Taylor & Francis
Taylor & Francis Group

New Zealand Veterinary Journal

Te Hautaka Tākuta Kararehe o Aotearoa

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tnzv20

Can we estimate herd-level prevalence of lameness in dairy cow herds kept at pasture by sampling part of the herd?

S Sapkota, RA Laven, KR Müller & DA Yang

To cite this article: S Sapkota, RA Laven, KR Müller & DA Yang (26 Mar 2025): Can we estimate herd-level prevalence of lameness in dairy cow herds kept at pasture by sampling part of the herd?, New Zealand Veterinary Journal, DOI: 10.1080/00480169.2025.2478599

To link to this article: https://doi.org/10.1080/00480169.2025.2478599

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 26 Mar 2025.

Submit your article to this journal

Article views: 74

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at https://www.tandfonline.com/action/journalInformation?journalCode=tnzv20

NEW ZEALAND VETERINARY JOURNAL
https://doi.org/10.1080/00480169.2025.2478599

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS   Check for updates

Can we estimate herd-level prevalence of lameness in dairy cow herds kept at pasture by sampling part of the herd?

S Sapkota[a], RA Laven [a], KR Müller [a] and DA Yang [b]

[a]Tāwharau Ora – School of Veterinary Science, Massey University, Palmerston North, New Zealand; [b]College of Veterinary Medicine, Nanjing Agricultural University, Nanjing, People's Republic of China

ABSTRACT

Aims: To assess whether herd-level lameness prevalence can be estimated on New Zealand dairy farms, by scoring the first, middle, or last 100 cows in the milking order. In pasture-based herds, whole herd locomotion scoring requires an assessor outside the milking parlour throughout milking. If sufficiently predictive, sampling a proportion of the herd based on milking order, could reduce the costs and time of welfare assessments.

Methods: Six pasture-based, spring-calving, dairy farms in the Manawatū region of New Zealand were conveniently selected. Visits occurred at approximately 6-week intervals between October 2021 and May 2022. Cows were scored using the DairyNZ lameness score (0–3). The assessor tallied cows as they left the parlour and recorded the milking order of those with a lameness score ≥ 2. Data were analysed to determine the association between farm, visit and the proportion of lame cows in the first, middle, and last 100 cows, and the agreement between the prevalence of lame cows in those groups and from whole herd scoring.

Results: Across all visits, 263 lame cows were recorded. Of these, 40.7% were in the last 100, 25.9% in the middle 100, and 14.4% in the first 100. Farm, visit and their interactions with group were all statistically significant ($p < 0.001$). While, overall, the last 100 cows had the highest proportion of lame cows, this pattern varied across farms and visits, Limits-of-agreement plots showed that as herd prevalence increased, agreement between the prevalence in each sample group and herd prevalence worsened. When herd prevalence exceeded 5%, only the middle 100 sampling group had a limits-of-agreement < 5%.

Conclusions: Variations across farms and seasons in the proportion of lame cows in each part of the milking order lead to variations in the accuracy of predicting overall lameness from such samples. Based on limits-of-agreement, observing the middle 100 cows is likely to be the most accurate sample, but is still likely to be of limited value on New Zealand dairy farms, especially as a single, one-off measurement.

Clinical relevance: On New Zealand dairy farms, locomotion scoring the middle 100 cows in the milking order as part of a welfare assessment would reduce costs and time but would not produce an accurate estimate of whole-herd lameness prevalence. However, it may be useful as a screening tool in herds routinely locomotion scoring throughout the year.

Abbreviations: EMM: Estimated marginal mean

ARTICLE HISTORY
Received 28 August 2024
Accepted 6 March 2025

KEYWORDS
Dairy cattle; lameness; prevalence; locomotion score; predictive ability; welfare assessment

Introduction

Lameness in dairy cattle has a deleterious effect on fertility (Alawneh *et al.* 2011) and milk production (Green *et al.* 2002) and leads to increased management and treatment costs (Ózsvári 2017). Moreover, it poses a significant welfare issue, causing pain and distress that can impact the health and behaviour of affected animals (Whay and Shearer 2017).

In New Zealand, the most commonly used method for actively detecting lameness is visual locomotion scoring, with the DairyNZ lameness score (0–3; DairyNZ 2017) being the industry-recognised system. Typically, cows are scored after they are released from the milking parlour as they walk along a flat, usually concrete, surface (Fabian *et al.* 2014). Thus, to locomotion-score the whole herd necessitates an observer who is present outside the milking parlour for the whole duration of milking. On many New Zealand dairy farms, this may be difficult to achieve on a routine basis due to the high ratio of cows to staff (DairyNZ 2022).

In addition, estimating the herd prevalence of lameness by locomotion scoring is an essential part of the welfare assessment of dairy cows (Laven and Fabian 2016). Thus, when developing a time-limited welfare

CONTACT RA Laven r.laven@massey.ac.nz

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

assessment for New Zealand dairy farms focused on assessment during milking time, Sapkota *et al.* (2022) either had an additional assessor outside the milking parlour solely to locomotion-score, or scored lameness at one milking and assessed the remaining animal-based measures at a subsequent milking. The requirement to locomotion-score thus significantly increased the cost of the welfare assessment.

