and not reliably predictive of herd-level prevalence. Cows of Holstein breed, parity >1 , and exiting the milking parlor in the last 20% of the pen had greater odds of lameness compared with other breeds, parities, and milking parlor exit order groups in a multivariate analysis. This study indicates that the sampling strategies using the middle of milking parlor exit order and a calculated sample distributed across the herd may be used to obtain an estimate of herd lameness prevalence.

Key words: lameness, locomotion, prevalence, sampling

INTRODUCTION

Lameness in dairy cattle has a negative effect on production efficiency and animal welfare. Lameness has been shown to decrease milk production (Warnick et al., 2001; Juarez et al., 2003; Hernandez et al., 2005) and increase culling (Sprecher et al., 1997; Booth et al., 2004). Reproductive efficiency is affected by lameness, with lame cows having longer calving to conception intervals, higher risk of conception failure (Hernandez et al., 2001), and a higher incidence of delayed ovarian cyclicity (Garbarino et al., 2004). Including production and reproduction losses as well as treatment expenses, a case of lameness has been estimated to cost \$120 to \$215, depending on the cause (Cha et al., 2010). Additionally, lameness is an indication of limb pain and therefore affects animal well-being (Whay et al., 1997; O'Callaghan et al., 2003; Dyer et al., 2007).

The prevalence of lameness in dairy herds has been assessed in previous studies. Espejo et al. (2006) studied freestall-housed Holstein dairy cows in groups designated as high production and found an average lameness prevalence of 24.6%. Cook (2003) found an average herd lameness prevalence of 23.9% in winter and 21.1% in summer across freestall and tiestall herds. Wells et al. (1993) assessed lameness in primarily tiestall herds and found a lameness prevalence of 13.7% in the summer and 16.7% in the spring. Benchmarks for lameness prevalence in dairy herds of 15% (Nordlund

Received April 2, 2013.

Accepted June 3, 2013.

¹Corresponding author: damoore@vetmed.wsu.edu

et al., 2004) and 10% (National Milk Producers Federation, 2010) have been recommended, although the latter could include young stock and nonlactating cows. Given that the average levels of lameness reported from research studies are greater than these recommendations, reduction of lameness prevalence is needed.

Mitigation of a problem typically involves assessment and quantification of the problem first. Therefore, it is important to accurately estimate the level of lameness in a dairy herd to incorporate a lameness reduction program. Previous studies have found that producers tend to underestimate the level of lameness in their herds compared with evaluations made by an outside observer (Wells et al., 1993; Espejo et al., 2006). Therefore, if an independent observer such as a veterinarian or other dairy advisor could quickly assess the lameness prevalence in a herd, goal levels could be established and steps to mitigate lameness taken.

Locomotion scoring systems for assessing cow gait based on criteria developed by Sprecher et al. (1997) have been used to measure lameness (Cook, 2003; Juarez et al., 2003; Espejo et al., 2006). These scoring systems have been validated for intra- and inter-observer agreement and correlated with hoof pathology (Winckler and Willen, 2001; O'Callaghan et al., 2003; Bicalho et al., 2007; Tadich et al., 2010; Thomsen et al., 2012). Locomotion scoring requires the observation of a cow standing and walking, which may take several minutes before a score can be determined. An opportunity commonly used to observe cows for locomotion scoring is as they walk between the milking parlor and the housing pens. However, milking parlor throughput of 164 to 363 cows per hour has been reported (Thomas et al., 1996), which suggests that on large dairy farms cows exit the milking parlor over several hours, making the process of locomotion scoring all animals time or cost prohibitive.

Scoring a smaller number of cows using strategic sampling could reduce the time and cost of prevalence estimation. Strategies to sample cows for locomotion scoring have been suggested. Main et al. (2010) developed a sampling strategy based on milking parlor exit order using a calculated sample size that estimated lameness prevalence within 5% of the true prevalence. Farms included in that study had fewer than 300 cows and most of the cows were housed in one pen. That sampling strategy has not been tested on large farms (>300 cows) with multiple pens. Another sampling strategy is utilized by National Dairy FARM (Farmers Assuming Responsible Management) Program assessors, in which a sample size is calculated for the herd, weighted by number of cows per pen to determine sample size per pen, and cows are selected systematically from the pen (National Milk Producers Federation, 2010). Validation

of this strategy for accuracy compared with true herd lameness prevalence has not been reported. Additionally, some have suggested that locomotion scoring certain pens or combinations of pens based on pen designation or production level would be predictive of herd-level prevalence. Hoof health industry representatives have suggested sampling 1 high-production, 1 low-production, and 1 hospital pen to determine herd prevalence. Previous studies assessing lameness prevalence have locomotion scored only a high-production pen to associate pen-level management, and are thus unable to represent the herd prevalence (Espejo et al., 2006; von Keyserlingk et al., 2012).

The purpose of the present study was to compare true herd lameness prevalence, as measured by locomotion scoring all lactating cows, to the estimated herd lameness prevalence based on 3 sampling strategies: (1) scoring the cows in the middle of the group based on the order they leave the milking parlor, (2) scoring based on a sample size distributed across pens, and (3) scoring groups based on their estimated production level.

