Identifying gram-negative and gram-positive clinical mastitis using daily milk component and behavioral sensor data

N. M. Steele,^{1,2*} A. Dicke,³ A. De Vries,⁴ S. J. Lacy-Hulbert,² D. Liebe,⁵ R. R. White,⁵ and C. S. Petersson-Wolfe¹

¹Department of Dairy Science, Virginia Tech, Blacksburg 24061

²DairyNZ Ltd., Private Bag 3221, Hamilton 3240, New Zealand

³Farm Credit, Bellefontaine, OH 43311

⁴Department of Animal Sciences, University of Florida, Gainesville 32611

⁵Department of Animal and Poultry Science, Virginia Tech, Blacksburg 24061

ABSTRACT

Opportunities exist for automated animal health monitoring and early detection of diseases such as mastitis with greater on-farm adoption of precision technologies. Our objective was to evaluate time series changes in individual milk component or behavioral variables for all clinical mastitis (CM) cases (ACM), for CM caused by gram-negative (GN) or gram-positive (GP) pathogens, or CM cases in which no pathogen was isolated (NPI). We developed algorithms using a combination of milk and activity parameters for predicting each of these infection types. Milk and activity data were collated for the 14 d preceding a CM event ($n = 170$) and for controls ($n = 166$) matched for breed, parity, and days in milk. Explanatory variables in the univariate and multiple regression models were the slope change in milk (milk yield, conductivity, somatic cell count, lactose percentage, protein percentage, and fat percentage) and activity parameters (steps, lying time, lying bout duration, and number of lying bouts) over 7 d. Slopes were estimated using linear regression between d -7 and -5 , d -7 and -4 , d -7 and -3 , d -7 and -2 , and d -7 and -1 relative to CM detection for all parameters. Univariate analyses determined significant slope ranges for explanatory variables against the 4 responses: ACM, GN, GP, and NPI. Next, all slope ranges were offered into the multivariate models for the same 4 responses using 3 baselines: d -10 , -7 , and -3 relative to CM detection. In the univariate analysis, no explanatory variables were significant indicators of ACM, whereas at least 1 parameter was significant for each of GN, GP, and NPI models. Superior sensitivity (Se) and specificity (Sp) estimates were observed for the best GP (Se = 82%, Sp = 87%) and NPI (Se = 80%, Sp

= 94%) multiple regression models compared with the best ACM (Se = 73%, Sp = 75%) and GN (Se = 71%, Sp = 74%) models. Sensitivity for the GN model was greater at the baseline closest to the day of CM detection (d -3), whereas the opposite was observed for the GP and NPI model as Se was maximized at the d -10 baseline. Based on this screening of relationships, milk and activity sensor data could be used in CM detection systems.

Key words: sensor data, pathogen type, slope change

INTRODUCTION

Mastitis remains an important and costly disease of dairy cows that has adverse effects on animal well-being and the quality and quantity of milk produced. The decision to treat a case of clinical mastitis (CM) is typically made without knowledge of the underlying cause of the disease because current diagnostic tools are based on culture or PCR technologies that delay the result by at least 1 d, limiting their use for prompt decision making. More than 130 bacterial species have been associated with bovine mastitis (Watts, 1988), but approximately 10 species, or groups of species, account for 95% of mastitis infections (Makovec and Ruegg, 2003). Most mastitis pathogens can be classified into gram-positive (GP) and gram-negative (GN) bacteria based on their cell wall characteristics. Prompt differentiation of CM cases as either GN or GP allows for more appropriate treatment and management approaches because GN infections are not considered suitable candidates for broad-spectrum antibiotic therapy (Pyörälä et al., 1994; Lago et al., 2011).

Currently, no cow-side diagnostic tool can differentiate CM cases caused by GN or GP upon detection. Up to 40% of CM cases in the United States are attributed to GN bacteria, with a further 10 to 40% of cases yielding no pathogen when cultured in the laboratory (Roberson, 2003). Considering that many producers treat

Received April 3, 2019.

Accepted November 6, 2019.

*Corresponding author: Nicole.Steele@dairynz.co.nz

most CM cases with antibiotic therapy regardless of the causative pathogen, up to 80% of the lactating cow antibiotics used to treat clinical cases could be deemed unnecessary (Roberson, 2012). Antimicrobial use in livestock production systems is under increasing scrutiny, and treatment of mastitis without prior knowledge of the pathogen could be considered imprudent.

Producers require accessible, comprehensible, and reliable tools to detect and diagnose CM to inform management decisions in their dairy herds. On-farm culture systems are increasingly available to provide a pathogen diagnosis, reducing the time to diagnosis to less than 1 d (Royster et al., 2014). However, these systems require training to assist interpretation and will not be suitable for all producers. The use of precision technologies on-farm is increasing, providing producers with new approaches for closely monitoring changes in cow performance and behavior (Rutten et al., 2013). The large amount of data generated daily provides an opportunity to monitor individual animals and may enable early detection of diseases such as mastitis. Most cow-side tests for detecting mastitis rely on changes in the milk, including SCC, sodium and chloride ion concentrations, or abundance of enzymes relating to tissue damage (Kitchen, 1981). Daily milk yield or lactose, protein, and fat components may also change in the days before visual detection of a CM case while remaining relatively stable over time in healthy cows (Quist et al., 2008; Forsbäck et al., 2010). Cow activity, including steps and lying behavior, also changes before detection of CM (Yeiser, 2011; Stangaferro et al., 2016; King et al., 2018). Individually, these changes may not be specific enough to indicate CM, but in combination, the accuracy of mastitis detection systems could be improved. Earlier detection and differentiation of CM cases could lead to quicker recovery for the cow and reduced economic losses for the producer (Milner et al., 1997).

The severity of mastitis is often related to the causative pathogen. Infections by GN tend to induce a more severe inflammatory response, potentially with systemic involvement, whereas many GP pathogens tend to cause a milder or delayed inflammatory response (Schukken et al., 2011). These differences are likely to be reflected in milking performance, milk components, and behavioral parameters in the lead up to clinical presentation of mastitis, providing an opportunity for mastitis detection using sensors in conventional milking systems as well as automatic milking systems. Few researchers have used sensor data to categorize CM cases by gram status, where electrical conductivity and milk color (Kamphuis et al., 2011) or parity, DIM, and CM or SCC history (Steenefeld et al., 2009) have been used as inputs for predictive models. With access to more

sensor inputs, such as milk components, the detection performance was improved for CM regardless of pathogen type (Jensen et al., 2016). We proposed that the inclusion of a wider range of sensor inputs in the week before CM detection and the separation of CM cases by the type of pathogen causing infection could generate models that can detect mastitis early and with improved accuracy. Therefore, the objective of this study was to evaluate whether time series changes in individual milk components or activity variables were indicators of CM events categorized by pathogen type (GN, GP, or no pathogen isolated; **NPI**) compared with all CM cases grouped together (**ACM**). Subsequently, a combination of milk and activity parameters were then used to develop separate multiple regression models to identify ACM, GN, GP, or NPI cases.

MATERIALS AND METHODS

Study Design

A data set containing milk yield, quality, and composition as well as activity parameters for the 14 d preceding CM detection by milkers was collated for 170 CM events and 166 matched controls (Tholen, 2012). All lactating cows from the Virginia Tech (**VT**) Dairy Center (n = 230; Blacksburg, VA) and the University of Florida (**UF**) Dairy Unit (n = 480; Gainesville, FL) were available for enrollment in a case-control observational study. Data were collected at VT for 12 mo beginning in March 2011 and at UF for 9 mo beginning in June 2011 (Tholen, 2012). Cows were milked twice daily at 12-h intervals in a double-8 herringbone parlor at VT or a double-12 herringbone parlor at UF. The herd consisted of Holstein, Jersey, and crossbred animals at VT and Holstein animals at UF. All animals were housed in freestall barns, with sawdust bedding at VT and sand bedding at UF. Bedding was changed twice weekly at both facilities.