Alternatives to whole-herd locomotion scoring are required. One possible alternative is scoring cows for lameness based on stance and activity during milking ("in-parlour scoring"; Werema *et al.* 2022), which could allow a single assessor to record lameness alongside other animal-based welfare measures. However, this method needs further development and testing before it can be recommended as an alternative to locomotion scoring. The same applies to automatic scoring and recording of lameness. O'Leary *et al.* (2020) suggested that any automatic recording system would need to have > 90% sensitivity and 99% specificity to be of significant value for lameness management on most farms. No systems with proven accuracy that exceed these targets are currently available in New Zealand.

One method used to reduce the time taken for locomotion-scoring during welfare assessment is to use a sampling strategy (Sørensen *et al.* 2007). In housed cows, random sampling of a herd can be used, as individual cows can be easily selected and locomotion-scored alongside other welfare assessments (Van Os *et al.* 2019), but this process is more difficult in cows at pasture, at least in part because they need to be moved away from pasture to be locomotion-scored. Scoring after milking greatly simplifies the random sampling process, but still means that the assessor has to be present outside the milking parlour for the whole of the milking (and therefore does not reduce costs or time).

An alternative to random sampling is targeted sampling of a specific subgroup. Main *et al.* (2010) reported that sampling a maximum of 100 cows from the middle of the milking order in 36 herds with >100 cows resulted in an estimated lameness prevalence that was within 5% of the prevalence identified by locomotion-scoring the whole herd on 83% of farms (i.e. 30/36 UK dairy farms). Observing up to 100 cows from the middle of the milking sequence could thus be a viable sampling strategy. In a subsequent Australian study, Beggs *et al.* (2019) found that the prevalence of lameness in the last 200 cows in the milking order on 50 dairy farms was always within 5.4% of the overall prevalence. However, no similar studies have been reported under New Zealand conditions.

The objective of the present study was therefore to build on the locomotion score sampling strategy used by Main *et al.* (2010) and assess whether scoring the first 100, middle 100 or last 100 cows in the milking order could be used to estimate the herd-level prevalence of lameness on New Zealand dairy farms. There were two parts to this analysis. Firstly, to determine whether time and farm affected the relationship between where a cow was in the milking order (i.e. first, middle or last 100 cows) and the probability that it was lame; and secondly, to calculate the agreement between lameness prevalence as identified by whole-herd scoring and that identified by scoring only the first, middle or last 100 cows.

Materials and methods

A convenience selection was made of six spring-calving, pasture-based dairy farms (see Table 1) in the Manawatū district of the North Island of New Zealand. Between 15 October 2021 (early lactation) and 10 May 2022 (late lactation), the study farms were visited five times at approximately 6-week intervals, except for Farm 5, where the last visit did not occur.

Cows were locomotion-scored on a flat, hard surface, immediately on leaving the milking parlour after being milked. Cows were scored by author SS, using the DairyNZ lameness score (0–3; DairyNZ 2017). The number of cows scored was tallied and if any cow had a lameness score of ≥ 2, the number on the tally for that cow was recorded, resulting in, for each farm/visit combination, a list of the milking order of each lame cow. On farms where multiple milking herds were present (2 farms), only the first herd to be milked was included in the study. Thus, if the farm had a lame cow group, it was never included in this study, but if the farm did not have such a group, then cows recently identified as lame by farm staff would have been included. Data were not included in the analysis if fewer than 200 cows were lameness-scored at a visit to an individual farm.

Statistical analysis

For each farm/visit combination, the prevalence of lameness (score ≥ 2) in the first, middle, and last 100 cows in the milking order were calculated alongside the whole herd prevalence of lameness. The effect of farm and visit number on the proportion of cows with a lameness score ≥ 2 identified by each sampling group was analysed using a repeated measures generalised estimating equation with a logistic regression setting with the proportion of all cows in the group identified as lame (classified on a binary scale as score ≥ 2) used as the outcome variable, and farm, visit number and sampled group as the predictor variables. The model included all main effects (farm, visit number, and sampled group) and interactions between visit number and group, and farm and

Table 1. Descriptive details of six farms used in a study assessing sampling strategies for the estimation of herd level lameness prevalence on New Zealand pastoral dairy farms.

Farm	Herds	Predominant breeds	Milking parlour type	Milking frequency	Number of cows scored[a]
1	Single	Holstein/cross breed	Herringbone	Once daily	268–290
2	Single	Holstein/cross breed	Herringbone	Once daily	207–263
3	Two	Jersey/cross breed	Rotary	Twice daily	261–325
4	Single	Holstein/cross breed	Rotary	Twice daily	398–421
5	Two	Holstein/cross breed	Rotary	Twice daily	460–510
6	Single	Holstein/cross breed	Herringbone	Once daily	210–220

[a]Range in number of cows locomotion-scored per visit; all farms visited six times except Farm 5, which was only visited five times, one visit excluded for Farm 6 as fewer than 200 cows scored.

group. For this analysis, all visits (n = 3) where no lame cows had been identified were excluded.