MATERIALS AND METHODS

This study was approved by the Washington State University animal care and use committee. Five dairy farms volunteered to participate in the study, 4 in Washington State and 1 in Oregon. Cows were locomotion scored using a 5-level scoring system based on the back posture and gait of a cow (Sprecher et al., 1997), where a score of 1 indicates a sound cow with a level back walking and standing, a score of 2 indicates the cow develops an arched back posture while walking but not standing and the gait remains normal, a score of 3 indicates the cow maintains an arched back posture walking and standing and short-strides one or more legs, a score of 4 indicates the cow additionally has reduced weight bearing on at least one limb, and a score of 5 indicates refusal to bear weight on a limb. If the cow was not observed while standing, the presence or absence of a short-striding gait (failure to track-up hind feet, increased speed of leg movement asymmetrically) at the walk was used to distinguish a score 2 from a score 3. Cows were classified as "lame" if they had a locomotion score of 3 or greater. Locomotion scoring was performed by one of the authors (DM, AH, or JV), who completed the same training using online lameness scoring materials (Zinpro Corp., 2010).

For calculation of inter-observer agreement, observers watched the same 33 videos of cows walking and assigned a locomotion score to each cow. The percentage agreement among all 3 observers and kappa (κ), weighted κ (Cohen, 1968), and prevalence-adjusted

bias-adjusted κ (**PABAK**; Byrt et al., 1993) values between each observer were calculated for each locomotion score. Percentage agreement and κ were calculated for classification of cows as lame (locomotion score ≥ 3).

Data Collection

Cows were locomotion scored as they exited the milking parlor except for 2 pens, 1 on farm 4 and 1 on farm 5, in which cows were locomotion scored while walking in the pen. The 3 sampling strategies were then modeled using the locomotion score data set. The order that cows were observed in the pen was used to model one of the sampling strategies, so cows were observed in their pens while they were restrained in headlocks along the feedbunk, and identification number (ID) was listed by pen in the order observed. Relative order of parlor exit was determined from the order in which ID number and locomotion score were recorded by observers. Individual cow DIM, reproductive status, parity, breed, and pen number were obtained. Data were collected in May and June 2011.

Data Analysis

Descriptive statistics and sampling strategy calculations were completed using a computerized spreadsheet program (Excel, Microsoft Corp., Redmond, WA). A multivariate logistic regression analysis for lameness risk factors was accomplished using the SAS statistical software package (version 9, SAS Institute Inc., Cary, NC) that included all cows for which milking parlor exit order and breed data were available. Exit order was standardized across different size pens by dividing the relative order of parlor exit for each cow by the total number observed in the pen to calculate relative exit order percentile. Exit order percentile category was used as a variable in the logistic regression model in addition to potential risk factors of herd, parity, breed, and DIM category. Milking order percentile categories were 1 to 10, 11 to 20, 21 to 30, 31 to 40, 41 to 50, 51 to 60, 61 to 70, 71 to 80, 81 to 90, and 91 to 100. The DIM were categorized as 1 to 100, 101 to 200, 201 to 300, 301 to 400, and >400 d. The final logistic regression model was determined using backward selection. Those variables with $P > 0.10$ were dropped from the model. Confidence interval of the proportion of lame (locomotion score ≥ 3) cows for herd and for each pen was calculated and used to compare sample prevalence to whole herd prevalence.

Sampling Strategy 1: Middle of Exit Order. Herd- and pen-level lameness prevalence (% locomotion score ≥ 3) were calculated for all cows scored and also from a sample of cows in the middle of the parlor exit

order. Sample number was determined using a quadratic equation ($R^2 = 0.31$, $P < 0.001$) published by Main et al. (2010): Sample size = $-0.001n^2 + 0.498n + 6.785$, where n = number of cows in the pen. This equation was based on the number of cows from the middle of the milking order needed to estimate prevalence within 5% of true prevalence on 67 UK farms. In the present study, a sample was taken from the milking order data set for each pen. The start and end of the sample was calculated for each pen by finding the cow in the middle of the exit order, subtracting half of the sample size to determine the first cow in the sample, and adding half the sample size to determine the last cow in the sample. Lameness prevalence within the sample was calculated for each pen and the pen samples were pooled to calculate herd sample prevalence. On 2 of the farms, farm 4 and farm 5, an entire pen was omitted from the exit order sampling strategy test because of no observations at the milking parlor exit (cows were scored in the pen). These pens were removed from the true and sample prevalence calculations for the purpose of comparison for exit order strategy only.

Sampling Strategy 2: Distributed Sample. Sample size for each herd was calculated based on the formula below for estimating prevalence based on a finite population (N), where $Z = 1.96$ (95% CI), expected prevalence (P) = 50%, and precision (d) = 10%:

$$\text{Herd sample size} = \frac{NZ^2P(1-P)}{d^2(N-1) + Z^2P(1-P)}.$$

The number of cows sampled per pen was weighted by the proportion of the lactating herd in that pen. The sample weighted by pen was then modeled in 2 ways. First, selection of cows for the sample was based upon the order they were observed restrained in stanchions. The total number of cows per pen was divided by the sample size for that pen to determine the number of cows (n) to count between selections of an individual cow, such that locomotion score is recorded for every n th cow that was observed as cows were restrained in a row of stanchions at the feed bunk. For example, if there were 100 cows in the pen and 20 needed to be sampled, a locomotion score was recorded for every fifth cow observed in that pen. The data were sorted by order in which they were restrained in stanchions and cows in the pen that were not restrained were sorted by random number (generated by the RAND function in Microsoft Excel) at the end of the data set, and the locomotion score of every n th cow was recorded. Second, to model selection of cows while they were loose in their housing pen, cows were selected randomly at the pen level for inclusion in the sample. The SURVEY-