Case and Control Definitions

A CM case was defined as a cow with CM in one quarter (i.e., abnormal milk with clots or flakes, with or without redness and swelling of the mammary gland) as identified by trained milking staff during premilking teat preparation. At the onset of CM, the case cow had to be ≥ 15 DIM and ≤ 291 DIM and free of clinical signs for at least 14 d before enrollment. Only the first case of CM per lactation per cow was included in the data set. A control animal was defined as not having visible signs of CM in the current lactation and was retrospectively assigned to each case based on DIM, breed, parity, and farm. Once a cow was enrolled as

a case animal, that cow was not eligible to serve as a control for the remainder of that lactation. A single cow could be used as a control for multiple cases. Because controls were retrospectively assigned to cases, milk samples were not collected from these animals to confirm absence of subclinical mastitis.

Milk Sample Collection and Diagnosis

Upon detection of CM, a quarter milk sample was aseptically collected by farm personnel following National Mastitis Council guidelines (Middleton et al., 2017). Samples were stored at -20°C until transportation to the Virginia Tech Mastitis and Immunology Laboratory (Blacksburg, VA). Samples from VT were transported on ice once weekly to the laboratory, and samples from UF were shipped once monthly on ice in an insulated cooler for overnight delivery.

Bacteriological procedures were completed in accordance with National Mastitis Council guidelines (Middleton et al., 2017). Briefly, 10 μL of milk was streaked on one quadrant of an esculin blood agar plate, and 100 μL of milk was streaked on one half of a MacConkey agar plate. Both plates were incubated at 37°C for 48 h. Bacterial identification was confirmed by colony morphology and biochemical testing (Middleton et al., 2017). A sample was considered contaminated if 3 or more dissimilar colony types grew in culture (Middleton et al., 2017). Samples that did not isolate bacteria under aerobic culture were cultured anaerobically for the presence of *Mycoplasma* spp.

Each CM case was assigned to 1 of 4 groups depending on the type of pathogen isolated: (1) GP pathogens, which included *Staphylococcus aureus*, coagulase-negative *Staphylococcus* spp., and *Streptococcus* spp.; (2) GN pathogens, which included *Klebsiella* spp., *Escherichia coli*, *Citrobacter* spp., *Enterobacter* spp., *Serratia* spp., *Pseudomonas* spp., and *Proteus* spp.; (3) other pathogens, which included *Prototheca*, yeast, and unknown microorganisms; and (4) NPI, in which no pathogens were isolated under aerobic or anaerobic culture. Cases that were contaminated were excluded, and cases with mixed isolations (i.e., more than 1 pathogen isolated from a sample regardless of whether the bacterial group was the same) were not available for inclusion in the study.

Data Collection

The VT and UF dairy farms were both equipped with herd management software (AfiFarm, Afimilk Ltd., Kibbutz Afikim, Israel) for data storage. Milk components (milk protein, fat, and lactose percentages) and SCC were determined at each milking at the cow level

using an in-line milk analyzer (AfiLab, Afimilk Ltd.). This system has been demonstrated to have moderate correlations with milk fat, protein, and lactose percentages determined by DHIA using the Bentley 2000 instrument (Kaniyamattam and De Vries, 2014), though no correlations were available for SCC measurement. The AfiFarm interface typically presents SCC data in categorical terms, but for this study, continuous numerical SCC data were retrieved from the Afimilk Ltd. database. Milk yield and electrical conductivity were determined at the cow level at each milking (AfiMilk MPC, Afimilk Ltd.).

Activity data were collected using pedometers attached to a rear-leg fetlock on each cow. At VT, the pedometers were capable of measuring cow activity as steps per hour as well as daily lying time, number of lying bouts, and lying bout duration (AfiPedometerPlus, Afimilk Ltd.). A lying bout was recorded only if the duration of inactivity was >3 min. At UF, each cow had a pedometer that measured only steps per hour (AfiPedometer, Afimilk Ltd.). Measurement of lying time and activity by the AfiPedometerPlus system has been previously evaluated (Higginson et al., 2010; Borchers et al., 2016). Milk and activity data were transmitted at milking time to the herd management software.

Milk and activity data were extracted for the 14 d preceding a CM case for the clinical cow and for the same period for the matched control cow. All data were collated into a database (Microsoft Excel, Microsoft Corp., Redmond, WA) and matched with the relevant bacteriological data. For each cow, electrical conductivity and milk components were averaged for the 2 milking sessions within a day. Hourly steps and lying bout duration were averaged across all measurements for a 24-h period. Milk yield for the 2 sessions as well as lying time and the number of lying bouts measured in a 24-h period were summed to provide daily measurements. Somatic cell count was averaged across the 2 sessions within a day and was log-transformed (log-SCC). Analyses of lying time, lying bout duration, and number of lying bouts were limited to VT cases and controls because these data were not available for UF.

Determination of Changes in Variables Before CM Detection

Ten explanatory variables were a priori considered for testing. These variables included daily measurements of milk lactose, protein, and fat percentages, milk yield, electrical conductivity, SCC, average activity, number of lying bouts, average lying bout duration, and total lying time. The slopes of each explanatory variable were estimated for each cow using linear regression in R ver-

sion 3.4 (R Core Team, 2017) for days ranging between $d - 7$ and $d - 5$, -4 , -3 , -2 , or -1 relative to CM detection ($d = 0$). We selected 7 d before CM detection as the appropriate baseline for our univariate models because we anticipated that this day range would best capture the changes occurring in the lead up to clinical infection. Slopes of explanatory variables rather than absolute values were used to identify changes within cows over time, using each cow as her own baseline. For example, the slope of lactose percentage between $d - 7$ and -1 was estimated by linear regression using values at $d - 7$, -6 , -5 , -4 , -3 , -2 , and -1 to represent the overall change in lactose over that time, whereas the slope of lactose percentage between $d - 7$ and -5 used lactose percentage at $d - 7$, -6 , and -5 to estimate the overall change within this day range. Unavailability of data, either due to data not being collected for some of the variables on one farm or due to system failures leading to missing data, had the potential to affect slope calculation for cases and controls. Missing values ranged from 0.4% to 10.4% depending on the variable (data not shown). As a result of missing data, the number of cases and matched controls used in model development was different for each model depending on the variables that remained significant in the model.

Univariate Models

Univariate logit models were fit for each combination of response and explanatory variables using all slope ranges for each explanatory variable. Four response variables were tested: ACM, GN, GP, and NPI. For the ACM univariate models, the response variable was all CM cases regardless of the cause, including NPI cases. Cows identified with CM were assigned a code of 1, and all controls were assigned a code of 0. For the GN models, the response was CM caused by GN pathogens; these cases were assigned a code of 1, and all other CM cases and controls were assigned a code of 0. Likewise, for the GP models, the response was CM caused by GP pathogens; these cases were assigned a code of 1, and all other cases and controls were assigned a code of 0. For the NPI model, the response variable was NPI; these cases were assigned a code of 1, and all other CM cases and controls were assigned a code of 0. For each response variable, 50 univariate models were tested (1 model for each of the 5 slope ranges for each of the 10 independent variables considered). In all models, farm was included as a random effect.

Models were compared using a 3-part procedure. First, the optimal slope range ($d - 7$ to -5 , -4 , -3 , -2 , or -1) was identified for each explanatory variable–response variable combination. The model with the lowest

P -value was selected as the optimal slope range for that independent–dependent variable combination. Next, an optimal cutoff was identified for each model. Defining an optimal cutoff was necessary because the response of interest (infected vs. healthy) was binary; however, the model predicts a continuous spectrum of values between 0 and 1. To identify a threshold prediction value above which an animal was considered infected, all possible thresholds between 0 and 1 were tested using a step size of 0.01. For each step, model accuracy was summarized using the number of correct identifications as a proportion of all predictions. A correct identification occurred when the model prediction was above the threshold for a clinical case or below the threshold for a noncase (a control or a case caused by a pathogen other than the response of interest for that particular model). The step value with the largest proportion of correct identifications (greatest accuracy) was used as the permanent threshold. Finally, to compare these models in terms of their ability to identify infections, the sensitivity (**Se**) and specificity (**Sp**) were calculated for each model. The Se and Sp values were then compared across explanatory variables to determine the strength of the relationships between the different explanatory variables and the responses (ACM, GN, GP, and NPI).