To determine the level of agreement between the overall lameness prevalence as determined by whole herd locomotion scoring (%) and lameness prevalence identified through locomotion scoring of the first, middle and last 100 cows in the milking order, three limits-of-agreement plots were created. This analysis used data from all visits where > 200 cows were scored. For each of the three groups, the mean of the sample and the herd prevalence ("mean") and their difference ("difference") were calculated for each farm/visit combination and a mean/difference plot created. A generalised linear mixed model, with difference as the outcome, mean as the fixed predictor variable, and farm as a random effect was then used to determine if there was a significant association between mean and difference (i.e. proportionate bias) for that sampled group (i.e. first, middle or last 100 in the milking order).

Then for all farm/visit combinations, sampling distributions were generated for each of the sample groups using bootstrapping (resampling with replacement), so that key parameters of that distribution could be calculated without assumptions as to the true distribution (e.g. that it is normal). Thus, for example, for $Farm_i$, $Visit_i$, Sample group first 100 cows, 100 cows were randomly selected with replacement from the first 100 cows in the milking order at $Visit_i$ on $Farm_i$. The cumulative count of selected cows whose LS had been ≥ 2 was used as the prevalence for this bootstrapped sample. A total of 1,000 bootstrap samples were simulated for each farm/visit (n = 28), and sample group (n = 3). The difference between the bootstrap estimates and the actual prevalence (from whole herd locomotion scoring) was then calculated by subtraction for each milking order group, and farm/visit. The 2.5% and 97.5% percentiles of these differences (n = 28 for each milking order group) were then calculated and plotted against the mean on the mean/difference plot. Two generalised linear mixed models, one with the 2.5% percentile and one with the 97.5% percentile as the outcome, mean as the fixed predictor variable, and farm as a random effect were then used to determine the limits of agreement (equivalent to the lines of best fit for the 2.5% and 97.5% percentiles) across the range of lameness prevalences found in this study. All residuals were checked to assess normality using histograms and the Kolmogorov–Smirnov test. Acceptable agreement was defined as the limits of agreement being < 5% over the range of herd lameness prevalences seen in this study (i.e. the analysis predicts that, for the particular milking order group, 95% of prevalence estimates will be within 5% of the value determined using whole herd scoring).

All statistical analyses were undertaken using SPPS v. 29 (IBM, Armonk, NY, USA) except for the bootstrapping, which used R (R Foundation for Statistical Computing, Vienna, Austria).

Results

Farm details and number of cows locomotion-scored are summarised in Table 1. Across the 29 visits, fewer than 200 cows were scored on only one visit (Farm 6, Visit 5); data are thus excluded from that visit, leaving scoring data from 28 visits.

Effect of farm and visit on the proportion of lame cows identified by each milking order group

Across the 28 visits, 8,872 lameness scores were recorded, of which 263 (3.0%) were recorded as lame (lameness score ≥ 2). Of the 28 farm visits, herd lameness prevalence was > 5 % (the threshold for acceptable welfare set by Sapkota *et al.* (2022)) on four occasions (see Figure 1). Of those four occasions, all but one were observed on the first visit to a farm (i.e. in mid-October to mid-November; see Figure 1).

These 263 lame cow observations across all visits came from cows within and without the sample groups. Of the 263 lame cow observations, 107 (40.7%) were made in the last 100 cows in the milking order, with 68 (25.9%) and 38 (14.4%) in the middle 100 and first 100, respectively. All interactions included in the model were statistically significant (p ≤ 0.009), i.e. the proportion of lame cows identified within a milking order group was affected by farm, and by visit. This is illustrated for the interaction between group and farm (Figure 2). Although,

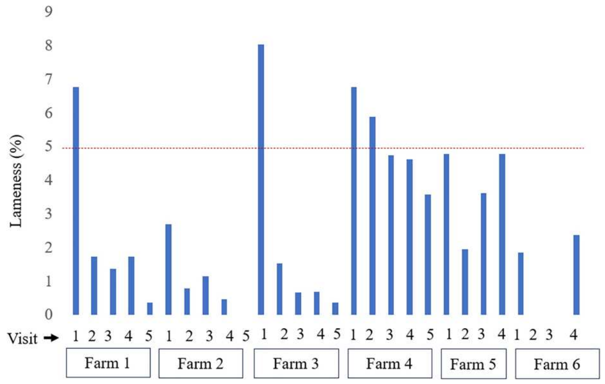

Figure 1. Proportion of cows from six North Island, New Zealand dairy farms, undergoing approximately 6-weekly whole herd locomotion score visits from mid-October to mid-May, using the DairyNZ 0–3 lameness scoring system, that were scored as lame (score ≥ 2) by farm and visit. There was no fifth visit on Farm 5, and at Visit 5 on Farm 6 < 200 cows were scored, so data from these visits are excluded. The horizontal dotted line represents the acceptable welfare threshold for lameness (%) on New Zealand dairy farms (Sapkota *et al.* 2022). On three occasions when no lame cows were observed, there is no vertical bar for this time point.

