Longitudinal study of herd udder hygiene and its association with clinical mastitis in pasture-based dairy cows

Sam Rowe,1* William Tranter,2,3 and Richard Laven4

1Faculty of Science, Sydney School of Veterinary Science, The University of Sydney, Camden, New South Wales 2570, Australia
2Tableland Veterinary Service, Malanda, Queensland 4885, Australia
3College of Public Health, Medical and Veterinary Sciences, James Cook University, Townsville, Queensland 4811, Australia
4Institute of Veterinary, Animal and Biomedical Sciences, Massey University, Palmerston North 4474, New Zealand

ABSTRACT

The objectives of this exploratory study were to (1) describe the association between herd-level udder hygiene and clinical mastitis and (2) investigate how sample size and milking stage affect the accuracy and precision of herd udder hygiene assessments made at milking time. A prospective longitudinal study was conducted in a dairy herd in Northern Australia as part of a previously published clinical trial of premilking teat disinfection. Video footage from 35 afternoon milkings was used to conduct 12,544 udder hygiene scores from 504 cows during an 89-d period and measure udder hygiene of the herd (proportion of cows with udder hygiene ≥3 out of 4). Linear interpolation was used to estimate herd udder hygiene on the days that were not scored, such that a herd-level udder hygiene measure was available for all cow-days in the study. Clinical mastitis events occurring during the study period were detected and recorded by farm staff according to a standardized definition. The relationship between herd udder hygiene on each of 1, 2, and 3 d before each study day (d −1, −2, and −3, respectively) and clinical mastitis at the cow level on each study day (each in turn being set as d 0) was determined using multivariable generalized estimating equations (family = Poisson, link = log), with the unit of analysis being the cow-day, adjusting for potential confounders and the clustering within the data. In addition, sampling strategies were evaluated by simulating herd udder hygiene assessments using a subset of cows in the herd. Herd udder hygiene from d −1, −2, and −3 was positively associated with clinical mastitis on d 0 (incidence rate ratio = 1.4 per 10-point increase in the percentage of cows with poor udder hygiene). Sampling strategy simulation found that at least 80 cows needed to be scored to achieve sufficiently precise estimations of herd udder hygiene. Furthermore, cows scored later during milking were slightly more likely to have poor udder hygiene than those scored earlier (risk ratio = 1.02 for cows that were 10% later in the milking order). More research is needed to evaluate risk factors for poor udder hygiene and potential interventions in pasture-based dairy cows.

Key words: udder hygiene, mastitis, pasture-based, environment

INTRODUCTION

Mastitis is a significant source of impaired health and welfare of dairy cows and reduced economic returns for dairy producers. On many farms, the risk of cow-to-cow transfer of mastitis pathogens is low due to the adoption of mastitis control practices such as postmilking teat disinfection, regular milking machine maintenance, lactating and dry cow antibiotic therapy, and strategic culling (Ruegg, 2017). Consequently, the environment is often the predominant cause of clinical mastitis in confinement (Breen et al., 2009; Lago et al., 2011; Levison et al., 2016) and pasture-based dairy systems (McDougall et al., 2007; Petrovski et al., 2009; Shum et al., 2009; Charman et al., 2012). There is currently great interest in describing the epidemiology of environmental mastitis (i.e., mastitis associated with IMI of environmental origin) and in the development of diagnostic tools for producers and advisors to better investigate environmental mastitis problems. For example, recent research has demonstrated that high concentrations of bacteria in bedding may increase risk of clinical mastitis (Rowbotham and Ruegg, 2016; Patel et al., 2019) and IMI (Rowe et al., 2019). No equivalent method has been validated for estimating infection pressure from environmental reservoirs in pasture-based dairy systems.

One potential method of measuring environmental contamination that is practical because it can be done
during milking is udder hygiene scoring. In confined dairy herds, observational studies have shown a relationship between udder hygiene and mastitis. For example, Breen et al. (2009) reported that cows with very dirty udders had 1.5 times higher odds of developing clinical mastitis during the next month than cows with clean udders, and cow-level associations between udder hygiene and SCC have also been observed (Schreiner and Ruegg, 2003; Reneau et al., 2005; Dohmen et al., 2010). At the herd level, Nyman et al. (2007) found that herds with high incidences of clinical mastitis had 5.4 times higher odds of poor leg hygiene than herds with low incidences of clinical mastitis, and Dohmen et al. (2010) reported that herd median cleanliness score was positively associated with bulk milk SCC in organic, but not conventional, dairy herds.