Multiple Regression Models

Separate multiple regression models were derived with the same response variables of interest as the univariate models: ACM, GN, GP, and NPI. Independent variables considered for inclusion in these models were DIM, parity, breed, BW, and the slope variables for milk and activity changes described previously. Mixed-effect logit models were used to relate response and explanatory variables. Farm was included as a random effect. Following procedures outlined in Roman-Garcia et al. (2016) and White et al. (2016), backward stepwise elimination was used to derive the models where all variables remaining were $P \leq 0.1$. Variance inflation factor (**VIF**) was used to check for multicollinearity of explanatory variables remaining in the model after backward elimination. Any explanatory variable with $VIF > 5$ was removed from the model.

Multivariate models for the 4 infection types were derived using 3 different baselines because the magnitude and timing of alterations in cow performance and behavior are expected to differ for different pathogen types before clinical signs of infection. First, all parameters for all slope ranges using $d - 7$ as the baseline (as in the univariate models) were offered into the model and backward elimination was used to reach the final model. Next, separate multivariate models were

developed using d -3 and -10 as baselines to elucidate whether changes occurring closer to the day of CM detection or more than 1 wk before detection were important for the different pathogen types. For the models using d -3 as the baseline, slopes were generated using linear regression for each variable between d -3 and -2 , -1 , or 0 . When d -10 was used as the baseline, slopes were generated between d -10 and -8 , -7 , -6 , -5 , -4 , or -3 relative to CM detection. All multiple regression models were evaluated using a procedure similar to that described for the univariate models. First, an optimal cut-off was identified by testing all possible thresholds between 0 and 1 using a step size of 0.01 . For each step, the number of correct identifications was identified and the threshold value was defined based on the greatest accuracy. The Se and Sp of each model were calculated to compare the different baselines tested within pathogen type.

RESULTS AND DISCUSSION

Data collected in a case-control study (Tholen, 2012) provided a unique data set for analyzing changes in milk and activity parameters before a CM case with known bacteriology. Studies comparing the physiological changes that precede a natural CM event are lacking due to the difficulty of predicting when a clinical case will occur and the logistics of identifying the causative pathogen. The increasing availability of precision technologies provides greater opportunities to measure milk and activity parameters for each cow at every milking and to identify deviations from normal, which may be indicative of disease. Further, we proposed to categorize CM cases into groups based on pathogen type to investigate whether changes occurring before CM detection were different for CM cases caused by GN bacteria or GP bacteria or for NPI cases.

Data Summary

Data were obtained for a total of 336 animals (170 cases and 166 controls) for the 14 d preceding detection of a CM case (Table 1). Of the CM case animals, 67 (39%) were from VT and 103 (61%) were from UF; of the control cows, 74 (45%) were from VT and 92 (55%) were from UF. The numbers of cases and controls used in analyses differed by farm due to underlying incidence of CM on each farm and exclusion of cases or controls with insufficient data records. Most cows were Holstein ($n = 255$), with the remainder being Jersey ($n = 29$) or mixed breed ($n = 52$), and 74% of animals were multiparous ($n = 246$). When categorized by bacterial group, most CM cases had NPI ($n = 90$), followed

by GN ($n = 37$), GP ($n = 35$), and other pathogens ($n = 8$). The majority of GP pathogen isolations were *Streptococcus* spp. (57%), whereas *E. coli* (46%) was the most frequent GN pathogen isolated.

Univariate Models

For each combination of response and explanatory variable, the model for the slope range that had the lowest P -value is presented in Table 2. All slope day ranges tested for each explanatory variable for the ACM, GN, GP, and NPI responses are presented in Supplemental Tables S1, S2, S3, and S4 (<https://doi.org/10.3168/jds.2019-16742>), respectively. Analysis of the 10 explanatory variables (slopes of protein percentage, lactose percentage, fat percentage, logSCC, milk yield, conductivity, activity, lying bouts, lying time, and lying bout duration) in univariate models revealed no significant indicators of CM when all cases were considered together (Table 2). Classification into GN, GP, and NPI indicated some significant indicators of CM caused by these pathogen types.

For CM caused by GN pathogens, changes in step activity, lying time, conductivity, and lactose correlated significantly with instances of CM. An increase in lying time between d -7 and -1 relative to CM detection

Table 1. Number of cases, split into pathogen type and bacteriological outcome, and controls in the model training data set, categorized by farm

Bacteriological outcome	Farm ¹		Total
	VT	UF	
Gram-positive pathogens			
<i>Staphylococcus aureus</i>	5	3	8
CNS	0	7	7
<i>Streptococcus</i> spp.	11	9	20
Total gram-positive	16	19	35
Gram-negative pathogens			
<i>Klebsiella</i> spp.	12	0	12
<i>Escherichia coli</i>	11	6	17
<i>Citrobacter</i> spp.	2	2	4
<i>Enterobacter</i> spp.	0	1	1
<i>Serratia</i> spp.	1	0	1
Unidentified gram-negative	1	1	2
Total gram-negative	27	10	37
Other pathogens			
<i>Prototheca</i> spp.	0	3	3
Unknown microorganisms	1	1	2
Yeast	2	1	3
Total other pathogens	3	5	8
No pathogen isolated	21	69	90
Total clinical mastitis cases	67	103	170
Total controls	74	92	166
Total cases and controls	141	195	336

¹VT = Virginia Tech Dairy Center (Blacksburg, VA); UF = University of Florida Dairy Unit (Gainesville, FL).

Table 2. Univariate models for each combination of response and explanatory variable at the slope range that had the lowest *P*-value using -7 d relative to clinical mastitis detection as the baseline¹