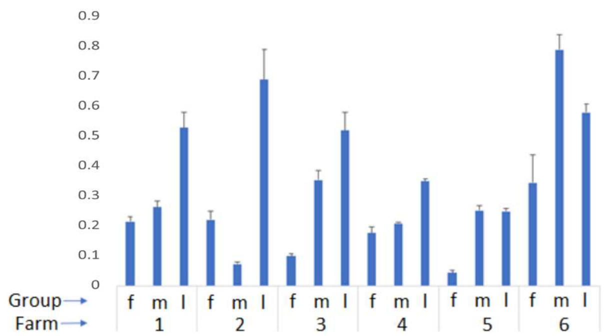

Figure 2. Estimated marginal means (95% CI) for the proportion of total lame cows observed in the first 100 (f), middle 100 (m), and last 100 (l) cows in the milking order for individual farms from a generalised estimating equation with farm, visit number, sampled group and interactions between visit number and group and between farm and group, as predictor variables. Data are from cows on six North Island, New Zealand dairy farms, undergoing approximately 6-weekly whole herd locomotion score visits from mid-October to mid-May, using the DairyNZ 0–3 lameness scoring system, that were scored as lame (score ≥ 2), by farm and visit.

overall, observing the last 100 cows in the milking order identified the highest proportion of lame cows (estimated marginal mean from the model (EMM) 0.5; 95% CI = 0.48–0.52), observing the middle 100 identified the next highest proportion (EMM 0.26; 95% CI = 0.25–0.28), and observing the first 100, the lowest (0.14; 95% CI = 0.12–0.17), this pattern was only seen on 3/6 farms (see Figure 2). For example, a lower proportion of lame cows was observed in the last 100 (compared to the middle 100) on Farm 6 (EMM 0.79; 95% CI = 0.74–0.84 *vs.* 0.58; 95% CI = 0.56–0.61, respectively), while on Farm 2 lame cows were more likely to be observed in the first 100 than the middle 100 (EMM 0.22; 95% CI = 0.20–0.25 *vs.* 0.07; 95% CI = 0.07–0.08).

Limits of agreement

Figure 3 shows the limits of agreement plots for each of the three milking order sample groups. For all of the analyses, residuals were normally distributed. For the first 100 and middle 100, once farm was accounted for, there was no fixed or proportional bias (p = 0.17 and 0.4, respectively) (see Figure 3(a, b)), i.e. across the range of means seen in this study, the expected mean difference between sample prevalence and herd prevalence is zero. In contrast, for the last 100, there was a strong positive association between the mean and the difference, with the line of best fit being: difference = 0.38 (95% CI = 0.31–0.46)*mean. Thus, if scoring the whole herd suggests that lameness prevalence is 2%, our data suggest that, on average, the prevalence of lame cows in the last 100 will be 2.9%,[1] and if whole herd scoring suggests that prevalence is 8% the prevalence of lame cows in the last 100 will be 11.8%.

For all three milking order groups, the limits of agreement widened as lameness prevalence increased (see Figure 3 and Table 2). For both first and last 100, our models predict that more than 5% of herds will have a difference in lameness prevalence of > 5% between the milking order sample and whole-herd scoring when true herd prevalence is ≥ 8%. This is still the case when the lower limits of the CI of our slope (see Table 2) are used rather than the model estimate (e.g. −0.87 for the 2.5% percentile for the first 100 rather than −0.99, and 1.1 rather than 1.26 for the 97.5% percentile for the last 100). In contrast, although our model prediction suggests that for the middle 100 more than 5% of herds will have a difference in lameness prevalence of > 5% between the milking order sample and whole herd scoring when true herd prevalence is ≥ 8%, the lower (absolute numeric value) for the confidence intervals for the 2.5% and 97.5% percentile are consistent with 95% of herds having a difference in lameness prevalence of ≤ 4% between the milking order sample and whole-herd scoring when true herd prevalence is ≤ 8% (the narrowest predictive interval is when the difference is ±0.5*mean).

[1]For the last 100 cow group, this is solved using simultaneous equations: if herd prevalence (p) = 2, mean of sample p (S) and herd p = (S + 2)/2 and difference between sample p and herd p = S − 2. But difference also = 0.38*mean = 0.38*(S+2)/2 = (0.76 + 0.38S)/2. So, S − 2 = (0.76 + 0.38S)/2. This is equivalent to 2 S − 0.38S = 4.76 and S = 2.9.

