We compared 2 methods for identifying lame cows and estimating the prevalence of lameness in tiestalls. Cows (n = 320) in 9 tiestall herds were scored as lame both by the presence of limping while walking and by stall lameness scores (SLS). The SLS was based on the number of the following behaviors that the cow showed while standing in the tiestall: weight shifting, standing on the edge of the stall, uneven weight bearing while standing, and uneven weight bearing while moving from side to side. Two observers watched video-recordings of the cows. Intraobserver agreements for the 4 SLS behaviors ranged from 92 to 100%, and interobserver agreement ranged from 81 to 100%. The overall prevalence of lameness based on an SLS of ≥2 was similar to that of limping (39 vs. 40%). The sensitivity of the classification based on the SLS was 0.63 and the specificity was 0.77 in identifying cows with a limp; accuracy varied across farms from 62.2 to 80.4%, with a mean of 71.7%. A cow with an SLS of ≥2 had 4.88 times the odds of limping than a cow with an SLS of <2. The prevalence of lameness on farms based on SLS was highly correlated with the prevalence of limping (Pearson correlation = 0.88; n = 9), and prevalence estimates from the 2 methods diverged most when the mean herd prevalence was lower. The SLS method provides an estimate of the prevalence of lameness in tiestall herds comparable with traditional gait scoring, but does not require that the cows be untied. The SLS method could be used to improve lameness detection on tiestall farms and obtain estimates of lameness prevalence without the need to walk the cows.

Key words: lameness, dairy cow, tiestall, welfare

Technical Note

Lameness is widely recognized as a serious welfare problem for dairy cows (von Keyserlingk et al., 2009). In freestalls, gait scoring is the most common method of identifying lame cows and is used to assess the prevalence of lameness for many animal welfare assessments (e.g., Whay et al., 2003; Welfare Quality, 2009) and for epidemiological research (e.g., Ito et al., 2010; von Keyserlingk et al., 2012). Lameness also affects cows in tiestalls (Cramer et al., 2008) but gait-scoring cows in tiestall barns can be challenging because the cows may not be used to walking and there may not be a suitable area in which to walk the cows. Recently, Leach et al. (2009) reported an alternative method of identifying lame cows kept in tiestalls, which did not require that the cows be walked. They evaluated the method using data from 98 cows in 4 dairy herds in Europe, and showed that the method had good sensitivity and specificity in identifying lame cows. Our objectives were to provide support for the use of this method by further testing the inter- and intraobserver repeatability of this method, and by comparing it with gait scoring as a means of assessing the prevalence of lameness in herds, using a larger number of dairy herds and cows in Canada.

The study was conducted on 320 Holstein cows in 9 herds on 7 tiestall farms located in Ontario and Quebec, Canada. The number of cows used in each herd varied between 13 and 68. All enrolled cows were released from their tiestalls every day and habituated to walk to access the milking parlor or to an outdoor exercise area. Each cow was scored for lameness while standing in the tiestall using the stall lameness score (SLS) adapted from Leach et al., (2009). First, the cow was encouraged to stand by an experienced cow handler, who stood behind the cow while saying “up up” in a stern voice. If the cow did not respond, she was tapped gently on the spine until she began to rise. Each cow was required to be standing for a minimum
of 3 min before the rest of the protocol was followed. The cow was filmed with the camcorder placed on a tripod. First, the cow’s hind legs were filmed for 10 s with the camera about 1 m behind the cow. Then, the camcorder was moved to the left of the cow to film a side view for 10 s, followed by filming from the right side for 10 s. The camera was then replaced behind the cow and the cow handler encouraged the cow to step from side to side 2 to 3 times. To do this, the cow handler first walked from the right to the left behind the cow in an attempt to encourage the cow to step in the opposite direction. The cow handler then repeated this in the opposite direction. If the cow did not respond to the handler’s movements, the movement was repeated while the handler tapped gently on the cow’s pin bone, on the side opposite to where the cow was being asked to move. If this still did not work, the cow was gently prodded using a pen on the pin bone. This process was repeated 2 to 3 times. The video operator continued to record for 1 min after the cow had completed the last step to the side.

To assess the gait of the cows, we used a simplified version of the gait-scoring system of Flower and Weary (2006), which involved identifying cows that walked with an obvious limp. This has been shown to be accurate in detecting hoof lesions (Flower and Weary, 2006; Chapinal et al., 2009) and is similar to the procedure used by Leach et al. (2009). A similar simplified version has been used previously to detect lame cows in commercial facilities (Ito et al., 2010). The cows were first habituated to being released from their tiestall and walked in a calm manner by an experienced handler down the passageway. This was repeated up to 3 times until each cow walked in a straight line at a steady pace for a minimum of 2 strides. Video-recordings of the cows were made while they were walking.