Substantially less research has been conducted to describe the relationship between udder hygiene and mastitis in pasture-based dairy cows. In an observational study of 708 primiparous cows from 30 herds in New Zealand, Compton et al. (2007) showed that IMI at 0 to 5 DIM was 1.3 times more likely in cows with mild to severely dirty udders at the time of milk sampling for bacteriology. However, udder hygiene after calving was not associated with postcalving clinical mastitis and precalving udder hygiene was not associated with any measures of postcalving udder health. Given that bacterial profiles in udder contaminants in pasture-based cows are likely to differ from those managed in confinement, more research is needed to investigate relationships between udder hygiene and clinical mastitis in pasture-based dairy cows.

Another area requiring further investigation is herd-level udder hygiene assessment strategies in pasture-based dairy herds. Given that cow-side udder hygiene scoring during milking is time consuming and occurs at a time when staff resources are already stretched, producers and milk quality advisors are unlikely to score all cows in the herd. Therefore, research is needed to describe how sample size (i.e., number of cows from the herd scored) affects accuracy and precision of herd udder hygiene assessments. Furthermore, investigations are needed to determine whether herd udder hygiene measurements can be biased if assessments are restricted to cows during a specific part of the milking order (i.e., early vs. late cows), as this has been demonstrated to occur for lameness scoring (Sauter-Louis et al., 2004). The objectives of this study were to (1) describe the association between herd-level udder hygiene and cow-level clinical mastitis and (2) investigate how sample size and milking stage affect the accuracy and precision of herd udder hygiene assessments made at milking time.

The Strengthening the Reporting of Observational Studies in Epidemiology–Veterinary Extension (STROBE-Vet) statement guidelines were followed in the reporting of this study (Sargeant et al., 2016). The prospective longitudinal study was conducted over a 3-mo period in summer and early autumn (January 14 to April 14) in 2016 in a pasture-based, year-round calving dairy herd in Malanda, Australia. Total rainfall during the study period was 360 mm and average daily peak temperature was 30.5°C. Rainfall is usually highest during this time in northern regions of Australia. Ethics approval was granted by the James Cook University Animal Ethics Committee (#A2249).

Study Herd

The predominant breed in the study herd was Holstein-Friesian. During the study, average daily milk production ranged from 23 to 26 L/cow and the milking cow diet consisted of tropical pastures, grain-based concentrate fed in the milking parlor, and a mixture of maize silage, protein meal, and minerals fed twice daily on a gravel feed pad. Cows had access to the feed pad between the morning and afternoon milkings and for up to 2 h after the afternoon milking. The herd was selected because it had experienced increases in clinical mastitis (11.6–24.3 cases/100 cows per month) in previous summers despite the adoption of many well-accepted preventative strategies, such as regular milking machine maintenance, postmilking teat disinfection, and whole-herd treatment with dry cow antibiotics and internal teat sealants at drying off. In addition, it was expected that udder hygiene during summer would fluctuate from day to day, with the major source of udder contamination being a shaded area adjacent to the feed pad where cows would congregate between the morning and afternoon milkings. The herd was concurrently being used for a clinical trial evaluating the effect of a premilking teat disinfection protocol on clinical mastitis incidence, with approximately half of the cows receiving the premilking teat disinfection protocol (Rowe et al., 2018).

Milking Management and Data Collection

Clinical mastitis detected by farm workers was the principal outcome event in this study and was detected by visual examination of foremilk of cows as part of the premilking routine. Cases were defined as the presence of abnormal milk (watery or clotted) after at least 3 strips of foremilk. Other signs such as increased milk
conductivity and swelling, pain, and heat in the udder were considered to be suggestive but not indicative of clinical mastitis. The herd manager recorded all cases of clinical mastitis on the day of detection, including date, cow ID, quarter affected, and treatment given. These records, along with calving, dry-off, and culling dates were used to determine whether cows experienced a case of clinical mastitis during the study.