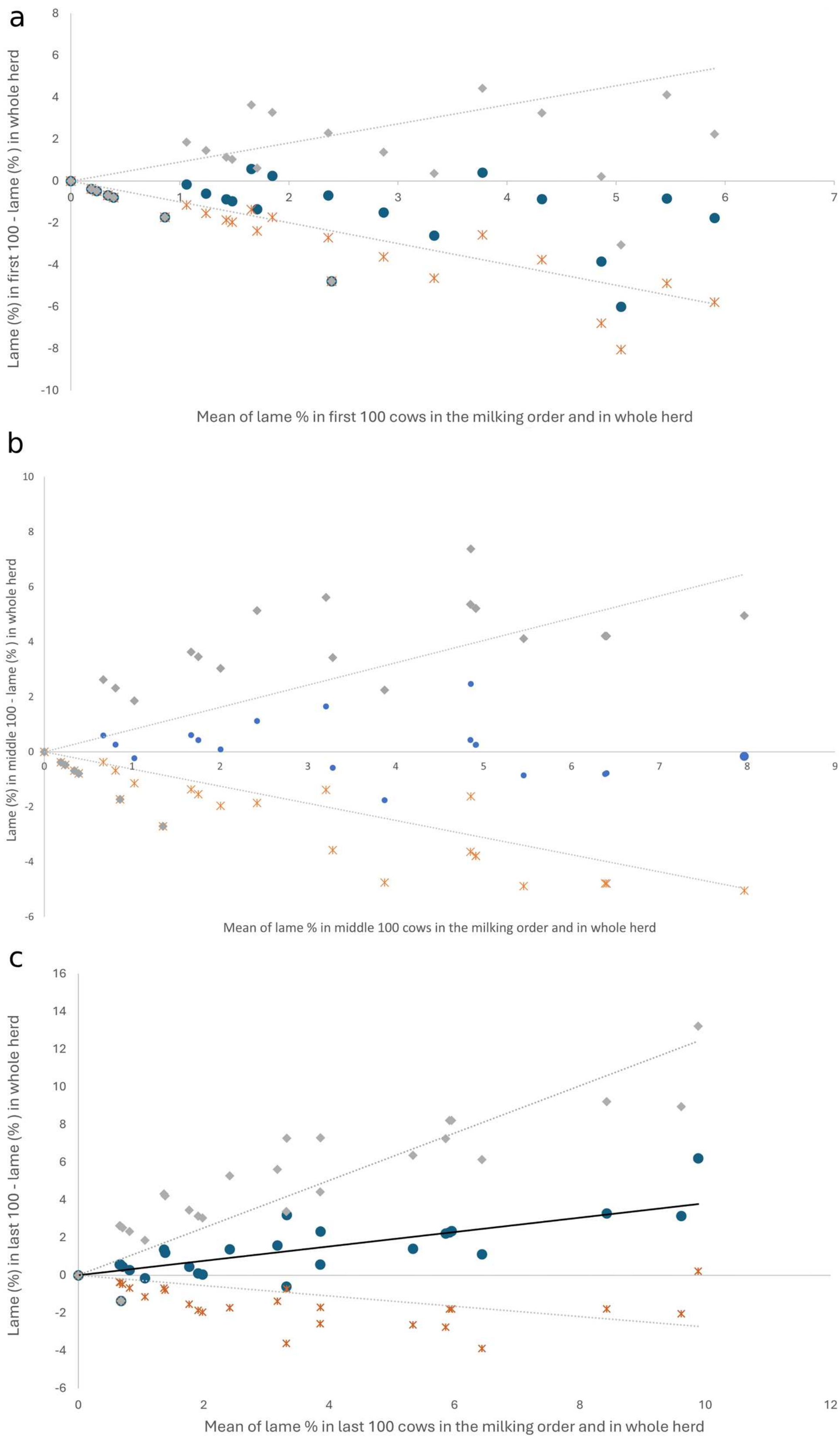

Figure 3. Limit-of-agreement plots comparing the difference between whole herd lameness prevalence and prevalence estimated from sampling the (a) first, (b) middle, and (c), last 100 cows in the milking order against the mean of sample and whole herd prevalence. Data are from six North Island New Zealand dairy farms undergoing approximately 6-weekly whole herd locomotion score visits from mid-October to mid-May, using the DairyNZ 0–3 lameness scoring system. Circles (●): mean difference from 1,000 bootstrapped samples; diamonds (◆): 97.5% percentile of the bootstrapped distribution; asterisks (✱): 2.5% percentile of the bootstrapped distribution. The solid line in part (c) is the line of best fit for mean *vs.* difference (for parts (a) and (b) there is no association between mean and difference ($p \geq 0.17$) and mean difference is 0, so the x-axis is effectively "the line of best fit"). The dotted lines are lines of best fit for the association between mean and the 2.5% and 97.5% percentiles from the generalised linear mixed model and thus mark the 95% limits-of-agreement, i.e. the range over which the model predicts that 95% of differences between the herd and sample prevalence will lie.

Table 2. Slope coefficients (95% CI) for estimating the line of best fit from a generalised mixed model fitted to 28 bootstrapped distributions of the 2.5% percentiles and the 97.5% percentiles of the difference in lameness prevalence between whole herd sampling and sampling the first, middle or last 100 cows for cows from 28 scoring occasions for six New Zealand pastoral dairy farms. Model-estimated prediction intervals for lameness prevalence in each of the 100-cow samples from three milking order groups are also shown at two levels of whole herd lameness prevalence (HP).