SELECT procedure of SAS (v. 9.3 SAS Institute Inc.) was used to simulate 100,000 random samples of cows within each herd. Sampling was performed within farm stratified by pen with sample size in a pen weighted based on proportional contribution of each pen to total milking cow herd number. The mean, standard deviation, and variance of percentage lame and percentage in each locomotion score, within farm, of the 100,000 random samples were calculated using the UNIVARIATE procedure of SAS (v. 9.3, SAS Institute Inc.).

Sampling Strategy 3: Selected Pens Based on Production Level. Sampling all cows in 1 high-production, 1 low-production, and 1 hospital pen was tested as a scheme for estimating prevalence. Production data by pen were not collected, so pens were designated based on pen-average DIM compared with typical lactation peak milk production: pens with average DIM 60–120 d were selected as the high-production pen; if no pen averages fell in this range, the pen with the lowest average DIM that was greater than 60 was selected. The pen with the highest average DIM was selected as the low-production pen, and the hospital pen was identified by the farm manager. Farm 2 and farm 4 did not have cows in a designated hospital pen at the time of scoring, so only high- and low-production pens were selected. Farm 3 had 3 pens total, so all pens were included in this sample. To determine the sample prevalence for strategy 3, locomotion score data from all 3 pens (high production, low production, and hospital) were pooled and the proportion of locomotion score ≥ 3 was calculated.

RESULTS AND DISCUSSION

Observer Agreement

Among all 3 observers, percentage agreement was 83% for actual locomotion score and 96% for lame versus not lame. The κ statistic for locomotion score was 0.4 to 0.5 between scorers without linear weighting and 0.57 to 0.68 with linear weighting. The κ statistic for lame versus not lame was 0.72 to 0.80. The PABAK for all 3 observers was 0.79. The inter-observer agreement for the 3 observers in this study was similar to inter-observer agreement previously reported for locomotion scores. O'Callaghan et al. (2003) reported inter-observer agreement for locomotion scores of 30%, with 66% near agreement. Winckler and Willen (2001) reported 68% inter-observer agreement for locomotion score, with 30% differing by only one score. Weighted κ for locomotion score reported by Bicalho et al. (2007) was 0.45 to 0.48 between observers. The inter-observer agreement for lame/not lame classification in the current study was within the range of 83.9 to 96.8% previ-

ously reported by Main et al. (2010). The PABAK in the present study was lower than the 0.84 reported by von Keyserlingk et al. (2012), although not significantly so ($P > 0.05$) as it was within the 95% confidence interval reported in that study (0.77 to 0.90).

Distribution of Locomotion Scores

Herd size ranged from 148 to 2,744 cows. Farm 1 had 18% Holstein, 59% Jersey, and 23% Holstein-Jersey crossbred cows. Farm 2 had all Jersey breed cows. Farms 3 and 5 had herds of only Holstein cows, and farm 4 had 82% Holstein, 5% Jersey, and 13% Holstein-Jersey crossbred cows, 5, and 13%, respectively. All cows were housed in freestall housing except for 2 open-lot pens in herd 1 (335 cows) and 1 loose-housed straw-bedded pen in herd 5 (29 cows). Of all the herds, 661 cows were not locomotion scored because they passed observers at a gait faster than a walk and 47 cows were deleted from the data set because of missing parity information. Information on 35 cows was deleted because of inconsistency in pen number between farm records and observations that could not be explained by pen movement procedures.

Of all cows for which locomotion score data were collected ($n = 4,422$), 48% were locomotion score 1 ($n = 2,130$), 34% were locomotion score 2 ($n = 1,501$), 13% were locomotion score 3 ($n = 578$), 4.6% were locomotion score 4 ($n = 203$), and 0.22% were locomotion score 5 ($n = 10$). Prevalence of locomotion score 3 ranged by herd from 4.7 to 26%. Overall lameness prevalence (locomotion score ≥ 3) was 10, 9.9, 37, 21, and 35% for herds 1 to 5, respectively, and ranged by pen from 3 to 69%. Overall prevalence of severe lameness (locomotion score ≥ 4) was 1.7, 5.6, 13, 8.6, and 8.7% for herds 1 to 5, respectively, and ranged by pen from 0 to 31%.

Pen-level prevalence for each herd is reported in Table 1. Out of 29 total pens, prevalence was different from herd prevalence ($P < 0.05$) in 15 of them. For one farm (farm 3), pen-level prevalence and herd-level prevalence did not differ ($P > 0.05$). Across herds, pens classified as "breeding" or "early lactation" ($n = 19$) had an average prevalence of 18.7% (range of 4–54%) and pens classified as "pregnant cows" or "late lactation" ($n = 7$) had an average prevalence of 22% (range of 2.5–51%). Pens for which the average DIM was less than 150 ($n = 15$) had an average prevalence of 25% (range of 4–75%) and pens for which the average DIM was ≥ 150 ($n = 15$) had an average prevalence of 23% (range of 2.5–51%). Pens for which average parity was less than 2.5 ($n = 16$) had an average lameness prevalence of 25% (range of 2.5–69%), and pens for which average parity was ≥ 2.5 ($n = 14$) had an average lameness prevalence of 22% (range of 5.9–75%).