Udder hygiene, which was assessed every 2 to 3 d over the 89-d study period, was the primary exposure of interest in this study; 35 afternoon milkings were scored. Udder hygiene scoring involved the placement of a smartphone (iPhone 4, Apple, Cupertino, CA) in a well-lit area of the shed to collect video footage of cows standing on the rotary parlor (Figure 1A). A specialized photography application [Hex (cam), Marsh Studio, Yokohama, Japan] was used to record video at a low frame rate (12 frames/s) to conserve the battery and memory of the device. Video records were downloaded every 2 wk onto a computer and stored on 2 backup drives. All recordings were viewed after the study period (July 2016) at 12× speed, allowing a 2-h milking to be scored in 10 min (Figure 1B). An example of the footage can be viewed on the analysis log webpage at https://rpubs.com/samrowe/UHS_CM-2020. Udder contamination was scored using a 4-point system developed by Schreiner and Ruegg (2002): 1 = no dirt or manure; 2 = 2 to 10% of surface area of the udder covered with dirt or manure; 3 = 10 to 30% of surface area of the udder covered with dirt or manure; 4 = >30% of surface area of the udder covered with dirt or manure. Cows with dark-colored udders (<10% of the herd) were excluded from scoring due to difficulties in visualizing udder contamination in the footage. An udder hygiene score (UHS) of 3 or 4 was considered indicative of poor hygiene. Cow ID was not recorded during scoring. Daily herd udder hygiene was calculated by dividing the number of cows with poor hygiene by the total number of cows scored for that milking.

As part of the concurrent clinical trial, cows were randomized to receive a premilking teat disinfection routine that consisted of prestripping, washing teats with low-pressure water, dipping with a registered premilking teat disinfectant, and drying with a paper towel. Control cows received the regular premilking routine used on that farm, which included prestripping, and teat washing with high-pressure water only when teats and udders were heavily contaminated with wet manure, or mud, or both. If the teats and udder were contaminated with dry material, milking staff removed contamination with a gloved hand. Farm workers who monitored cows for clinical mastitis were not blinded to treatment status of cows.

**Coding of Exposure and Outcome Variables**

The data set was formatted to accommodate for a time-varying predictor (herd udder hygiene) by creating a separate row for each cow-day at risk (i.e., the unit of analysis was the cow-day). Cows were eligible for inclusion into the study if they were present in the lactating herd for at least 1 d during the study period. This was determined using individual health records, which were downloaded from herd software (Dairy-WIN 2004 v. 99.91.148, Massey University, Palmerston North, New Zealand). The first day of the study (January 16, 2016) or, for cows not in the lactating herd on that date, the day of their next calving after this date, was the first day at risk for each cow. Cows that were dried off during the study were re-enrolled after the subsequent calving. A herd inventory was conducted before the study to confirm that records matched the

![Figure 1.](https://rpubs.com/samrowe/UHS_CM-2020)
cows in the lactating herd. Cow-days were excluded from analyses after cows left the lactating herd (dry-off, culling) or experienced a case of clinical mastitis.

A herd udder hygiene value (proportion of the herd with scores ≥3 out of 4) was assigned for each day of the study. This was determined using udder hygiene scores conducted on 35 of the 89 d of the study period, as indicated by the blue data points in Figure 2. The herd udder hygiene values during the remaining 54 d were estimated by interpolating between measurements using a linear method, as indicated by the black line in Figure 2. These data were used to assign a herd udder hygiene value for each cow-day at risk in the final data set. Given that experimental challenge studies have found incubation periods of >12 h for clinical mastitis (Erskine and Bartlett, 1993), herd udder hygiene from −1, −2, and −3 d were evaluated as the primary exposures of interest. Consequently, each row (i.e., cow-day at risk) in the final data set included variables that indicated whether the cow was diagnosed with clinical mastitis on that day and the herd udder hygiene values from each of the previous 1, 2, and 3 d (−1, −2, and −3 d, respectively).