The videos were watched twice (at an interval of at least 6 wk) by 2 independent, trained observers who were not present at the time of recording. One observer watched the videos of all 320 cows and the second observer watched the videos of 195 cows. The observers first watched the SLS videos and then the gait videos. Analyses of SLS videos involved observing the hind legs for 4 distinctive behaviors described by Leach et al. (2009):

1. Weight shift: regular, repeated shifting of weight from one hoof to another, defined as lifting each hind hoof completely off the ground at least twice. The hoof had to be lifted and returned to the same location and did not include stepping forward or backward.
2. Stand on edge: the cow places one or more hooves on the edge of the stall while standing station-ary. This did not include times when both hind hooves were in the gutter or when the cow briefly placed her hoof on the edge during a movement or step.
3. Uneven weight: repeatedly resting one foot more than the other, indicated by the cow raising a part or the entire hoof off the ground. This did not include raising of the hoof to lick or during kicking.
4. Uneven movement: uneven weight bearing between feet when the cow was encouraged to move from side to side. This was demonstrated by a more rapid movement by one foot than the other or by an evident reluctance to bear weight on a particular foot.

The SLS for a cow was the number of behavioral indicators that were observed. A cow was judged as lame in 3 ways: by an SLS of at least 1 (i.e., the presence of at least 1 of the 4 behaviors), an SLS of at least 2 (i.e., the presence of 2 or more of the behaviors), or an SLS of at least 3 (i.e., the presence of 3 or more of the behaviors). Because the use of at least 2 behaviors was the most accurate (Table 1), we used this method in all analyses except where stated.

The videos of the cows walking were watched by the same 2 independent, trained observers, and a cow was judged as lame or not depending on whether she showed an obvious limp, which was defined as uneven weight bearing of one or more limbs (Flower and Weary, 2006; Chapinal et al., 2009). This is similar to the method used by Leach et al. (2009). If a cow walked at a trot or a run, then the video was disregarded and not included in the analysis.

Intra- and interobserver reliability was examined at the start, midway, and end of the video-watching period. The percentage exact agreement was calculated.

Table 1. Prevalence of lameness, sensitivity, specificity, and false positives and negatives when cows were judged as being lame or not using a stall lameness score (SLS) of at least 1, at least 2, or at least 3

<table>
<thead>
<tr>
<th>Item</th>
<th>SLS$^1$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>≥1</td>
</tr>
<tr>
<td>Prevalence (%)</td>
<td>74</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.87</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.34</td>
</tr>
<tr>
<td>False positives (%)</td>
<td>53</td>
</tr>
<tr>
<td>False negatives (%)</td>
<td>21</td>
</tr>
</tbody>
</table>

$^1$SLS $= \geq 1$: decision as to whether a cow was lame or not was based on the observed occurrence of at least 1 of the 4 behaviors; SLS $= \geq 2$: decision as to whether a cow was lame or not was based on the observed occurrence of at least 2 of the 4 behaviors; SLS $= \geq 3$: decision as to whether a cow was lame or not was based on the observed occurrence of at least 3 of the 4 behaviors.
between the 2 observers as number of exact agreements/total number of observations × 100, and the kappa coefficient was calculated. Intraobserver agreement for the 2 observers was ≥89% for limping while walking and 100% for weight shift, and ≥95% for stand on edge, 95% for uneven weight, and 92% for uneven movement in the SLS system. Based on the 195 cows that were successfully scored by both methods, interobserver agreement between the 2 observers was ≥89% for limping while walking, and was 100% for weight shift, 94% for stand on edge, 95% for uneven weight, 81% for uneven movement, and 89% for lameness using 2 behaviors in the SLS system. Based on the 195 cows that were successfully scored by both methods, interobserver agreement between the 2 observers was 96% for limping while walking, and was 100% for weight shift, 94% for stand on edge, 95% for uneven weight, 95% for uneven weight, and 92% for uneven movement in the SLS system. Based on the 195 cows that were successfully scored by both methods, interobserver agreement between the 2 observers was ≥89% for limping while walking, and was 100% for weight shift, 94% for stand on edge, 95% for uneven weight, 81% for uneven movement, and 89% for lameness using 2 behaviors in the SLS system. The correlations between the two observers in the estimates of prevalence of lameness were 0.96 (P < 0.001) using limping while walking, and 0.91 (P = 0.004) using the SLS system with 2 behaviors. The kappa coefficient for limping while walking was 0.97 (95% confidence limit = 0.95–1.00) and for lameness was 0.77 (95% confidence limit = 0.68–0.86) based on the SLS.

These estimates of intra- and interobserver reliability for scoring the occurrence of the behaviors were high, similar to those reported by Leach et al. (2009), and within the ranges often reported for gait scores taken while the cows are walking (Winckler and Willen, 2001; Flower and Weary, 2006). However, the kappa coefficients were somewhat lower for the SLS. Reliabilities were lowest for uneven weight bearing while moving from side to side, which depended more on the cow cooperating with the attempts of the handler to encourage her to step from side to side.