Table 1. Pen-level lameness prevalence (% locomotion score ≥ 3), pen description, and pen-average DIM and parity for 5 dairy herds

Herd	Pen description ¹	Cows locomotion scored	Average DIM	Average parity	Pen prevalence, % (95% CI)
1. Herd prevalence 10%	Early lactation	216	29.5	2.9	8 (5.3–13)
	AI breeding	247	122	1.3	4* (2.1–7.4)
	AI breeding	265	116	3.0	9 (6.1–13)
	Breeding, Holsteins	274	119	3.8	23* (18–28)
	Pregnant	249	238	3.4	11 (7.9–16)
	AI breeding	272	121	3.5	6* (3.6–9.4)
	Pregnant	278	239	2.3	3* (1.1–5.2)
	AI breeding	238	332	2.9	6 (3.8–10)
	AI breeding	264	161	3.3	15* (11–20)
	Hospital	71	172	2.9	23* (14–34)
2. Herd prevalence 9.9%	Early lactation/AI breeding	96	138	2.3	8 (4.1–16)
	Early lactation/AI breeding	104	148	3.4	16* (10–25)
	1st lactation/AI breeding	93	152	1.1	6 (2.7–14)
	Pregnant/bull breeding	82	295	2.7	6 (2.3–14)
3. Herd prevalence 37%	Hospital	4	100	4.3	75 (29–97)
	All others lactating	66	173	3.2	35 (24–47)
	All others lactating	65	198	2.2	37 (26–49)
4. Herd prevalence 21%	AI breeding	154	188	2.5	22 (16–29)
	AI breeding	154	142	2.2	13* (8.5–19)
	Early lactation	108	37	2.1	6* (3.0–13)
	Pregnant, culls	131	293	2.2	42* (34–51)
5. Herd prevalence 34%	Early pregnant	139	161	1.7	33 (26–41)
	Early lactation	83	27.3	2.0	33 (23–43)
	Late pregnant/bull breeding	135	326.4	1.8	30 (23–39)
	Late pregnant	100	346.4	2.1	51* (41–61)
	AI breeding	135	145.8	2.2	41 (33–49)
	AI breeding	98	144.8	1.1	13* (7.8–22)
	AI breeding	136	151.0	1.4	18* (12–25)
	AI breeding, older cows	135	144.6	3.3	54* (46–62)
	Hospital	29	123	2.3	69* (51–83)

¹Pen description given by herd owner or manager at the time of locomotion scoring.

*Pen-level prevalence differed from whole-herd prevalence ($P < 0.05$).

Lameness prevalence in herds examined in this study was within the range identified in previous studies of freestall-housed dairy cattle using similar locomotion scoring systems. Cook (2003) found a prevalence range of 7.3 to 35% for score 3 and 0 to 17% for score 4 in Wisconsin herds. Espejo et al. (2006) found a prevalence range of 3 to 39% for score 3 and 0 to 18% for score 4 in high-production pens in Minnesota. von Keyserlingk et al. (2012) reported an average clinical lameness prevalence of about 30% in British Columbia and California and 55% in the northeastern United States and a severe lameness prevalence of 4% in California and 8% in both the northeastern United States and British Columbia. Both of these prevalence measurements had great variability within regions. Pen-level lameness prevalence for all pens within the herds was not reported in these previous studies, and only the high-production pen was evaluated in 2 of them (Espejo et al., 2006; von Keyserlingk et al., 2012). Based on the ranges of pen-level prevalence in the present study, pen-level prevalence appears to be variable regardless of pen designation or average DIM or parity status of the cows in the pen.

Although the above-mentioned studies used similar locomotion scoring criteria to that in the present study, differences among gait scoring methods exist, especially in the definition for locomotion score 3. Cook (2003) describes a score 3 as “makes deliberate short steps with an arched back” walking and standing, whereas a locomotion score 2 makes “shorter strides with an arched back” and “stands with a level back.” In the present study, the cow would be considered a locomotion score 3 if a short stride was noted. von Keyserlingk et al. (2012) utilized the gait scoring system described by Flower and Weary (2006), which includes as criteria for locomotion score 3 an arched back, failure of hind hooves to track up, and asymmetric gait similar to the present study, but they additionally include “joints show signs of stiffness,” “slight limp can be discerned,” and “steady head carriage” (as opposed to head bobbing movement). The addition of these criteria may mean that less emphasis is placed on back posture for this scoring system compared with scoring systems such as that used in the present study that are based upon criteria described by Sprecher et al. (1997). Criteria for defining severe lameness (locomotion score

Table 2. Final logistic regression model describing variables associated with lameness (locomotion score ≥ 3) in 5 dairy herds¹