**Statistical Analysis**

**Sample Size Calculation.** Sample size calculations were conducted for the clinical trial evaluating premilking teat disinfection, which was conducted during the same period (Rowe et al., 2018). No power calculations were conducted before undertaking the current study.

**Variable Management.** Herd demographic information and laboratory findings were recorded in Excel spreadsheets (Microsoft Corp., Redmond, WA), which, along with electronic herd records, were imported into the R Statistical Programming Environment (R Core Team, 2018) for analysis. Data sets and analysis log can be found at https://rpubs.com/samrowe/UHS_CM -2020. No data were missing, and consequently, no imputation methods were used. Variables were checked for missingness using the “summarytools” package in R (version 0.9.6). Variables available for potential inclusion in models were herd udder hygiene (main explanatory variable of interest), parity, age, clinical mastitis history at enrollment, DIM, and treatment group allocation for the clinical trial.

**Association Between Herd Udder Hygiene and Clinical Mastitis.** The relationship between herd udder hygiene on each of 1, 2, and 3 d before each study day (d −1, −2, and −3, respectively) and cow-level clinical mastitis on each study day (each in turn being set as d 0) was evaluated using Poisson regression (models 1, 2, and 3, respectively) using generalized estimating equations (GEE) to adjust for the clustering of cows within study day. Consequently, the exponent of β coefficients from models produced incidence rate ratios (IRR). Because all rows in the data set represented a single cow-day at risk, no offset was used. A directed acyclic graph was drawn to summarize our understanding of the underlying causal relationships between the variables available in our data set (Hernán and Robins, 2020). Variables that were hypothesized to cause clinical mastitis and not be affected by herd udder hygiene were included in all multivariable models as covariates to mitigate potential confounding. Consequently, fitted in the final multivariable models were parity, DIM, lifetime history of clinical mastitis before enrollment (0 = no cases, 1 = ≥1 case), and treatment group in the concurrently run clinical trial (0 = control, 1 = premilking teat disinfection). Cow age was highly correlated with parity (Pearson correlation coefficient >0.7) and was consequently excluded from all models to prevent multicollinearity. An independence working covariance structure was used. As a sensitivity analysis, equivalent models were run using different covariance structures within GEE in addition to using generalized linear mixed models with random intercepts for cow and study day, all of which produced similar estimates and standard errors to the Poisson GEE model (results in analysis log). Nonlinear relationships between the herd udder hygiene and clinical mastitis (on the log scale) were explored by fitting a quadratic term. If the
**Association Between Milking Order and Udder Hygiene.** The relationship between milking order (explanatory variable) and cow-level udder hygiene (outcome variable, 0 = scores 1 and 2, 1 = scores 3 and 4) was evaluated using log-binomial GEE (family = binomial, link = log). Consequently, the exponent of β-coefficients from the model produced risk ratios (RR). A milking order value (i.e., a percentile rank) was assigned for each cow and score (n = 12,544) by dividing the position of the score by the total number of scores for that milking and multiplying by 10. A multiplication factor of 10 was used so that the RR indicated the expected change in risk when the milking order percentile rank increases by 10 percentage points. For example, the 50th cow to be scored from a milking with 400 scores had a milking order value of 50/400 × 10 = 1.25 (percentile rank = 12.5%). Clustering of scores within milkings was controlled for using an independence working covariance structure. Model fit and assessment for nonlinearity on the log scale were assessed using the same approach outlined for models evaluating the relationship between herd udder hygiene and clinical mastitis.

**RESULTS**

**Descriptive Results**

A total of 31,383 cow-days from 504 cows were recruited into the study. The mean, median, minimum, and maximum days at risk for enrolled cows were 62.3, 73.5, 1, and 89 d. Average age at enrollment was 52.1 mo (SD = 21.4), and the proportion of cows for each parity group was as follows: 1, 33.7%; 2, 24.0%; 3, 21.0%; and ≥4, 21.2%. Average DIM at enrollment was 124 (SD = 112). As part of the concurrent clinical trial, 50.2 and 49.8% of cows received a premilking teat disinfection routine and the standard premilking routine (control), respectively. During the study period, 117 cases of clinical mastitis were recorded. Consequently, the incidence rate for clinical mastitis during the study period was 3.7 cases per 1,000 cow-days, or 12.3 cases per 100 cow-months.