Response ²	Explanatory variable	Slope range	No. of cases	Intercept ($\pm$ SE)	Slope estimate ($\pm$ SE)	<i>P</i> -value	Cutoff	Accuracy (%)	Se (%)	Sp (%)
ACM	Fat %	d -7 to -3	147	0.01 $\pm$ 0.12	-1.12 $\pm$ 0.70	0.11	0.50	55.3	62.6	48.0
	Lactose %	d -7 to -5	129	0.01 $\pm$ 0.13	-1.32 $\pm$ 1.79	0.46	0.52	54.5	20.9	88.2
	Protein %	d -7 to -1	157	0.03 $\pm$ 0.11	-3.27 $\pm$ 2.97	0.27	0.51	52.9	49.0	56.9
	Conductivity	d -7 to -1	157	0.05 $\pm$ 0.12	-0.79 $\pm$ 0.59	0.18	0.52	55.8	36.3	75.8
	Milk yield	d -7 to -1	157	0.06 $\pm$ 0.12	0.09 $\pm$ 0.08	0.27	0.46	53.5	95.5	10.5
	LogSCC	d -7 to -4	143	0.04 $\pm$ 0.12	-0.26 $\pm$ 0.28	0.34	0.56	51.1	4.2	99.3
	Activity	d -7 to -1	157	0.01 $\pm$ 0.12	-0.03 $\pm$ 0.02	0.13	0.54	53.2	12.7	94.8
	Lying time	d -7 to -1	54	-0.15 $\pm$ 0.19	0.01 $\pm$ 0.01	0.17	0.49	61.2	40.7	79.0
	Lying bouts	d -7 to -4	52	-0.13 $\pm$ 0.19	0.10 $\pm$ 0.12	0.42	0.49	57.1	26.9	83.3
	Lying bout duration	d -7 to -1	54	-0.16 $\pm$ 0.19	0.05 $\pm$ 0.06	0.43	0.49	58.6	27.8	85.5
	Fat %	d -7 to -5	33	-2.10 $\pm$ 0.57	0.94 $\pm$ 0.71	0.19	0.16	61.9	75.8	59.8
	Lactose %*	d -7 to -2	37	-2.08 $\pm$ 0.52	8.57 $\pm$ 4.39	0.05	0.16	68.4	73.0	67.8
	Protein %	d -7 to -5	33	-2.09 $\pm$ 0.55	3.12 $\pm$ 2.10	0.14	0.13	60.3	78.8	57.6
	Conductivity*	d -7 to -4	35	-2.15 $\pm$ 0.56	-1.48 $\pm$ 0.69	0.03	0.11	63.3	82.9	60.5
GN	Milk yield	d -7 to -3	35	-2.14 $\pm$ 0.56	0.12 $\pm$ 0.09	0.20	0.08	61.3	82.9	58.4
	LogSCC	d -7 to -3	35	-2.12 $\pm$ 0.57	-0.30 $\pm$ 0.55	0.59	0.16	62.2	77.1	60.2
	Activity*	d -7 to -5	33	-2.17 $\pm$ 0.60	0.03 $\pm$ 0.01	0.01	0.12	62.3	81.8	59.4
	Lying time*	d -7 to -1	23	-1.48 $\pm$ 0.25	0.02 $\pm$ 0.01	0.05	0.22	72.4	52.2	77.4
	Lying bouts	d -7 to -5	22	-1.40 $\pm$ 0.24	-0.13 $\pm$ 0.11	0.26	0.18	46.4	81.8	37.5
	Lying bout duration	d -7 to -5	22	-1.40 $\pm$ 0.24	0.03 $\pm$ 0.03	0.31	0.17	37.3	95.5	22.7
	Fat %	d -7 to -2	31	-2.18 $\pm$ 0.19	-1.10 $\pm$ 1.56	0.48	0.10	42.9	64.5	40.4
	Lactose %	d -7 to -3	30	-2.19 $\pm$ 0.20	-3.94 $\pm$ 4.19	0.35	0.11	72.9	36.7	77.0
	Protein %	d -7 to -3	30	-2.20 $\pm$ 0.20	4.27 $\pm$ 3.87	0.27	0.12	79.7	23.3	86.2
	Conductivity	d -7 to -1	31	-2.18 $\pm$ 0.19	-1.27 $\pm$ 0.91	0.16	0.11	75.2	32.3	79.9
	Milk yield	d -7 to -5	28	-2.13 $\pm$ 0.21	0.12 $\pm$ 0.08	0.15	0.12	73.9	39.3	78.2
	LogSCC	d -7 to -5	28	-2.10 $\pm$ 0.20	-0.29 $\pm$ 0.37	0.42	0.13	83.5	14.3	92.1
	Activity	d -7 to -5	28	-2.15 $\pm$ 0.21	-0.01 $\pm$ 0.01	0.10	0.11	71.2	42.9	74.7
	Lying time*	d -7 to -5	11	-2.42 $\pm$ 0.38	0.01 $\pm$ 0.01	0.03	0.10	64.5	72.7	63.6
NPI	Lying bouts	d -7 to -5	11	-2.27 $\pm$ 0.34	0.17 $\pm$ 0.14	0.20	0.10	57.3	63.6	56.6
	Lying bout duration	d -7 to -2	12	-2.22 $\pm$ 0.32	0.05 $\pm$ 0.07	0.48	0.12	82.8	33.3	88.5
	Fat %*	d -7 to -5	64	-1.22 $\pm$ 0.39	-1.49 $\pm$ 0.60	0.01	0.19	56.8	85.9	47.2
	Lactose %	d -7 to -2	81	-1.15 $\pm$ 0.44	-5.14 $\pm$ 3.52	0.14	0.22	58.8	76.8	52.3
	Protein %	d -7 to -3	77	-1.15 $\pm$ 0.43	-2.71 $\pm$ 2.63	0.30	0.29	59.8	74.0	54.7
	Conductivity	d -7 to -4	74	-1.18 $\pm$ 0.41	0.77 $\pm$ 0.44	0.08	0.31	65.7	62.2	67.0
	Milk yield	d -7 to -5	64	-1.18 $\pm$ 0.38	-0.08 $\pm$ 0.06	0.19	0.29	64.6	65.6	64.2
	LogSCC	d -7 to -3	77	-1.15 $\pm$ 0.43	-0.16 $\pm$ 0.35	0.65	0.18	59.1	77.9	52.3
	Activity*	d -7 to -3	77	-1.19 $\pm$ 0.42	-0.03 $\pm$ 0.01	0.02	0.27	60.6	76.6	54.9
	Lying time	d -7 to -5	16	-1.79 $\pm$ 0.28	-0.00 $\pm$ 0.00	0.39	0.13	39.1	87.5	30.9
	Lying bouts	d -7 to -5	16	-1.82 $\pm$ 0.28	0.15 $\pm$ 0.12	0.20	0.18	80.9	31.3	89.4
	Lying bout duration	d -7 to -2	17	-1.73 $\pm$ 0.26	-0.05 $\pm$ 0.08	0.47	0.16	67.2	41.2	71.7

¹For each model, the number of cases, intercept and slope estimates ($\pm$ SE), and the accuracy, sensitivity (Se), and specificity (Sp) at the optimal cutoff are presented. Models for all slope ranges are presented in Supplemental Tables S1, S2, S3, and S4 (<https://doi.org/10.3168/jds.2019-16742>).

²ACM = all clinical mastitis cases; GN = gram-negative cases; GP = gram-positive cases; NPI = cases with no pathogen isolated.

*Model $P \leq 0.05$.

provided an Se of 52% and an Sp of 77%, whereas increasing step activity between d -7 and -5 had an Se of 82% and an Sp of 59% for CM caused by GN pathogens (Table 2). Increased lying time is an indicator of sickness behavior in cows (Fogsgaard et al., 2015), supporting our finding of more time spent lying over the 7 d preceding clinical diagnosis. One study demonstrated reduced lying time at d -2 and -1 relative to naturally occurring CM compared with uninfected control animals, which might be indicative of udder pain causing the cow to lie down less (Yeiser, 2011). Although our data suggest an increase in step activity over d -7 to -5 relative to CM diagnosis, other studies have shown the opposite trend when comparing mastitic cows with healthy cows in the 5 d preceding CM diagnosis. Stangaferro et al. (2016) reported reduced activity in cows from 5 d before diagnosis of *E. coli* CM compared with healthy cows. Similarly, mastitic cows had reduced activity compared with healthy counterparts in the 3 d before diagnosis of CM caused by any pathogen (King et al., 2018). Reduced activity closer to CM detection makes biological sense, especially if the cow is spending more time lying. Using a cow as her own baseline is a useful approach because lying and activity behavior are highly variable among cows. However, activity alone is not a specific indicator of CM because changes can be associated with several other diseases (Edwards and Tozer, 2004) and with estrus (Kiddy, 1977).

The statistically significant milk parameters in the GN univariate analysis were an increase in lactose percentage between d -7 and -2 and a reduction in conductivity between d -7 and -4 relative to CM detection. The Se and Sp estimates for lactose slope were 73 and 68%, respectively, and for conductivity were 83 and 61%, respectively (Table 2). This finding of increased lactose percentage and decreased conductivity contrasts with our understanding of the biological processes occurring in the mammary gland around CM, where lactose percentage declines and conductivity increases for mastitic cows (reviewed by Kitchen, 1981). King et al. (2018) reported greater maximum conductivity (highest quarter-level conductivity) in mastitic cows compared with nondiseased cows from 14 d before CM diagnosis. Mastitic cows also deviated from baseline levels of conductivity 12 d before CM diagnosis (King et al., 2018). In our study, milk constituents were measured at the cow level, so changes occurring within one infected gland were diluted by other quarters. Nevertheless, the opposite conductivity effect seen here was unexpected. When d -3 (relative to CM detection) was used as the baseline, lactose declined between d -3 and -1 and between d -3 and 0, and the conductivity slope increased over the same day ranges (data not shown). Although the d -3 baseline slopes reflect the expected

changes in milk around CM, these were not significant indicators of GN CM.

An increase in lying time between d -7 and -5 relative to diagnosis was the only significant individual indicator of CM caused by GP pathogens (Table 2). This model correctly identified 65% of all cases and controls, with an Se of 73% and an Sp of 64%. Fewer parameters were found to be significant indicators of GP compared with GN. In the study by Stangaferro et al. (2016), the changes in activity were less pronounced for GP CM than for GN CM when comparing infected and healthy cows. Mastitis is typically considered to be an acute health disorder (King et al., 2018); however, some cases of mastitis are chronic in nature (e.g., *Staph. aureus*) or may exist in subclinical form until cure or until developing into CM (Schukken et al., 2011). In such cases, it is difficult to represent a cow's baseline because deviations from normal may already be occurring 7 d before clinical signs are detected. The use of an earlier baseline (e.g., 10 d before CM detection) resulted in more important parameters for detecting GP. There was a tendency for lactose percentage, conductivity, lying time, and milk yield changes between d -10 and -3 relative to CM detection to be associated with GP (data not shown).