Variable	Level	Odds ratio	95% CI		P-value
			Lower limit	Upper limit	
Parity	1	Referent			<0.001
	2	2.09	1.58	2.78	
	3+	6.17	4.76	8.01	
Herd	1	Referent			<0.001
	3	2.12	1.37	3.50	
	2	2.08	1.33	3.24	
	4 ²	0.92	0.63	1.36	
	5 ²	3.03	2.19	4.20	
Breed	Jersey	Referent			<0.001
	Holstein	4.62	3.16	6.75	
	Holstein-Jersey cross	1.28	0.80	2.03	
Milking order percentile	10	Referent			<0.001
	20	1.501	0.981	2.296	
	30	1.312	0.852	2.021	
	40	1.097	0.703	1.711	
	50	1.473	0.960	2.261	
	60	1.460	0.953	2.234	
	70	1.437	0.940	2.197	
	80	2.083	1.378	3.148	
	90	1.569	1.029	2.392	
	100	2.208	1.462	3.333	

¹n = 3,433; cows without milking order or breed information were omitted from this analysis.

²One entire pen omitted due to missing milking order information.

≥ 4) compared with definition of locomotion score 3 are more similar across studies and possibly less subject to observer variability, as in all the previously mentioned prevalence studies locomotion score 4 and above is considered to mean decreased weight bearing on a limb. These differences in locomotion scoring criteria should lead to cautious interpretation of direct comparisons of lameness prevalence in different studies. Nonetheless, lameness prevalence results are similar across studies and most of the criteria for gait evaluation utilized in the above-mentioned studies have been subjected

to validation by correlation with hoof pathology or nociception (Flower and Weary, 2006; Tadich et al., 2010). In a review of dairy cattle gait assessment, the advantages of gait scoring systems (easy to do on farm) and disadvantages of gait scoring systems (inter- and intra-observer reliability) and variability of predictive value for hoof or leg pathology (Flower and Weary, 2009) highlight the need to better define their use—for research or for on-farm assessment.

Because cows were not locomotion scored if they moved past observers faster than a walk and it is un-

Table 3. Herd lameness prevalence (% locomotion score ≥ 3) on 5 dairies obtained by 3 different sampling strategies (exit order, distributed, selected pens)

Herd	Whole-herd prevalence (no.)	Exit order ¹		Distributed ²		Selected pens ³	
		Prevalence (no.)	95% CI	Prevalence (no.)	95% CI	Prevalence (no.)	95% CI
1	0.10 (2,374)	0.09 (657)	0.07–0.12	0.05 (97)	0.02–0.12	0.096 (574)	0.074–0.12
2	0.099 (375)	0.12 (188)	0.08–0.17	0.10 (80)	0.05–0.19	0.073 (178)	0.042–0.12
3	0.37 (135)	0.36 (80)	0.27–0.47	0.40 (59)	0.28–0.53	0.37 (135)	0.29–0.45
4	0.21 ⁴ (547)	0.14 ⁴ (169)	0.09–0.20	0.17 (85)	0.10–0.27	0.26 (285)	0.22–0.32
5	0.35 ⁵ (991)	0.34 ⁵ (417)	0.30–0.39	0.34 (90)	0.25–0.45	0.48* (264)	0.42–0.54

¹Sample size calculated per pen sampled from the middle of parlor exit order.

²Sample size calculated by herd and sample distributed evenly across and within pens based on order observed while cows restrained in stanchions.

³Sample of all cows in high-production, low-production, and hospital pens.

⁴Herd prevalence for the population used in the exit order sampling strategy comparison was 0.15 because 1 pen of cows (n = 131) was excluded from the exit order test due to no observation at milking time.

⁵Herd prevalence for the population used in the exit order sampling strategy comparison was 0.34 because 1 pen of cows (n = 29) was excluded from the exit order test due to no observation at milking time.

*P < 0.05 compared with whole-herd prevalence in the same row.

likely for cows with lameness to move at a faster gait, it is possible that the lameness prevalence measured in the current study was overestimated. However, the findings were similar to previous studies of lameness prevalence so this was likely not a large effect.

Variables Associated with Lameness

Parity, herd, breed, and milking order percentile were significantly ($P < 0.001$) associated with locomotion score ≥ 3 at the cow level in a final logistic regression model (Table 2). The DIM group was associated with locomotion score ≥ 3 ($P = 0.002$) in univariate analysis but not in the final multivariate analysis.

Cows of the Holstein breed were more likely to be lame compared with cows of the Jersey breed (Table 2). This result is similar to reports in the literature. A case study comparing performance of Holstein pens to a pen of Jersey and Holstein $\times$ Jersey crossbred cows found that Holsteins had 13 percentage points more lameness over a 12-mo period (Anderson et al., 2007). Alban (1995) found that Danish-Jersey cows had the lowest risk of lameness compared with other breeds.

Cows that were in their second or greater lactation had higher odds of lameness (Table 2). This association is similar to that of previous reports (Wells et al., 1993; Sogstad et al., 2005). However, in a Danish study (Alban, 1995), primiparous cows had a higher risk of being diagnosed with lameness compared with parity 2 or 3 cows, but parity 4 cows had the greatest risk. This may be due to differing etiologies of lameness in different areas.