	Slope coefficient to estimate line of best fit (95% CI)[a]		Approximate 95% prediction interval for group lameness prevalence[b]	
Milking order group	2.5% percentile	97.5% percentile	2% HP	8% HP
First 100	−0.99 (−1.12 to −0.87)	0.91 (0.63–1.19)	0.02–3.8	0.08–15.3
Middle 100	−0.62 (−0.75 to −0.5)	0.81 (0.5–1.12)	0.8–3.6	3.0–14.5
Last 100	−0.27 (−0.4 to −0.15)	1.26 (1.1–1.41)	1.3–5.1	5.3–20.4

[a]Slope of the line of best fit of the 2.5% and the 97.5% percentiles of the difference in lameness prevalence between whole herd sampling and sampling the first, middle or last 100 cows for cows against the mean of the sample and whole herd prevalence i.e. expected value of 2.5% percentile = slope * mean of lame (%) in sample and whole herd lame (%).

[b]When prevalence from whole herd locomotion scoring (HP) is 2% or 8%, the model predicts that 95% of results from the milking order sample will be between these values.

Calculation of approximate 95% prediction interval: For the first and middle 100, there is no association between mean and difference (see Figure 3), so when HP is 8%, on average the prevalence of lameness in the sample will also be 8%. So, the approximate 95% prediction interval of prevalence is calculated by (1) multiplying the slope for the 2.5% line and the 97.5% line by HP and (2) adding those differences to HP. – For example, for the middle 100 and 8% HP, this gives −4.96 (from: 8*−0.62) and 6.48 (from: 8*0.81), giving to 1 d.p. 3.0 (from: 8 − 4.96) and 14.5 (from: 8 + 6.48). For the last 100 cows, as there is an association between mean and difference, the expected mean prevalence must be calculated for each HP which is then used to calculate the expected differences. (See footnote to Results section)

Interpretation of approximate 95% prediction interval: Agreement is defined as "acceptable" if the model predicts that 95% of milking order sample results will be within 5% of the HP (i.e. the 95% prediction interval does not extend more than 5% above or below HP). For HP 2%, the 95% prediction interval for the middle 100 is 1.24% below to 1.6% above HP, i.e. acceptable agreement. However, for HP 8%, the 95% prediction interval for the middle 100 is 5.0% below to 6.5% above HP, i.e. agreement is not acceptable.

Additionally, the slope is an estimate from the sample data for which CI have been calculated. The lowest absolute value of each CI can be used to estimate the narrowest prediction interval which is compatible with the data. For the middle 100 group at HP 8%, the prediction interval is 4–12%, so the data are compatible with a difference of ± 4%.

Discussion

The aim of this study was to assess whether a sampling strategy (based on scoring a specific group of 100 cows within the milking order) could reliably predict herd level prevalence of lameness on New Zealand dairy farms, with the aim of eliminating the need for an additional assessor to lameness score during a time-limited welfare assessment.

The study sampled three groups of cows: the first 100, the middle 100 and last 100 in the milking order, building on the study by Main *et al.* (2010). Our initial analysis identified that, although overall the proportion of lame cows identified in each group was significantly different ($p < 0.001$ for all comparisons), the effect of group varied markedly by farm and by visit. Consequently, even within a farm, there was no consistent proportion of lame cows that were found in each sample group. Thus, although our results are consistent with those of Sauter-Louis *et al.* (2004), who reported that cows that became lame were more likely to be observed in the last quarter of the walking order after milking (milking order), they also show that farm level and seasonal factors seem to affect the position of lame cows within the milking order. Both Sauter-Louis *et al.* (2004) and the current study are small studies (10 and 6 farms, respectively), so our results highlight the need for larger studies of lameness and milking order. In particular, we need to study the impact, in cows permanently kept at pasture, of lameness on milking order, as we believe that it is likely that farms in which lameness has less impact on milking order are managing their cows better between pasture and milking parlour than farms where lameness has more impact. This is because cows have a relatively narrow range of preferred positions within the milking order (Sauter-Louis *et al.* 2004), and therefore if lameness is having less impact on milking order, this means that cows are able to maintain their preferred position within the herd.

Our limits-of-agreement analysis suggested that the difference in all sample groups was within the target set by Main *et al.* (2010) of estimating prevalence (i.e. ± 5%) only when mean lameness prevalence was low (< 5%) but not if mean prevalence was > 5%. This is despite the exclusion of visit number from our model for calculating the limits of agreement. That is, we assumed that the results for Farm 1 at Visit 1 were independent of the results for the other five farms at Visit 1, and that this independence continued for the remaining visits. If this assumption is incorrect then our limits of agreement may be too narrow (i.e. there is even more variation in the agreement between group- and herd-level prevalence than we have predicted), as we would have overestimated the information in the data. Our data are thus in contrast to the conclusion by Main *et al.* (2010) that sampling the middle 100 was sufficiently accurate to determine herd lameness prevalence to within 5%. However, there are two caveats to this conclusion. Firstly, our data are compatible with the lameness percentage in the middle 100 estimating actual prevalence to within 5% on 95% of farms (see lower absolute values for the CI in Table 2) at least up to 10% true prevalence; and secondly our criteria (within 5% on 95% of farms) are stricter than Main *et al.* (2010), who reported that actual prevalence was within 5% for 30/36 herds. This is 83% rather than 95%, and may, actually, be even lower (95% CI = 68–92%).