Reduced milk fat percentage between d -7 and -5 and reduced step activity between d -7 and -3 relative to CM detection were significant indicators of NPI cases, with Se between 77 and 86% and Sp between 47 and 55% (Table 2). A culture-negative result from a truly clinical milk sample can occur for many reasons. Bacterial concentrations may be lower than the detection limit in culture through either low shedding levels or death of bacteria during inappropriate storage conditions. Alternatively, bacteria growth may be inhibited due to unsuitable medium, culture conditions, or the presence of antibiotics (Middleton et al., 2017). For GN infections, signs of inflammation can persist after the bacteria have been eliminated by the immune system, leading to no pathogen being isolated from a clinically mastitic milk sample (Pyörälä et al., 1994). These NPI cases are thought to be more like GN cases for this reason. However, our findings suggest that NPI cases were unique because there were differences in the significant parameters and direction of the relationship. The current study had 90 CM cases out of 170 with NPI in culture, at a rate of 53%. The proportion of culture-negative cases was greater in CM samples from UF (67%) compared with VT (31%), which may be at least partially explained by the increased shipping distance for the samples from UF. Overall, the rate of culture-negative samples collected from CM cases in this study was greater than that reported in the literature (10–40%; Roberson, 2003; Taponen et al., 2009).

Multiple Regression Models

Three different baselines (d -10, -7, and -3 relative to CM detection) were used to develop models for the 4 responses of interest. All but 1 of the models had some explanatory variables remaining in the model with $P < 0.1$ and VIF < 5 . For the ACM model (Table 3), the greatest detection accuracy was achieved using the d -10 baseline model (reflecting the changes occurring between d -10 and -3 relative to CM detection). These parameters included a change in conductivity, lactose percentage, and SCC between d -10 and -7 and between d -10 and -8 as well as protein percentage between d -10 and -8 and number of lying bouts between d -10 and -3 relative to CM detection. This model correctly identified 74% of cases and controls, with an Se of 73% and an Sp of 75%. Lactose and protein percentages were important across all baselines tested, indicating that changes in these parameters are consistently occurring before CM caused by any pathogen type. Lactose is the main osmotic regulator of milk; changes during mastitis are well documented (Kitchen, 1981; Auldist et al., 1995; Pyörälä, 2003). Lactose percentage declines due to reduced synthesis by mammary epithelial cells as well as leakage into interstitial fluid as tight junction integrity becomes compromised (Kitchen, 1981). Milk proteins decline due to reduced synthesis and degradation by bacterial products, whereas serum proteins increase due to tight junction leakiness (Auldist et al., 1995). The overall change in protein percent-

age is dependent on the degree of damage sustained by the mammary epithelium. With greater technology uptake on modern dairy farms, daily measurement of lactose and protein percentages is becoming more common, enabling incorporation of these components into mastitis detection models.

For the GN model, changes in fat and lactose percentages remained in models for all 3 baselines, and the change in activity was important for d -10 and -7 baselines (Table 4). The most accurate model for predicting both cases and controls was the d -10 model, with 82% accuracy, an Se of 41%, and an Sp of 88%. However, the Se was greater for the model incorporating data from the 3 d before CM detection (73%), but Sp was less (74%) than for the d -10 baseline. Model Se increased as the baseline moved closer to the day of CM diagnosis. The d -3 baseline may better capture the drastic changes occurring immediately before CM caused by GN pathogens because these typically cause acute infections with rapid induction and regression (Schukken et al., 2011).

The most accurate GP model also used d -10 as the baseline (Table 5). Explanatory variables in this model included breed and changes in activity (d -10 to -4), number of lying bouts (d -10 to -3), lying bout duration (d -10 to -8), and conductivity (d -10 to -7 and d -10 to -8). The accuracy was 86%, with an Se of 82% and an Sp of 87%. The d -7 baseline was the next best model at identifying GP, with 80% accuracy, 73% Se, and 81% Sp. In contrast to the

Table 3. Final multivariate regression models for all clinical mastitis cases grouped together using 3 baselines (d -10, -7, and -3 relative to clinical mastitis detection)¹

Baseline	Model	Slope estimate ($\pm$ SE)	P-value	Cases (no.)	Cutoff	Accuracy (%)	Se (%)	Sp (%)
d -10	Intercept	0.0 $\pm$ 0.2	0.89	51	0.48	74.1	72.5	75.4
	Lying bouts, d -10 to -3	-0.9 $\pm$ 0.4	0.02					
	Conductivity, d -10 to -7	-5.1 $\pm$ 1.7	<0.01					
	Conductivity, d -10 to -8	3.6 $\pm$ 1.2	<0.02					
	Lactose %, d -10 to -7	-24.9 $\pm$ 8.6	<0.01					
	Lactose %, d -10 to -8	17.9 $\pm$ 5.9	<0.01					
	Protein %, d -10 to -8	9.1 $\pm$ 3.2	<0.01					
	LogSCC, d -10 to -7	-2.9 $\pm$ 1.0	<0.01					
d -7	LogSCC, d -10 to -8	1.6 $\pm$ 0.7	0.02	51	0.37	66.4	88.2	47.5
	Intercept	-0.1 $\pm$ 0.2	0.74					
	Lying time, d -7 to -1	0.0 $\pm$ 0.0	0.10					
	Conductivity, d -7 to -5	-1.2 $\pm$ 0.5	0.03					
	Lactose %, d -7 to -2	21.5 $\pm$ 8.8	0.01					
d -3	Protein %, d -7 to -1	12.5 $\pm$ 8.4	0.14	46	0.4	64.6	76.1	54.7
	Intercept	-0.2 $\pm$ 0.2	0.41					
	Lying time, d -3 to -1	0.0 $\pm$ 0.0	0.05					
	Lactose %, d -3 to -2	3.9 $\pm$ 1.9	0.04					
	Protein %, d -2 to -1	5.2 $\pm$ 2.6	0.04					
	LogSCC, d -3 to -2	0.6 $\pm$ 0.3	0.02					

¹For each model, the slope estimate ($\pm$ SE) for the intercept and significant explanatory variables are presented at the cutoff where the percentage of cases and controls correctly identified (accuracy) is maximized, with the model sensitivity (Se) and specificity (Sp) at this cutoff.

Table 4. Final multivariate regression models for gram-negative clinical mastitis cases using 3 baselines (d −10, −7, and −3 relative to clinical mastitis detection)¹

Baseline	Model	Slope estimate (±SE)	P-value	Cases (no.)	Cutoff	Accuracy (%)	Se (%)	Sp (%)
d −10	Intercept	−2.2 ± 0.2	<0.01	32	0.16	82.1	40.6	87.5
	Activity, d −10 to −3	0.1 ± 0.0	0.01					
	Activity, d −10 to −6	−0.0 ± 0.0	0.01					
	Fat %, d −10 to −3	−3.2 ± 1.3	0.02					
	Lactose %, d −10 to −5	−9.8 ± 5.9	0.09					
d −7	Intercept	−2.0 ± 0.2	<0.01	33	0.14	71.2	63.6	72.3
	Fat %, d −7 to −2	4.9 ± 2.0	0.02					
	Lactose %, d −7 to −2	12.3 ± 6.4	0.05					
	Activity, d −7 to −5	0.0 ± 0.0	0.02					
d −3	Intercept	−1.6 ± 0.3	<0.01	21	0.24	73.3	71.4	73.8
	Lying bout duration, d −3 to 0	0.1 ± 0.0	0.08					
	Fat %, d −3 to −2	1.6 ± 0.7	0.02					
	Lactose %, d −3 to −2	3.5 ± 2.0	0.08					

¹For each model, the slope estimate (±SE) for the intercept and significant explanatory variables are presented at the cutoff where the percentage of cases and controls correctly identified (accuracy) is maximized, with the model sensitivity (Se) and specificity (Sp) at this cutoff.

GN model, Se declined for GP models as the baseline neared the CM diagnosis. The GP model was able to detect twice as many clinical cases at d −10 than the GN model at the same baseline (Se 81% vs. 41%), indicating that the d −10 baseline is more suitable for detecting true cases of CM caused by GP pathogens. This baseline represents a longer period of time series changes, which is sensible because GP infections are often of longer duration and may persist in subclinical form before clinical presentation of mastitis (Schukken et al., 2011).