A total of 12,544 udder hygiene scores from footage of 35 afternoon milkings were conducted as part of this study, with average of 358 cows scored at each milking. The number of cows with scores 1, 2, 3, and 4 were 158 (1.3%), 7,396 (59.0%), 3,813 (30.4%), and 1,177 (9.4%; Table 1). Consequently, the overall prevalence of poor udder hygiene (scores 3 and 4) was 39.8%. Herd udder hygiene (i.e., the proportion of cows with scores 3 and 4) at 35 milkings ranged from 12.8 to 68.3%, with a mean of 39.1% (SD = 14.2%).

Based on findings from the analysis described above, we compared true and estimated herd udder hygiene when a subset of 80 consecutive cows was used (i.e., the first, middle, and last 80 cows to be scored). The concordance between estimated and true herd udder hygiene was calculated using CCC and by calculating limits of agreement, as outlined earlier.

The number of cows with scores 1, 2, 3, and 4 were 158 (1.3%), 7,396 (59.0%), 3,813 (30.4%), and 1,177 (9.4%; Table 1). Consequently, the overall prevalence of poor udder hygiene (scores 3 and 4) was 39.8%. Herd udder hygiene (i.e., the proportion of cows with scores 3 and 4) at 35 milkings ranged from 12.8 to 68.3%, with a mean of 39.1% (SD = 14.2%).

P-value for the quadratic term was >0.05, then the linear model was retained. If the P-value was <0.05, then model fit for the linear and quadratic models was compared by visualizing each of regression curves (produced with estimated marginal means) and crude incidence rates of clinical mastitis each plotted against herd udder hygiene. Model validity was also evaluated by checking whether model-predicted probabilities for each row in the data set were between 0 and 1. All predicted probabilities for final models were between 0 and 1.

**Scoring Subsets of the Herd.** As a secondary objective, we aimed to compare herd udder hygiene using all available scores (referred to as true herd udder hygiene) and herd udder hygiene values based on subsets of the herd (referred to as estimated herd udder hygiene). Udder hygiene scores (n = 12,544) from each scored milking (n = 35) were used. For each milking, 13 estimated herd udder hygiene values were calculated, one each from a randomly selected subset of 10, 20, 30, 40, 50, 60, 80, 100, 150, 200, 250, 300, and 350 scored cows. Consequently, for each scored milking, there was 1 true herd udder hygiene value and 13 estimated herd udder hygiene values. For example, on the first day of the study, the true herd udder hygiene was 0.179 (i.e., 17.9% of cows had scores of 3 or 4), and the estimated herd udder hygiene values using 10, 20, 30, 40, 50, 60, 80, 100, 150, 200, 250, 300, and 350 of the scores were 0.20, 0.30, 0.20, 0.25, 0.22, 0.20, 0.13, 0.24, 0.19, 0.19, 0.18, 0.18, and 0.17, respectively. Agreement between true and estimated herd udder hygiene was evaluated using concordance correlation coefficients (CCC; Lawrence and Lin, 1992). Average differences and 99% limits of agreement (average difference ± 2.58 × standard deviation of differences) were calculated, such that nonzero average differences indicate bias and 99% limits of agreement indicate the variation in differences between milkings (i.e., precision; Bland and Altman, 1999). In addition, 99% confidence intervals were calculated for CCC, mean differences, and lower and upper limits of agreement values. We used 99% limits of agreement and 99% confidence intervals (i.e., a more conservative measure than 95%) because some degree of autocorrelation between days may have been present, and CCC and Bland-Altman calculations assumed that observations (i.e., study days) were independent from each other.