Exact agreement was also calculated between the number of cows in each herd based on the presence of limping while walking and the SLS. The prevalence of limping was 40%. The prevalence of lameness, the sensitivity and specificity, and the occurrence of false negatives and positives of the SLS are shown in Table 1. The accuracy (percentage of cows classified as lame or not lame) of SLS in identifying cows classified as lame or not lame based on limping ranged across the 9 herds from 62.2 to 80.4%, with a mean of 71.7%. A logistic regression showed that the odds ratio of a cow having a limp was 1.43 (CI = 0.73–2.80; P < 0.10) for an SLS of ≥1, 4.88 (CI = 2.73–8.62; P < 0.0001) for an SLS of ≥2, and 1.03 (CI = 0.38–2.80; P < 0.10) for an SLS of ≥3.

The sensitivity and specificity of the SLS system were highest when the occurrence of at least 2 of the behaviors was used to classify a cow as lame or not. Using the occurrence of at least 1 behavior gave a large number of false positives, whereas using the occurrence of at least 3 behaviors gave a large number of false negatives. When the occurrence of at least 2 behaviors was used, the scores gave an accurate estimate of the overall prevalence of limping. These results differed from those of Leach et al. (2009), who showed that the occurrence of only 1 of 4 behaviors provided the best estimates of lameness, whereas the use of 2 behaviors may overestimate lameness prevalence. This discrepancy may reflect the fact that Leach et al. (2009) used direct observation, whereas we used video-recordings, which may have allowed more detailed observation of the cows. A follow-up study is needed to address this issue by comparing live and video assessment conducted by the same pair of observers.

Figure 1 shows the prevalence of lameness in each herd based on the 2 methods. The Pearson correlations between the estimates of prevalence of lameness using limping while walking and the estimates using the SLS system was 0.88 (P = 0.002). The difference between the prevalence of lameness as measured by the SLS and that measured by limping (SLS − Limp) varied from −16.2 to 19.2%, with a median value of 5.6% (Q1 = 0%, Q3 = 8.6%, where Q = quartile). The SLS underestimated lameness prevalence in 2 herds and overestimated lameness prevalence in 6 herds (Figure 1).

The discrepancy between the 2 methods might be expected to be larger in smaller herds or in herds with a low prevalence of lameness, because misclassification of a small number of cows would have a greater effect. Within our small sample, we found that the size of the discrepancy was negatively correlated with the herd prevalence of lameness (r = −0.77; P = 0.02) but not with herd size (r = −0.24; P > 0.10). We can only speculate as to the other causes of the discrepancy. Possibly the different types of lesions in different locations in the hooves or injuries to the upper limb that are responsible for the lameness cause different levels of

![Figure 1](image-url)
discomfort depending on whether the cow is standing or walking.

Overall, our results support those of Leach et al. (2009) in showing that the method provides estimates of lameness comparable to gait scoring and thus it is a useful tool for estimating the prevalence of lameness on tie-stall farms either for research purposes or for purposes of animal welfare assessments. We found no strong evidence of bias, although overestimation of lameness prevalence, rather than an underestimation, was found on more farms, especially where the prevalence of lameness was low. However, the moderate values for sensitivity and specificity suggest that the method is less useful for identifying which particular cows are lame, although, on farms where no alternative method exists, 77% of cows that limp will be correctly identified. This compares favorably with the 20 to 30% accuracy often reported for dairy producers (Whay et al., 2003; Espejo et al., 2006). Because gait scores are themselves imperfectly associated with the presence of hoof lesions (e.g., Chapinal et al, 2009), further research is needed to determine how accurate the present method is at identifying the occurrence of hoof lesions.

ACKNOWLEDGMENTS

This study was funded by Agriculture and Agri-Food Canada (Agassiz, BC, Canada) and Dairy Farmers of Canada (Ottawa, ON, Canada) as part of the Dairy Science Cluster Initiative, Fonds Québécois de la Recherche sur la Nature et les Technologies (FQRNT; Québec city, QC, Canada) – Novalait (Québec city, QC, Canada) – Ministère de l’Agriculture, des Pêcheries et de l’Alimentation du Québec (MAPAQ; Québec city, QC, Canada), and by Alberta Milk. We especially thank Gosia Zdanowicz from Agriculture and Agri-Food Canada (Agassiz, BC, Canada) for video analysis and data handling. We thank the collaborators, students, and co-op students from Agriculture and Agri-Food Canada (Agassiz, BC, Canada), University of Guelph, (Guelph, ON, Canada), Université Laval (Québec City, QC, Canada), and Valacta (Sainte-Anne-de-Bellevue, QC, Canada) for data collection.

REFERENCES