The NPI model with the greatest case and noncase classification accuracy was the d −3 baseline, which included breed and changes in lying bout duration and

fat percentage between d −3 and −2 relative to CM diagnosis (Table 6). This model correctly identified 84% of cases and controls, with 62% Se and 88% Sp. The next best NPI model used d −10 as the baseline, which included breed, DIM, and changes in protein, lactose, fat, conductivity, activity, and number of lying bouts. This had the greatest Se (94%) of all NPI models combined with 77% Sp. The improved Se for detecting NPI cases at d −10 relative to diagnosis suggests some differences between these cases and GN cases because greater Se was achieved for the GN models at the baseline closer to clinical diagnosis. This finding supports our univariate findings of NPI cases being dissimilar to the other infection types.

Table 5. Final multivariate regression models for gram-positive clinical mastitis cases using 3 baselines (d −10, −7, and −3 relative to clinical mastitis detection)¹

Baseline	Model	Slope estimate (±SE)	P-value	Cases (no.)	Cutoff	Accuracy (%)	Se (%)	Sp (%)
d −10	Intercept	−3.9 ± 0.8	<0.01	11	0.18	86.1	81.8	86.6
	Breed	2.0 ± 0.8	0.02					
	Activity, d −10 to −4	−0.2 ± 0.6	0.02					
	Lying bouts, d −10 to −3	−1.9 ± 0.8	0.02					
	Lying bout duration, d −10 to −8	−0.1 ± 0.1	0.03					
	Conductivity, d −10 to −7	−7.7 ± 2.9	0.01					
	Conductivity, d −10 to −8	3.7 ± 1.8	0.04					
d −7	Intercept	−3.3 ± 0.6	<0.01	11	0.12	80.0	72.7	80.8
	Breed	1.9 ± 0.8	0.01					
	Lying time, d −7 to −5	0.0 ± 0.0	0.08					
	Lactose %, d −7 to −5	−14.0 ± 6.9	0.04					
d −3	Intercept	−2.9 ± 0.4	<0.01	19	0.19	85.5	47.4	89.4
	Breed	1.7 ± 0.6	<0.01					
	Activity, d −3 to 0	0.0 ± 0.0	0.04					
	Protein %, d −3 to −2	−4.0 ± 1.7	0.02					

¹For each model, the slope estimate (±SE) for the intercept and significant explanatory variables are presented at the cutoff where the percentage of cases and controls correctly identified (accuracy) is maximized, with the model sensitivity (Se) and specificity (Sp) at this cutoff.

Table 6. Final multivariate regression models for clinical mastitis cases with no pathogen isolated using 3 baselines (d -10, -7, and -3 relative to clinical mastitis detection)¹

Baseline	Model	Slope estimate ($\pm$ SE)	P-value	Cases (no.)	Cutoff	Accuracy (%)	Se (%)	Sp (%)
d -10	Intercept	-5.3 ± 1.3	<0.01	16	0.13	79.6	93.8	77.2
	DIM	0.0 ± 0.0	0.05					
	Breed	2.3 ± 0.8	0.01					
	Protein %, d -10 to -3	-73.3 ± 36.8	0.05					
	Protein %, d -10 to -4	73.6 ± 31.0	0.02					
	Lactose %, d -10 to -5	-69.3 ± 23.0	<0.01					
	Fat %, d -10 to -5	-9.5 ± 4.9	0.06					
	Conductivity, d -10 to -5	-7.7 ± 3.8	0.04					
	Activity, d -10 to -7	0.1 ± 0.0	0.09					
	Lactose %, d -10 to -8	10.8 ± 6.5	0.10					
	Lying bouts, d -10 to -7	-0.9 ± 0.4	0.03					
	Lying bouts, d -10 to -8	0.4 ± 0.2	0.08					
d -7	Intercept	-1.1 ± 0.2	<0.01	64 ²	0.24	59.1	76.6	53.4
	Breed	-2.3 ± 1.0	0.03					
	Activity, d -7 to -1	-0.1 ± 0.0	0.04					
	Milk yield, d -7 to -4	0.2 ± 0.1	0.08					
	Milk yield, d -7 to -5	-0.2 ± 0.1	0.02					
d -3	Intercept	-2.8 ± 0.5	<0.01	13	0.20	84.2	61.5	87.5
	Breed	1.2 ± 0.7	0.06					
	Lying bout duration, d -3 to -2	-0.0 ± 0.0	0.08					
	Fat %, d -3 to -2	-2.0 ± 0.9	0.03					

¹For each model, the slope estimate ($\pm$ SE) for the intercept and significant explanatory variables are presented at the cutoff where the percentage of cases and controls correctly identified (accuracy) is maximized, with the model sensitivity (Se) and specificity (Sp) at this cutoff.

²The number of cases is greater for the d -7 baseline because cases from both farms are included due to the lack of significance of activity parameters that were specific to the Virginia Tech Dairy Center (Blacksburg, VA) only.

Performance comparisons across the 4 response variables are best done by assessing the Se and Sp estimates rather than the accuracy of the model. Sensitivity and Sp are test characteristics and are independent of disease prevalence, whereas accuracy represents the correct classification rate of both cases and noncases achieved by a model. A greater number of noncases compared with cases will result in an accuracy more reflective of the model's ability to correctly identify noncases (Sp), but we are equally interested in model Se, or ability to correctly identify cases. The ratio of cases to noncases was balanced for the ACM model. However, for the pathogen type models, a smaller proportion of cases was available. Therefore, consideration of Se and Sp together is more appropriate when comparing models.

When considering both Se and Sp, the GP model at d -10 baseline was the best performing model for detecting GP cases of CM, with an Se of 82% and an Sp of 87%. The NPI model also had superior Se (94%) compared with the ACM model, whereas the Se and Sp of the best GN model (71 and 74%, respectively) were similar to those of the ACM model (73 and 75%, respectively). The test characteristics for our models are similar to those reported in the literature. A recent study that used milk component and electrical conductivity data produced a dynamic linear model with a Bayesian classifier that detected CM with 80% Se and

81% Sp without differentiating CM cases by pathogen type. When classifying a known CM case by gram status, Steeneveld et al. (2009) reported 73% accuracy, which increased to 95% when a probability threshold of >0.9 was implemented. Similarly, Kamphuis et al. (2011) reported 91% accuracy when classifying known CM cases as GP or GN based on the training data set, but accuracy declined to 55% when an independent test data set was used. Across 37 sensor systems developed for automated mastitis detection, the reported Se ranged from 55 to 89% and Sp ranged from 56 to 99% (Rutten et al., 2013). However, none of these systems or the models in the current study met the International Organization for Standardization/Final Draft International Standard 20966 requirements of 80% Se and 99% Sp (Hogeveen et al., 2010). An Sp below 99% will result in many false alerts, but some producers may accept lower levels of Sp to create attention lists as a screening tool to assist in mastitis detection (Kamphuis et al., 2008).

Data Challenges, Study Limitations, and Opportunities for Future Work

The current study offered the first steps in the process of building models to alert farmers to cows with mastitis caused by GN or GP bacteria. We aimed to

understand the important changes occurring before CM detection for these pathogen types and for NPI infections that can be observed using sensor data. Relevant sensor inputs and the best timeframe before CM detection were determined for each of the responses, allowing comparison between ACM and separated GN, GP, and NPI. However, our study was not without limitations and has provided opportunities for future work.

This study used linear models where slope changes in milk and activity variables were estimated by linear regression and input into logistic regression models to estimate the log probability of being a case. Studies with similar objectives have used a decision tree model (Kamphuis et al., 2011) or naïve Bayesian networks (Steenefeld et al., 2009), which demonstrated some success in classifying a known CM case as GN or GP. Another potential approach could have used multinomial regression and generated a single model with multiple outcomes rather than 4 separate models. However, our approach was selected over other techniques to (1) retain some biological understanding of the changes occurring before detection of CM, (2) include just those explanatory variables that were identified as important for a particular pathogen type, and (3) simultaneously detect CM and indicate pathogen type rather than classify a known CM case by gram status. We included nonsensor information such as DIM, breed, and parity in our multiple regression models. Breed was important for indicating GP, and both breed and DIM were important for NPI cases. We did not have previous CM history and SCC available for inclusion into analyses, but previous studies have demonstrated improvements in predictive performance of algorithms for detecting mastitis when this information is incorporated into models (Steenefeld et al., 2010; Jensen et al., 2016). Ongoing work by our research group investigates the use of more data-driven analyses to enhance CM detection.