Cows exiting the parlor in the last part of the pen (milking order percentile > 80) had greater odds of being lame than cows exiting in the first 10% of cows in the pen (Table 2). This positive association between milking order and lameness is similar to the results of Main et al. (2010), who found that the median percentage rank milking order was higher for cows with increased locomotion score.

Sampling Strategies

Sampling Strategy 1: Middle of Exit Order. Prevalence estimates obtained from the milking parlor exit order sampling (strategy 1) did not differ from true prevalence ($P > 0.05$) and the numeric differences did not exceed 2.4 percentage points (Table 3). On a pen level, sample prevalence of lameness using the milking order sample strategy was, on average, 3 percentage points different from whole-pen lameness prevalence (range 0 to 6 percentage points).

This sampling strategy was developed by Main et al. (2010) on a herd-level basis. However, 37 out of

the 67 herds investigated in that study were housed in only one group. The present study shows that the same sample size calculation and middle of milking parlor exit order sample taken on a pen level can be used to accurately estimate herd lameness prevalence in large herds with multiple pens. Additionally, the percentage difference from the true prevalence per pen was low so this sampling strategy may be acceptable for pen-level prevalence determination. A limitation of this sampling strategy on large farms is that it requires locomotion scoring over the course of an entire milking shift, a similar time commitment to locomotion scoring every cow at milking parlor exit but unlike locomotion scoring all cows, creating a complete list of lame cows requiring treatment is not accomplished.

Sampling Strategy 2: Distributed Sample. The sample prevalence obtained using the distributed sampling (strategy 2) did not differ from true prevalence ($P > 0.05$), but numeric differences were found, with the greatest being 5 percentage points from herd prevalence in herd 1 (Table 3). Sample prevalence of lameness on a pen level using the distributed sample strategy was, on average, 7.5 percentage points different from pen prevalence, ranging from 0.03 to 31 percentage points. The prevalence of each locomotion score within the sample population using the distributed sampling strategy was compared with the whole-herd prevalence of each locomotion score, and found to differ by 2.5 percentage points on average, with a minimum difference of 0 percentage points. The maximum difference was 12 percentage points for the percentage locomotion score 1 in herd 1; for all other herds, the sample locomotion score prevalence did not differ by more than 7 percentage points.

In a simulation of 100,000 repeated samples of each herd's data using the distributed sampling strategy, the mean lameness prevalence of all of the repeated samples differed by less than 1 percentage point from the herd prevalence for all 5 herds. Of all the repeated samples, 95% were within 10 percentage points of the herd prevalence and 80% were within 6 percentage points of the true prevalence. The minimum and maximum sample prevalence obtained during repeated sampling differed from the herd prevalence by as much as 21 percentage points. The standard deviation for herd prevalence was between 2.9 and 4.6. The variance was between 8.7 and 21%. The standard deviation for prevalence of each locomotion score in each herd was between 0 and 5 and the variance was between 0 and 25%.

Sampling strategy 2, using a calculated sample size distributed across the lactating herd, required the lowest number of total cows to locomotion score and estimated lameness prevalence within 5 percentage points of whole-herd prevalence in the 5 study herds. However,

because accurate pen-level prevalence was not obtained by this sampling strategy, this strategy would only be useful for herd-level assessment.

The purpose of testing sampling strategy 2 was to test the method for sampling cows, not the sample size calculation itself, because sample size calculations assume random sampling of a population, and the method for sampling described by National Milk Producers Federation (2010) is systematic and nonrandom. Although this empirical test on only 5 farms does not have adequate power to fully assess the accuracy of this systematic sampling method, it provides examples of its results compared with a known prevalence for each population, which indicates that it may have merit for estimation of herd lameness prevalence.

It would be most convenient in implementing the distributed sampling strategy to sample across the pen at any time including when cows are loose in the pen or when cows are locked in stanchions by releasing selected individuals. Ideally, to validate these processes, the sampling would have to be performed in all the different ways on many farms. Time and cost restrictions prevented this and simulations using the data sets from the 5 farms were performed instead. Therefore, 2 models of sampling were completed using the data set: one to simulate systematic selection of cows as they are locked in stanchions along the feed bunk, and the other to simulate selection of cows as they are loose in their pen. Both of these tests use the nonrandom sampling method of weighting the sample size by pen size, but they use different methods of selecting cows within the pen. The first test using this sampling strategy used the pen order data (the order in which cows were observed standing in a row of stanchions at the feed bunk) to select cows according to the described systematic sampling method. The second used random selection of cows within the pen as a proxy for systematically selecting cows loose within a pen, and to address the issue of a low number of sample repetitions, computer modeling was used to take repeated samples from each of the 5 farms. Therefore, it was assumed that randomly selecting cows from the data set would be similar to selecting cows loosely distributed throughout the pen. However, cows may not be randomly distributed when they are loose in the pens. Juarez et al. (2003) found that locomotion score varied with distance from the pen entrance. Even in cases of unequal location of lame cows within a pen, systematic selection of cows evenly distributed across the pen by selecting every n th cow should minimize sampling bias. Indeed, modeled sampling based upon random order and order restrained in stanchions yielded average prevalence of each locomotion score that were not different from true herd prevalence on the 5 farms included in this study.