Based on findings from the analysis described above, we compared true and estimated herd udder hygiene when a subset of 80 consecutive cows was used (i.e., the first, middle, and last 80 cows to be scored). The concordance between estimated and true herd udder hygiene was calculated using CCC and by calculating limits of agreement, as outlined earlier.
Effect estimates from 3 multivariable GEE models estimating the effect of herd udder hygiene at −1, −2, and −3 d (models 1, 2, and 3, respectively) on clinical mastitis on d 0 are shown in Table 2. Incidence rate ratios (95% CI) for models 1, 2, and 3 were 1.4 (1.2, 1.6), 1.4 (1.3, 1.6), and 1.4 (1.2, 1.6), respectively. Herd udder hygiene was coded such that IRR indicated the expected change in incidence rate when herd udder hygiene changed by 0.1 (10%). Model-predicted incidence rates (estimated marginal means from model 1) for herd udder hygiene values ranging from 0.0 to 0.6 are shown in Figure 3. For example, according to model 1, the predicted incidence rates when 30 and 20% of cows had poor udder hygiene (i.e., scores ≥3) were 7.2 and 5.1 cases per 100 cow-months, respectively (i.e., 7.2/5.1 = IRR = 1.4).

### Scoring Subsets of the Herd

**Proportion of Cows Needed.** Figure 4 shows the deviations from true herd udder hygiene for each milking (n = 35) when subsets of the herd were used (estimated herd udder hygiene). As expected, when fewer cows were scored, the precision of estimated herd udder hygiene decreased, as evidenced by wider limits of agreement and lower CCC values. The 99% CI for mean differences ranged from ±0.003 (350 cows) to ±0.063 (10 cows). The 99% CI for lower and upper 99%

---

**Table 1.** Frequencies from 12,544 udder hygiene scores from 35 afternoon milkings during an 89-d period

<table>
<thead>
<tr>
<th>Score¹</th>
<th>No.</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>158</td>
<td>1.3</td>
</tr>
<tr>
<td>2</td>
<td>7,396</td>
<td>59.0</td>
</tr>
<tr>
<td>3</td>
<td>3,813</td>
<td>30.4</td>
</tr>
<tr>
<td>4</td>
<td>1,177</td>
<td>9.4</td>
</tr>
</tbody>
</table>

¹Udder contamination was scored using a 4-point system developed by Schreiner and Ruegg (2002): 1 = no dirt or manure; 2 = 2 to 10% of surface area of the udder covered with dirt or manure; 3 = 10 to 30% of surface area of the udder covered with dirt or manure; 4 = >30% of surface area of the udder covered with dirt or manure. The percentage of “dirty” scores (i.e., 3 or 4) was 39.8%.

**Table 2.** Results from 3 generalized estimating equation models (family = Poisson, link = log, independence working correlation structure) estimating the relationship between herd udder hygiene (% of cows with udder hygiene scores of ≥3 out of 4) and cow-level clinical mastitis rates

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>β</th>
<th>SE</th>
<th>Incidence rate ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Herd udder hygiene: 1 d ago</td>
<td>0.340</td>
<td>0.0642</td>
<td>1.4 (1.2, 1.6)¹</td>
</tr>
<tr>
<td>Model 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Herd udder hygiene: 2 d ago</td>
<td>0.358</td>
<td>0.0618</td>
<td>1.4 (1.3, 1.6)</td>
</tr>
<tr>
<td>Model 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Herd udder hygiene: 3 d ago</td>
<td>0.333</td>
<td>0.0669</td>
<td>1.4 (1.2, 1.6)</td>
</tr>
</tbody>
</table>

¹All models controlled for parity, DIM, lifetime history of clinical mastitis (0 = no cases, 1 = ≥1 case), and treatment group in the concurrently run clinical trial (0 = control, 1 = premilking teat disinfection). Coefficients for covariates are not shown for clarity but can be seen in the analysis log at https://rpubs.com/samrowe/UHS_CM-2020.

²For every 10% increment in herd udder hygiene (proportion of the herd on the previous day with udder hygiene scores of ≥3 out of 4), the rate of cow-level clinical mastitis increased by a factor of 1.4. For example, the incidence rate of clinical mastitis is expected to be 40% higher if the proportion of cows with poor udder hygiene increases from 20 to 30%.
limits of agreement ranged from ±0.005 (350 cows) to ±0.107 (10 cows). Our assessment was that the lowest acceptable number was 80 cows, which had a CCC of 0.96 and lower and upper limits of agreement of −0.07 and 0.07, respectively, and was within ±0.1 of true herd udder hygiene at all assessments.