The selection of a threshold for determining a case or noncase in our study was based on maximizing the proportion of correct predictions, giving an equal weighting to Se and Sp. This approach was taken because both Se and Sp are important when considering the success of a mastitis detection model. This study formed the initial step in building models for detecting gram-specific CM, and an equal weighting of correctly predicting cases and noncases was deemed appropriate for our current objective of understanding the changes occurring in the lead up to CM caused by GN, GP, or NPI compared with ACM cases. Alternatively, selecting a threshold using receiver operating characteristic curves is another common approach but was not used in this study. Furthermore, positive and negative predictive values can be computed to provide more information on the pre-

dictive abilities of a model and are considered useful in our ongoing work.

A limitation commonly reported in CM detection studies is the limited availability of CM cases to train and test models (Hogeveen et al., 2010; Kamphuis et al., 2011). Our analyses included data relating to 170 cases of CM and 166 controls, but for pathogen type models as few as 11 cases (e.g., GP models) were used to derive models. Originally, the data set contained 268 cases of CM and an equal number of controls (Tholen, 2012), but incomplete data in the period before CM detection meant that not all cases could be included. Moreover, behavioral data from the UF herd was limited to activity due to the more simplistic Afipedometer used on this farm, which meant that only VT CM cases could be used in models that included behavioral parameters other than step activity. Our study used 4 different sensor systems to measure 10 milk and activity parameters and BW. Missing data are common in sensor systems due to errors in automatic data reading or entry in the parlor. For simplicity and to represent real data collected using on-farm technologies, no data manipulation techniques were implemented in this study to make up for the missing data. However, a model designed for use on farm must have the ability to handle missing observations. One potential approach to deal with missing data is to generate the missing value by taking 95% of the previous day's value and 5% of the herd's average for the current day (Sørensen et al., 2016). Ongoing work by our research group involves some manipulation to reduce the effect of missing data.

Sensor systems such as the Afimilk technologies provide information at each milking, or more regularly for activity data, but we combined measurements into daily averages or sums for each variable to focus on the day-to-day variability in these variables in the lead up to CM. Improvements may be made in detection accuracy by using session data, similar to the approach used by Jensen et al. (2016), where morning and evening session data were modeled separately. Alternatively, instead of using averages and sums for each variable, variability within a day could be represented by including maximum and minimum values alongside means (e.g., Kamphuis et al., 2011). These alternative approaches have shown promise in CM detection by other researchers but were outside of the scope of the current study. The Afimilk meter and analyzer units provide cow composite measures of milk yield, conductivity, SCC category, and fat, protein, and lactose percentage. However, the effect of CM in one quarter is diluted by other, uninfected quarters. Quarter-based measurements could improve predictions by detecting more pronounced changes at the quarter level (e.g., Mollenhorst et al., 2010). This approach is commonly used in automated

milking systems but cannot be applied in our analysis, which used data from conventional farms. Further, our findings are limited to farms that have adopted technologies for measuring cow performance and behavior.

In this study, model performance was evaluated internally using the model training data set, and we recognize the potential bias of this approach. Given the small number of cases in this data set, it was not feasible to split the data into a training and validation data set. An independent evaluation of multiple regression models derived in this study is underway to investigate the predictive ability of these models using an independent test data set and to provide a more in-depth assessment of success rates of gram-specific models. The intended use of these models in the future is as an on-farm screening tool for early indication of infections caused by these different pathogen types, which, if accurate, could support decisions around appropriate intervention.

CONCLUSIONS

Individual activity or milk parameters were important for the categorized infection types (GN, GP, and NPI), but no significant parameters were revealed in the univariate analyses for ACM. Using a combination of milk and activity measures in multiple regression models, we demonstrated better model performance for GP and NPI categorized CM cases compared with ACM, although GN model performance was similar to ACM model performance. Sensitivity was greater for detecting GN when the baseline was closer to the day of diagnosis ($d - 3$), whereas Se was greatest for detecting GP and NPI when data up to 10 d before clinical diagnosis were included in models. Further evaluation of these models using independent data is necessary to support our findings. Based on our analyses, changes in cow performance and behavior that precede a CM case differ for GN, GP, and NPI cases. On-farm sensor data may therefore be useful as an additional screening tool for clinically mastitic cows.

ACKNOWLEDGMENTS

We acknowledge the farm staff at the Virginia Tech Dairy Center (Blacksburg, VA) and the University of Florida Dairy Unit (Gainesville, FL) for assistance with data collection and Afimilk Ltd. (Kibbutz Afikim, Israel) for data extraction. The funder of this research, Virginia Agricultural Council (Richmond), is gratefully acknowledged. Nicole Steele is partially supported by New Zealand dairy farmers through DairyNZ Inc. (Hamilton, New Zealand). The authors have not stated any conflicts of interest.

REFERENCES

- Auldust, M., S. Coats, G. Rogers, and G. McDowell. 1995. Changes in the composition of milk from healthy and mastitic dairy cows during the lactation cycle. *Aust. J. Exp. Agric.* 35:427–436. <https://doi.org/10.1071/EA9950427>.
- Borchers, M. R., Y. Chang, I. Tsai, B. Wadsworth, and J. Bewley. 2016. A validation of technologies monitoring dairy cow feeding, ruminating, and lying behaviors. *J. Dairy Sci.* 99:7458–7466. <https://doi.org/10.3168/jds.2015-10843>.
- Edwards, J. L., and P. R. Tozer. 2004. Using activity and milk yield as predictors of fresh cow disorders. *J. Dairy Sci.* 87:524–531. [https://doi.org/10.3168/jds.S0022-0302\(04\)73192-6](https://doi.org/10.3168/jds.S0022-0302(04)73192-6).
- Fogsgaard, K. K., T. W. Bennedsgaard, and M. S. Herskin. 2015. Behavioral changes in freestall-housed dairy cows with naturally occurring clinical mastitis. *J. Dairy Sci.* 98:1730–1738. <https://doi.org/10.3168/jds.2014-8347>.
- Forsbäck, L., H. Lindmark-Mansson, A. Andren, M. Akerstedt, L. Andree, and K. Svennersten-Sjaunja. 2010. Day-to-day variation in milk yield and milk composition at the udder-quarter level. *J. Dairy Sci.* 93:3569–3577. <https://doi.org/10.3168/jds.2009-3015>.
- Higginson, J. H., S. T. Millman, K. E. Leslie, and D. F. Kelton. 2010. Validation of a new pedometer system for use in behavioral research and lameness detection in dairy cattle. Pages 2–5 in *Proc. 1st North Am. Conf. Precision Dairy Management*, Toronto, ON, Canada. Progressive Dairy Operators, Elora, ON, Canada.
- Hogeveen, H., C. Kamphuis, W. Steeneveld, and H. Mollenhorst. 2010. Sensors and clinical mastitis—The quest for the perfect alert. *Sensors (Basel)* 10:7991–8009. <https://doi.org/10.3390/s100907991>.
- Jensen, D. B., H. Hogeveen, and A. De Vries. 2016. Bayesian integration of sensor information and a multivariate dynamic linear model for prediction of dairy cow mastitis. *J. Dairy Sci.* 99:7344–7361. <https://doi.org/10.3168/jds.2015-10060>.
- Kamphuis, C., H. Mollenhorst, and H. Hogeveen. 2011. Sensor measurements revealed: Predicting the Gram-status of clinical mastitis causal pathogens. *Comput. Electron. Agric.* 77:86–94. <https://doi.org/10.1016/j.compag.2011.03.012>.
- Kamphuis, C., D. Pietersma, R. Van der Tol, M. Wiedemann, and H. Hogeveen. 2008. Using sensor data patterns from an automatic milking system to develop predictive variables for classifying clinical mastitis and abnormal milk. *Comput. Electron. Agric.* 62:169–181. <https://doi.org/10.1016/j.compag.2007.12.009>.
- Kaniyamattam, K., and A. De Vries. 2014. Agreement between milk fat, protein, and lactose observations collected from the Dairy Herd Improvement Association (DHIA) and a real-time milk analyzer. *J. Dairy Sci.* 97:2896–2908. <https://doi.org/10.3168/jds.2013-7690>.
- Kiddy, C. A. 1977. Variation in physical activity as an indication of estrus in dairy cows. *J. Dairy Sci.* 60:235–243. [https://doi.org/10.3168/jds.S0022-0302\(77\)83859-9](https://doi.org/10.3168/jds.S0022-0302(77)83859-9).
- King, M. T., S. LeBlanc, E. Pajor, T. Wright, and T. DeVries. 2018. Behavior and productivity of cows milked in automated systems before diagnosis of health disorders in early lactation. *J. Dairy Sci.* 101:4343–4356. <https://doi.org/10.3168/jds.2017-13686>.
- Kitchen, B. J. 1981. Bovine mastitis: Milk compositional changes and related diagnostic tests. *J. Dairy Res.* 48:167–188. <https://doi.org/10.1017/S0022029900021580>.
- Lago, A., S. Godden, R. Bey, P. Ruegg, and K. Leslie. 2011. The selective treatment of clinical mastitis based on on-farm culture results: I. Effects on antibiotic use, milk withholding time, and short-term clinical and bacteriological outcomes. *J. Dairy Sci.* 94:4441–4456. <https://doi.org/10.3168/jds.2010-4046>.
- Makovec, J. A., and P. Ruegg. 2003. Results of milk samples submitted for microbiological examination in Wisconsin from 1994 to 2001. *J. Dairy Sci.* 86:3466–3472. [https://doi.org/10.3168/jds.S0022-0302\(03\)73951-4](https://doi.org/10.3168/jds.S0022-0302(03)73951-4).
- Middleton, J., L. Fox, G. Pighetti, and C. Petersson-Wolfe. 2017. *Laboratory Handbook on Bovine Mastitis*. National Mastitis Council, New Prague, MN.
- Milner, P., K. Page, and J. Hillerton. 1997. The effects of early antibiotic treatment following diagnosis of mastitis detected by a change