Repeated random sampling of all 5 farms demonstrated the repeatability of the distributed sampling strategy. Frequency distributions from this repeated sampling suggest that one can be 95% confident that the sample prevalence is within 10 percentage points of the true prevalence and 80% confident that the sample prevalence is within 6 percentage points of the true prevalence. We suggest that this level of confidence in the sample proportion is adequate for characterization of the risk level of a herd compared with averages or goal benchmarks if the sample prevalence is greater than 6 to 10 percentage points different from the goal herd prevalence level. Monitoring herd lameness prevalence has been used in interventions to reduce dairy herd lameness (Main et al., 2012). However, with this level of variability, this sampling strategy would not be useful for monitoring small (less than 6–10 percentage points) changes in lameness prevalence over time within a herd. Additionally, it should be used with caution as an auditing tool if there are negative consequences of incorrect classification.

Sampling Strategy 3: Selected Pens Based on Production Level. Prevalence obtained using sampling strategy 3 was not different from whole-herd prevalence for 4 out of the 5 farms ($P > 0.05$). Farm 3 had 3 pens total so the sample using this strategy included the entire herd and was therefore equal to whole-herd prevalence. On the one farm on which sample prevalence differed (farm 5), it was 13 percentage points higher than whole herd prevalence. If only the high and low production pens were assessed (hospital pen excluded) for those farms that had designated hospital pens (farms 1, 3, and 5), the sample prevalence was not different from herd prevalence for farms 1 and 3 and was different for farm 5 (45%, 95% CI: 39–52%). Therefore, inclusion of the hospital pen did not significantly change the results of this sampling strategy for these 5 herds. Of all the potentially high-producing pens, nonhospital pens for which average DIM was 30 to 150, lameness prevalence was higher in 4 pens, lower in 4 pens, and the same in 3 pens as the prevalence in the whole herd.

The inconsistency of sampling strategy 3 for predicting whole-herd prevalence is most likely due to the variability in lameness prevalence by pen, as more than half the pens in this study had lameness prevalence different from herd prevalence ($P < 0.05$). Housing is known to affect lameness prevalence (Cook and Nordlund, 2009) so if housing design differs among pens, that could be a source of confounding if only select pens are sampled to estimate herd prevalence. Additionally, sampling cows based on pen designation allows for more possible confounding of DIM, parity, and breed, which were found to be associated with lameness in this or other studies (Wells et al., 1993; Alban, 1995; Sogstad et al., 2005;

Espejo et al., 2006; Anderson et al., 2007). Further, this strategy would be difficult to standardize across farms because of the variability in how pens are designated. For example, none of the farms participating in this study designated pens based on production level.

CONCLUSIONS

The first step in reducing lameness on a dairy farm is to determine the level within the herd. Estimating lameness prevalence in a dairy herd may have utility for setting herd health goals or for animal welfare assessment. Both the sampling strategy using parlor exit order and that using a sample distributed across pens estimated herd prevalence within 5 percentage points of true prevalence on the 5 farms sampled, and estimates were not statistically different from true prevalence. The data presented in this study suggest that estimation of herd-level lameness prevalence for managerial evaluation can be achieved using these sampling strategies. However, the estimate should be interpreted carefully, with consideration of confidence level. Distributed sampling across the pens requires locomotion scoring fewer cows than the milking parlor exit order strategy and does not depend on milking time for evaluation so could be completed at times convenient to the observer. Because not all cows are locomotion scored in the sampling strategies presented here, they do not allow for identification of all lame cows in the herd for treatment or epidemiologic analysis of the disease at the cow level. Although these data from the sample prevalence of 5 farms provide evidence for methods of selecting cows to be locomotion scored, they do not validate or provide support for the locomotion scoring system or the definition of lameness. However, because the prevalence of each locomotion score group in the sample was accurate using the distributed strategy, a modification of the scoring system or changing the cutoff for lameness definition would likely still yield accurate prevalence estimation results. Reducing the number of cows to locomotion score and the need to be present for an entire milking shift would decrease the time needed for lameness prevalence assessment and therefore minimize cost of assessment.

ACKNOWLEDGMENTS

This project was supported by the Washington State Dairy Federation (Elma) and Washington Dairy Products Commission (Lynnwood). The authors thank the collaborating dairy producers and veterinarians Fred Muller (Sunnyside, WA), Robert Holt and Jacob Steiger (Lynden, WA), and Mike Anderson and Gale DeJong (Lynden, WA) for sharing their valuable time and input to the project.