However, we question whether this 5% target is suitable for estimating lameness prevalence on dairy farms in New Zealand. In a recent study, Mason *et al.* (2023) conducted lameness assessments on 120 farms during two distinct periods, once in the spring and once in the summer. Using this data, out of the 240 total assessment occasions, Laven *et al.* (2025) identified instances where more than 10% of cows were lame on only 14 occasions. Thus, if across all of those scoring occasions we had simply assumed a constant 5% lameness prevalence on every farm, we would have been within 5% of the whole herd prevalence from locomotion scoring on 226/240 occasions (94 (95% CI = 90−96) %) without scoring any cattle (as setting a constant 5% prevalence means that any farm with a herd prevalence ≤ 10% is within 5% of our estimate). Similarly, if we had simply recorded that there were 0% lame cows on all farms, we would have been within 5% of whole herd prevalence on 188/240 occasions (78 (95% CI = 73-83) %) (Laven *et al.* 2025), as only 52 scoring occasions identified a herd prevalence > 5%.

Secondly, our recommended acceptable welfare threshold for lameness of 5% (Sapkota *et al.* 2022), is much lower than that used in other countries with well-developed welfare assessment schemes (e.g. UK: 15% annual average prevalence (RSPCA 2018); Netherlands: < 33% moderately lame cows (de Vries 2013); Canada: ≤ 10% (NFACC 2023)). Thus, setting a target accuracy of within ± 5% of the true value will not allow us to distinguish between farms with very low levels of lameness and ones with lameness prevalence equivalent to the unacceptable threshold of 10% set by Sapkota *et al.* (2022).

We thus need a more accurate measurement of lameness than suggested by either Main *et al.* (2010) or Beggs *et al.* (2019). Our data suggest that measuring the middle 100 is probably the best approach, and may be reasonably accurate if whole herd prevalence based on locomotion scoring is < 8%. However, as our data are from only six farms, we need far more data to confirm this suggestion. Secondly, if we are going to use sampling to determine lameness prevalence, even being within 2% may not be sufficient to accurately categorise farms as having adequate welfare. Approximately 33% of the scoring occasions reported by Mason *et al.* (2023) reported a herd prevalence of between 3 and 7% (i.e. within 2% of our adequate threshold of 5%). If we had a sampling strategy that estimated herd prevalence to ± 2%, all of those herds could potentially be misclassified as marginal when they were adequate, or vice versa. Thus, if a welfare assessment on a New Zealand dairy farm is to use a sampling strategy to estimate herd level prevalence, it would be better to use it as a screening test rather than an absolute measure of lameness prevalence. This determination is similar to the conclusion from our recent analysis of the repeatability of whole locomotion scoring (Laven *et al.* 2025) based on the data from Mason *et al.* (2023). We identified that a single, one-off, whole-herd locomotion score did not accurately predict future lameness prevalence, and was not a good predictor of lameness-associated welfare status, as not only was the proportion of poor welfare herds identified too low, but the one-off assessment did not distinguish herds where lameness was a temporary problem from those with ongoing lameness issues. We concluded that, for welfare assessment purposes, the standard one-off independent assessment of lameness status should be combined with regular locomotion scoring by trained farmers or external scorers. If this is done, then that one-off independent assessment is no longer being used as *the* determinant of lameness-related welfare status but as one (important but not definitive) piece of evidence in determining such status. If this is the case, then we can afford a less accurate sampling strategy. We thus need further testing of more herds to identify and test potential sampling strategies. Once the best strategy is identified, we then need to determine how best to integrate that strategy into a time-limited welfare assessment protocol.

Conclusion

This study showed that although, as expected, there was a significant association between milking order and risk of lameness, the proportion of lame cows identified within a particular group within that order (i.e. first, middle or last 100 cows) was not consistent and was significantly affected by farm and by timing of the visit. This suggests that there will be variations across farms and seasons in the proportion of lame cows identified by sampling and thus variations in the accuracy of predicting overall lameness from such a sample.

This was confirmed using the limits-of-agreement analysis which showed that none of the three milking order groups could predict whole herd prevalence within a 5% margin. We need more research on a larger number of farms to better characterise the accuracy of scoring the middle 100 cows, and to develop and test strategies that use scoring the middle 100 cows as a screening test alongside additional herd lameness scoring by trained farmers or external scorers.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

RA Laven http://orcid.org/0000-0002-8938-8595
KR Müller http://orcid.org/0000-0001-7023-5356
DA Yang http://orcid.org/0000-0003-3893-9403

References

Alawneh J, Laven R, Stevenson M. The effect of lameness on the fertility of dairy cattle in a seasonally breeding pasture-based system. *Journal of Dairy Science* 94, 5487–93, 2011. https://doi.org/10.3168/jds.2011-4395

Beggs D, Jongman E, Hemsworth P, Fisher A. Lame cows on Australian dairy farms: a comparison of farmer-identified lameness and formal lameness scoring, and the position of lame cows within the milking order. *Journal of Dairy Science* 102, 1522–9, 2019. https://doi.org/10.3168/jds.2018-14847