- in the electrical conductivity of milk. *J. Dairy Sci.* 80:859–863. [https://doi.org/10.3168/jds.S0022-0302\(97\)76008-9](https://doi.org/10.3168/jds.S0022-0302(97)76008-9).
- Mollenhorst, H., P. P. J. van der Tol, and H. Hogeveen. 2010. Somatic cell count assessment at the quarter or cow milking level. *J. Dairy Sci.* 93:3358–3364. <https://doi.org/10.3168/jds.2009-2842>.
- Pyörälä, S. 2003. Indicators of inflammation in the diagnosis of mastitis. *Vet. Res.* 34:565–578. <https://doi.org/10.1051/vetres:2003026>.
- Pyörälä, S., L. Kaartinen, H. Käck, and V. Rainio. 1994. Efficacy of two therapy regimens for treatment of experimentally induced *Escherichia coli* mastitis in cows. *J. Dairy Sci.* 77:453–461. [https://doi.org/10.3168/jds.S0022-0302\(94\)76973-3](https://doi.org/10.3168/jds.S0022-0302(94)76973-3).
- Quist, M. A., S. LeBlanc, K. Hand, D. Lazenby, F. Miglior, and D. Kelton. 2008. Milking-to-milking variability for milk yield, fat and protein percentage, and somatic cell count. *J. Dairy Sci.* 91:3412–3423. <https://doi.org/10.3168/jds.2007-0184>.
- R Core Team. 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Roberson, J. R. 2003. Establishing treatment protocols for clinical mastitis. *Vet. Clin. North Am. Food Anim. Pract.* 19:223–234. [https://doi.org/10.1016/S0749-0720\(02\)00071-3](https://doi.org/10.1016/S0749-0720(02)00071-3).
- Roberson, J. R. 2012. Treatment of clinical mastitis. *Vet. Clin. North Am. Food Anim. Pract.* 28:271–288. <https://doi.org/10.1016/j.cvfa.2012.03.011>.
- Roman-Garcia, Y., R. R. White, and J. L. Firkins. 2016. Meta-analysis of postruminal microbial nitrogen flows in dairy cattle. I. Derivation of equations. *J. Dairy Sci.* 99:7918–7931. <https://doi.org/10.3168/jds.2015-10661>.
- Royster, E., S. Godden, D. Goulart, A. Dahlke, P. Rapnicki, and J. Timmerman. 2014. Evaluation of the Minnesota Easy Culture System II Bi-Plate and Tri-Plate for identification of common mastitis pathogens in milk. *J. Dairy Sci.* 97:3648–3659. <https://doi.org/10.3168/jds.2013-7748>.
- Rutten, C. J., A. Velthuis, W. Steeneveld, and H. Hogeveen. 2013. Invited review: Sensors to support health management on dairy farms. *J. Dairy Sci.* 96:1928–1952. <https://doi.org/10.3168/jds.2012-6107>.
- Schukken, Y. H., J. Günther, J. Fitzpatrick, M. Fontaine, L. Goetze, O. Holst, J. Leigh, W. Petzl, H.-J. Schuberth, A. Sipka, D. G. E. Smith, R. Quesnell, J. Watts, R. Yancey, H. Zerbe, A. Gurjar, R. N. Zadoks, and H.-M. Seyfert. 2011. Host-response patterns of intramammary infections in dairy cows. *Vet. Immunol. Immunopathol.* 144:270–289. <https://doi.org/10.1016/j.vetimm.2011.08.022>.
- Sørensen, L. P., M. Bjerring, and P. Lovendahl. 2016. Monitoring individual cow udder health in automated milking systems using online somatic cell counts. *J. Dairy Sci.* 99:608–620. <https://doi.org/10.3168/jds.2014-8823>.
- Stangaferro, M. L., R. Wijma, L. S. Caixeta, M. A. Al-Abri, and J. O. Giordano. 2016. Use of rumination and activity monitoring for the identification of dairy cows with health disorders: Part II. Mastitis. *J. Dairy Sci.* 99:7411–7421. <https://doi.org/10.3168/jds.2016-10908>.
- Steeneveld, W., L. C. van der Gaag, H. W. Barkema, and H. Hogeveen. 2009. Providing probability distributions for the causal pathogen of clinical mastitis using naive Bayesian networks. *J. Dairy Sci.* 92:2598–2609. <https://doi.org/10.3168/jds.2008-1694>.
- Steeneveld, W., L. C. van der Gaag, W. Ouweltjes, H. Mollenhorst, and H. Hogeveen. 2010. Discriminating between true-positive and false-positive clinical mastitis alerts from automatic milking systems. *J. Dairy Sci.* 93:2559–2568. <https://doi.org/10.3168/jds.2009-3020>.
- Taponen, S., L. Salmikivi, H. Simojoki, M. T. Koskinen, and S. Pyörälä. 2009. Real-time polymerase chain reaction-based identification of bacteria in milk samples from bovine clinical mastitis with no growth in conventional culturing. *J. Dairy Sci.* 92:2610–2617. <https://doi.org/10.3168/jds.2008-1729>.
- Tholen, A. 2012. The use of animal activity data and milk components as indicators of clinical mastitis. MS Thesis. Virginia Polytechnic Institute and State University, Blacksburg, VA.
- Watts, J. L. 1988. Etiological agents of bovine mastitis. *Vet. Microbiol.* 16:41–66. [https://doi.org/10.1016/0378-1135\(88\)90126-5](https://doi.org/10.1016/0378-1135(88)90126-5).
- White, R. R., Y. Roman-Garcia, and J. L. Firkins. 2016. Meta-analysis of postruminal microbial nitrogen flows in dairy cattle. II. Approaches to and implications of more mechanistic prediction. *J. Dairy Sci.* 99:7932–7944. <https://doi.org/10.3168/jds.2015-10662>.
- Yeiser, E. E. 2011. The use of activity measures in combination with physiological factors as indicators of disease in dairy cattle. MS Thesis. Virginia Polytechnic Institute and State University, Blacksburg, VA.

ORCIDS

- N. M. Steele <https://orcid.org/0000-0001-9915-0954>
 A. De Vries <https://orcid.org/0000-0003-4511-0388>
 S. J. Lacy-Hulbert <https://orcid.org/0000-0002-9847-1294>
 D. Liebe <https://orcid.org/0000-0003-4447-4120>
 C. S. Petersson-Wolfe <https://orcid.org/0000-0002-2766-1306>