Scoring at Different Stages of Milking. Figure 5 shows the deviations from true herd udder hygiene for each milking (n = 35) when subsets of 80 cows from different stages of milking (i.e., the first, middle, and last 80 cows to be scored) were used to calculate estimated herd udder hygiene. Concordance correlation coefficient values between estimated and true for the first, middle, and last 80 cows were 0.88, 0.95, and 0.87, respectively. Precision and accuracy were slightly worse when the first 80 cows (underestimated) and last 80 cows (overestimated) were used to estimate herd udder hygiene. This apparent trend was supported by results from a GEE model that found that the probability of a cow having poor udder hygiene (score ≥3) increased by a factor 1.02 for cows that were 10% later in milking order (RR = 1.02 per 10 percentage points; 95% CI: 1.01, 1.03).

DISCUSSION

Poor Herd Udder Hygiene was Associated with Clinical Mastitis Incidence Rates

Results from this study show that clinical mastitis rates increased by a factor of 1.4 for every 10% increase in the proportion of cows with poor udder hygiene. Effect estimates for −1, −2, and −3 d were the same (IRR = 1.4), likely because herd udder hygiene did not vary substantially over such a short time period and because linear interpolation was used to estimate herd udder hygiene on days that were not directly assessed.

These findings are consistent with Breen et al. (2009), who conducted a prospective cohort study of 1,677 cows from 8 UK dairy herds using the same udder hygiene scoring system as the current study (Schreiner and Ruegg, 2002). They found that cows with an UHS of 4 had 1.5 times higher odds of developing clinical mastitis during the next month compared with cows with UHS of ≤2. The studies are not directly comparable, as Breen et al. (2009) focused on udder hygiene at the individual level with assessments on individual
cows conducted once per month, whereas we evaluated the effect of udder hygiene at the herd level estimated each day from assessments every 2 to 3 d on cow-level risk of clinical mastitis 1 to 3 d later.

Our findings are also consistent with the only pasture-based study of udder hygiene and intramammary infection conducted to date, which found that early-lactation IMI was 1.3 times higher in first-lactation cows with UHS >1 at the time of sampling than in first-lactation cows with UHS of 1 (Compton et al., 2007).

At Least 80 Cows Were Needed for Accurate and Precise Estimation of Udder Hygiene

Our findings indicate that under conditions of this study, scoring fewer than 80 cows is likely to yield erroneous results. Scoring 80 cows was within ±0.1 of true herd udder hygiene in 35 out of 35 (100%) assessments. Scoring larger numbers of cows will provide marginal increases in sampling precision. These conclusions are consistent with epidemiologic sample size formulae outlined in Dohoo et al. (2003). For example, if the expected prevalence is 35%, approximately 87 cows of the herd are needed for a margin of error (i.e., half of the desired 95% CI width) of 0.1: \(1.96^2 \times \text{expected prevalence} \times (1 - \text{expected prevalence}) / \text{margin of error}^2 = 1.96^2 \times 0.35 \times 0.65 / 0.1^2 = 87 \text{ cows.}\) It should be noted that this number (87) is slightly larger than what was found using the simulation methods in this study (i.e., at least 80 cows). According to this formula, which assumes an infinite-sized population, herds with an expected prevalence of 50% would require 97 cows to achieve similar levels of precision. Furthermore, for herds with lower expected prevalences of poor udder hygiene (e.g., 10%), fewer cows would be required (n = 35). Therefore, for herds with prevalences ranging from 10 to 90%, 35 to 97 cows would need to be scored. It should be noted that for herds that are much smaller or larger than the study herd, the required numbers of cows would be a little less and more, respectively, than the 80 cows recommended for this herd, where, on average, 358 cows were assessed at each milking when scoring was performed.

Cows Milked Later Were More Likely to Have Poor Udder Hygiene than Cows Earlier in the Milking Order

Our findings indicate that cows that are milked later in the order may be more likely to have poor udder hygiene than cows milked earlier. Therefore, routine scoring of cows exclusively from early or late in the milking period may bias herd udder hygiene assessments in the long run, as evidenced by mean difference values not equaling zero. The finding has parallels with a study conducted in New Zealand by Sauter-Louis et al. (2004), which found that cows in the last quarter of the milking order had 1.5 times higher odds of being lame.