REFERENCES

- Alban, L. 1995. Lameness in Danish dairy cows: Frequency and possible risk factors. *Prev. Vet. Med.* 22:213–225.
- Anderson, T., R. Shaver, P. Bosma, and V. De Boer. 2007. Case study: Performance of lactating Jersey and Jersey-Holstein crossbred versus Holstein cows in a Wisconsin confinement dairy herd. *Prof. Anim. Sci.* 23:541–545.
- Bicalho, R. C., S. H. Cheong, G. Cramer, and C. L. Guard. 2007. Association between a visual and an automated locomotion score in lactating Holstein cows. *J. Dairy Sci.* 90:3294–3300.
- 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.
- Byrt, T., J. Bishop, and J. B. Carlin. 1993. Bias, prevalence and kappa. *J. Clin. Epidemiol.* 46:423–429.
- Cha, E., J. A. Hertla, D. Barb, and Y. T. Gröhn. 2010. The cost of different types of lameness in dairy cows calculated by dynamic programming. *Prev. Vet. Med.* 97:1–8.
- Cohen, J. 1968. Weighted kappa: Nominal scale agreement with provision for scales disagreement of partial credit. *Psychol. Bull.* 70:213–220.
- Cook, N. B. 2003. Prevalence of lameness among dairy cattle in Wisconsin as a function of housing type and stall surface. *J. Am. Vet. Med. Assoc.* 223:1324–1328.
- Cook, N. B., and K. V. Nordlund. 2009. The influence of the environment on dairy cow behavior, claw health, and herd lameness dynamics. *Vet. J.* 179:360–369.
- Dyer, R. M., N. K. Neerchal, U. Tasch, Y. Wu, P. Dyer, and P. G. Rajkondawar. 2007. Objective determination of claw pain and its relationship to limb locomotion score in dairy cattle. *J. Dairy Sci.* 90:4592–4602.
- 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.
- 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.
- Flower, F. C., and D. M. Weary. 2009. Gait assessment in dairy cattle. *Animal* 3:87–95.
- Garbarino, E. J., J. A. Hernandez, J. K. Shearer, C. A. Risco, and W. W. Thatcher. 2004. Effect of lameness on ovarian activity in postpartum Holstein cows. *J. Dairy Sci.* 87:4123–4131.
- Hernandez, J. A., E. J. Garbarino, J. K. Shearer, C. A. Risco, and W. W. Thatcher. 2005. Comparison of milk yield in dairy cows with different degrees of lameness. *J. Am. Vet. Med. Assoc.* 227:1292–1296.
- Hernandez, J. A., J. K. Shearer, and D. W. Webb. 2001. Effect of lameness on the calving-to-conception interval in dairy cows. *J. Am. Vet. Med. Assoc.* 218:1611–1614.
- Juarez, S. T., P. H. Robinson, E. J. DePeters, and E. O. Price. 2003. Impact of lameness on behavior and productivity of lactating Holstein cows. *Appl. Anim. Behav. Sci.* 83:1–14.
- Main, D. C. J., Z. E. Barker, K. A. Leach, N. J. Bell, H. R. Whay, and W. J. Browne. 2010. Sampling strategies for monitoring lameness in dairy cattle. *J. Dairy Sci.* 93:1970–1978.
- Main, D. C. J., K. A. Leach, Z. E. Barker, A. K. Sedgwick, C. M. Maggs, N. J. Bell, and H. R. Whay. 2012. Evaluating an intervention to reduce lameness in dairy cattle. *J. Dairy Sci.* 95:2946–2954.
- National Milk Producers Federation. 2010. Animal Care Manual. National Dairy Farmers Assuring Responsible Management (FARM) Program. Dairy Management Inc., Rosemont, IL.
- Nordlund, K. V., N. B. Cook, and G. R. Oetzel. 2004. Investigation strategies for laminitis problem herds. *J. Dairy Sci.* 87(E. Suppl.):E27–E35.
- O’Callaghan, K. A., P. J. Cripps, D. Y. Downham, and R. D. Murray. 2003. Subjective and objective assessment of pain and discomfort due to lameness in dairy cattle. *Anim. Welf.* 12:605–610.
- Sogstad, A. M., T. Fjeldaas, and O. Østeras. 2005. Lameness and claw lesions of the Norwegian Red dairy cattle housed in free stalls in relation to environment, parity, and stage of lactation. *Acta Vet. Scand.* 46:203–217.

- Sprecher, D. J., D. E. Hostetler, and J. B. Kaneene. 1997. A lameness scoring system that uses posture and gait to predict dairy cattle reproductive performance. *Theriogenology* 47:1179–1187.
- Tadich, N., E. Flor, and L. Green. 2010. Associations between hoof lesions and locomotion score in 1098 unsound dairy cows. *Vet. J.* 184:60–65.
- Thomas, C. V., M. A. DeLorenzo, and D. R. Bray. 1996. A network simulation model of large herringbone and parallel milking parlors. *J. Dairy Sci.* 79:1960–1971.
- Thomsen, P. T., L. Munksgaard, and J. T. Sørensen. 2012. Locomotion scores and lying behavior are indicators of hoof lesions in dairy cows. *Vet. J.* 193:644–647.
- 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.
- 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.
- Wells, S. J., A. M. Trent, W. E. Marsh, and R. A. Robinson. 1993. Prevalence and severity of lameness in lactating cows in a sample of Minnesota and Wisconsin herds. *J. Am. Vet. Med. Assoc.* 202:78–82.
- Whay, H. R., A. E. Waterman, and J. F. Webster. 1997. Associations between locomotion, claw lesions and nociceptive threshold in dairy heifers during the peri-partum period. *Vet. J.* 154:155–161.
- Winckler, C., and S. Willen. 2001. The reliability and repeatability of a lameness scoring system for use as an indicator of welfare in dairy cattle. *Acta Agric. Scand. A Anim. Sci.* 51:103–107.
- Zinpro Corp. 2010. Dairy Locomotion Scoring (video). Accessed May 10, 2011. <http://www.zinpro.com/video-library/dairy-locomotion-videos>.