***DairyNZ**. *Lameness Field Guide*. https://www.dairynz.co.nz/media/5787958/lameness_field_guide_july_2017_v2_web.pdf (accessed 29 August 2024). DairyNZ, Wellington, NZ, 2017

***DairyNZ**. *Latest Dairybase Benchmarks*. https://www.dairynz.co.nz/business/dairybase/benchmarking/latest-dairybase-benchmarks/ (accessed 29 August 2024). DairyNZ, Wellington, NZ, 2022

***de Vries M.** Assuring dairy cattle welfare. *PhD thesis*, Wageningen University, Wageningen, Netherlands, 2013

Fabian J, Laven R, Whay H. The prevalence of lameness on New Zealand dairy farms: a comparison of farmer estimate and locomotion scoring. *The Veterinary Journal* 201, 31–8, 2014. https://doi.org/10.1016/j.tvjl.2014.05.011

Green L, Hedges V, Schukken Y, Blowey R, Packington A. The impact of clinical lameness on the milk yield of dairy cows. *Journal of Dairy Science* 85, 2250–6, 2002. https://doi.org/10.3168/jds.S0022-0302(02)74304-X

Laven R, Fabian J. Applying animal-based welfare assessments on New Zealand dairy farms: feasibility and a comparison with United Kingdom data. *New Zealand Veterinary Journal* 64, 212–7, 2016. https://doi.org/10.1080/00480169.2016.1149523

Laven RA, Mason W, Laven LJ, Müller KR. The repeatability of whole herd lameness scoring: an analysis of a New Zealand dataset. *New Zealand Veterinary Journal* 73, 46–52, 2025. https://doi.org/10.1080/00480169.2024.2394554

Main D, Barker Z, Leach K, Bell N, Whay H, Browne W. Sampling strategies for monitoring lameness in dairy cattle. *Journal of Dairy Science* 93, 1970–8, 2010. https://doi.org/10.3168/jds.2009-2500

Mason W, Müller K, Huxley J, Laven R. Prevalence of lameness on pasture-based New Zealand dairy farms: an observational study. *Preventive Veterinary Medicine* 220, 106047, 2023. https://doi.org/10.1016/j.prevetmed.2023.106047

***NFACC**. *Code of Practice for the Care and Handling of Dairy Cattle.* https://www.nfacc.ca/codes-of-practice/dairy-cattle/code2023#section5 (accessed 29 August 2024). National Farm Animal Care Council, Lacombe, Canada, 2023

O'Leary N, Byrne D, Connor O, Shalloo AL. Cattle lameness detection with accelerometers. *Journal of Dairy Science* 103, 3895–911, 2020. https://doi.org/10.3168/jds.2019-17123

Ózsvári L. Economic cost of lameness in dairy cattle herds. *Journal of Dairy, Veterinary and Animal Research* 6, 00176, 2017. https://doi.org/10.15406/jdvar.2017.06.00176

***RSPCA**. *RSPCA Welfare Standards for Dairy Cattle*. https://science.rspca.org.uk/documents/1494935/9042554/RSPCA+welfare+standards+for+dairy+cattle+%28PDF+7.76MB%29.pdf/41638530-20de-c6cc-5e9c-7b73f9c8f4b7?t=1557731468543 (accessed 29 August 2024). Royal Society for the Prevention of Cruelty to Animals, Horsham, UK, 2018

Sapkota S, Laven R, Müller KR, Kells N. Practicability of a time-limited welfare assessment protocol for pasture-based dairy farms, and a preliminary assessment of welfare outcome thresholds. *Animals* 12, 2481, 2022. https://doi.org/10.3390/ani12182481

Sauter-Louis C, Chesterton R, Pfeiffer D. Behavioural characteristics of dairy cows with lameness in Taranaki, New Zealand. *New Zealand Veterinary Journal* 52, 103–8, 2004. https://doi.org/10.1080/00480169.2004.36414

Sørensen JT, Rousing T, Møller SH, Bonde M, Hegelund L. On-farm welfare assessment systems: what are the recording costs? *Animal Welfare* 16, 237–9, 2007. https://doi.org/10.1017/S0962728600031420

Van Os JMC, Weary DM, Costa JHC, Hötzel MJ, von Keyserlingk MAG. Sampling strategies for assessing lameness, injuries, and body condition score on dairy farms. *Journal of Dairy Science* 102, 8290–304, 2019. https://doi.org/10.3168/jds.2018-15134

Werema CW, Yang DA, Laven LJ, Mueller KR, Laven RA. Evaluating alternatives to locomotion scoring for detecting lameness in pasture-based dairy cattle in New Zealand: in-parlour scoring. *Animals* 12, 703, 2022. https://doi.org/10.3390/ani12060703

Whay HR, Shearer JK. The impact of lameness on welfare of the dairy cow. *Veterinary Clinics of North America: Food Animal Practice* 33, 153–64, 2017. https://doi.org/10.1016/j.cvfa.2017.02.008

*Non-peer-reviewed