Practical Implications

Our findings help confirm a long-understood but poorly described risk factor for clinical mastitis in pasture-based dairy cows. These findings indicate that
prevention of udder contamination may prevent clinical mastitis in pasture-based dairy herds. Although very few surveys have explicitly identified risk factors for poor udder hygiene in pasture-based dairy cows, it is generally expected that wet conditions will cause paddocks, tracks, feeding areas, shaded areas, and other congestion points within farms to become waterlogged and potentially become sources of udder contamination. More research is needed to evaluate the efficacy of interventions to prevent udder contamination from these sources. Such interventions include strategic restriction of cow access to contaminated areas, including around trees, increasing the frequency of scraping areas with manure build-up, use of portable shade infrastructure, and improvement of drainage. Findings from several clinical trials in Australasia suggest that wholesale use of premilking teat disinfection is unlikely to be helpful in many pasture-based herds (Depiazzi and Bell, 2002; Williamson and Lacy-Hulbert, 2013; Morton et al., 2014; Rowe et al., 2018). Although teat end cleanliness is likely to be a more direct measure of clinical mastitis risk from environmental sources, it is much less practical than udder hygiene scoring methods used in this study.

**Study Strengths and Limitations**

We acknowledge that this study used observational data (i.e., exposures other than premilking teat disinfection protocol were not randomly allocated to subjects). Therefore, despite the use of carefully considered causal models in our analysis, unknown or uncontrolled confounders may have biased our effect estimates. One important potential confounder in this study is time. That is, herd-level risk factors for mastitis other than udder hygiene could have varied over time (e.g., milking machine function), such that background mastitis risk may not have been consistent across all herd udder hygiene levels. This potential source of bias is important to consider because udder hygiene was evaluated at the herd level and not at the cow level. One advantage of this prospective longitudinal single-herd study design is that confounding from non-time-varying herd-level variables is less likely than in multitherd, cross-sectional studies, which represent most studies that have evaluated the relationship between udder hygiene and udder health.

Misclassification of udder hygiene score is possible, given that cows were scored from video footage instead of cow-side. We are confident that the quality of the footage was sufficient to enable accurate scoring, and we encourage readers to view the footage on our analysis log webpage. One important limitation to the external validity of this study is that cows were from a single herd in a subtropical dairy farming region in Queensland, Australia, that was known to have seasonal challenges with environmental mastitis before the study. We encourage readers to consider the herd characteristics outlined in the Materials and Methods before generalizing results to other herds. Furthermore, we recommend that findings from our study be replicated in multiherd studies that include herds from different environmental and climatic conditions.

**Further Research**

Multiherd studies should be conducted to replicate the findings from this study. Further research is needed to identify risk factors for poor udder hygiene in pasture-based dairy farms and potential interventions to prevent udder contamination.

**CONCLUSIONS**

In a prospective longitudinal study of 504 cows from a pasture-based dairy herd, we demonstrated that increasing proportions of cows with poor udder hygiene were associated with higher rates of clinical mastitis in the following 1 to 3 d (RR = 1.4). Our findings indicate that for similarly sized herds (~380 cows), milk quality advisors should score at least 80 cows at milking time, and that scoring cows in the middle of the milking order or from a mixture of stages in the milking order will provide the most accurate estimates of herd udder hygiene.

**ACKNOWLEDGMENTS**

We thank Dennis Byrnes and the milking staff at Lakeside Dairy (Yungaburra, QLD, Australia). We also thank Richard Maclehose (University of Minnesota, Minneapolis) for advice on statistical approach, and we acknowledge the input of the anonymous reviewers whose comments significantly improved the paper. This study received financial support from Dairy Australia (Southbank, Victoria, Australia). Sam Rowe conducted fieldwork, data analysis, and manuscript writing. William Tranter and Richard Laven were involved in manuscript editing and supervision. The authors have no competing interests to declare.

**REFERENCES**


ORCIDs

Sam Rowe  https://orcid.org/0000-0001-8336-6523
Richard Laven  https://orcid.org/0000-0002-8938-8505