Animal 15 (2021) 100205

Contents lists available at ScienceDirect

Animal
The international journal of animal biosciences

Evaluation of infrared thermography combined with behavioral biometrics for estrus detection in naturally cycling dairy cows

H.J. Perez Marquez [a,*], D.J. Ambrose [a,b], A.L. Schaefer [a,c], N.J. Cook [a,d], C.J. Bench [a]

[a] Department of Agricultural, Food, and Nutritional Science, University of Alberta, Edmonton, AB T6G 2P5, Canada
[b] Livestock Research and Extension Branch, Alberta Agriculture and Forestry, Edmonton, AB T6H 5T6, Canada
[c] Animal Inframetrics Inc., Box 5451, Lacombe, AB T4L 1X2, Canada
[d] Livestock Research and Extension Branch, Alberta Agriculture and Forestry, Lacombe Research Centre, 6000 C & E Trail, Lacombe, AB T4L 1W1, Canada

ARTICLE INFO

Article history:
Received 17 August 2020
Revised 5 February 2021
Accepted 9 February 2021
Available online 23 June 2021

Keywords:
Combined-parameters
First-lactating
Movement-frequency
Preovulation
Skintemperature

ABSTRACT

Low estrus detection rates (>50%) are associated to extended calving intervals, low economic profit and reduced longevity in Holstein dairy cows. The objective of this study was to evaluate the accuracy of infrared thermography and behavioral biometrics combined as potential estrus alerts in naturally (not induced) cycling dairy cows housed in a tie-stall barn. Eighteen first lactation cows were subjected to transrectal ultrasonography to determine spontaneous ovulation. The dominant follicle (**DF**) disappearance was used retrospectively as an indirect indicator of ovulation, and to establish the estrus period (48–24 h prior the DF disappearance). Raw skin temperature (**Raw IR**) and residual skin temperature (**Res IR**) were recorded using an infrared camera at the Vulva area with the tail (**Vtail**), Vulva area without the tail (**Vnotail**), and Vulva's external lips (**Vlips**) at AM and PM milking from Day 14 until two days after ovulation was confirmed. Behavioral biometrics were recorded on the same schedule as infrared scan. Behavioral biometrics included large hip movements (**L-hip**), small hip movements (**S-hip**), large tail movements and small tail movements to compare behavioral changes between estrus and nonestrus periods. Significant increases in Raw IR skin temperature were observed two days prior to ovulation (Vtail; 35.93 ± 0.27 °C, Vnotail; 35.59 ± 0.27 °C, and Vlips; 35.35 ± 0.27 °C) compared to d −5 (Proestrus; Vtail; 35.29 ± 0.27 °C, Vnotail; 34.93 ± 0.31 °C, and Vlips; 34.68 ± 0.27 °C). No significant changes were found for behavioral parameters with the exception of S-hip movements, which increased at two days before ovulation (d −2; 11.13 ± 1.44 Events/5min) compared to d −5 (7.30 ± 1.02 Events/5min). To evaluate the accuracy of thermal and behavioral biometrics, receiver operating characteristic curve analysis was performed using Youden index (**YJ**), diagnostic odds ratio, positive likelihood ratio (**LR+**), Sensitivity, Specificity and Positive predicted value to score the estrus alerts. The greatest accuracy achieved using thermal parameters was for Res IR Vtail PM (YJ = 0.34) and L-hip PM (YJ = 0.27) for behavioral biometrics. Combining thermal and behavioral parameters did not improve the YJ index score but reduced the false-positive occurrence observed by increasing the diagnostic odds ratio (26.62), LR+ (12.47), Specificity (0.97) and positive predicted value (0.90) in a Res IR Vtail PM, S-hip AM, S-hip PM combination. The combination of thermal and behavioral parameters increased the accuracy of estrus detection compared to either thermal or behavioral biometrics, independently in naturally cycling cows during milking.

Implications

The use of infrared thermography is prompt to optimize the identification of estrus occurrence in dairy cows in a noninvasive way without additional labor input. The combination of behavioral and thermal data could be combined in one platform and create estrus alerts more accurately than when used in isolation. Additionally, sophisticated machine-learning methodologies could monitor thermal and behavioral data at the individual level (e.g. individual cows) that can increase the estrus detection rates and identify the optimum time to inseminate dairy cows. Automate infrared platforms can be used on different housing systems (e.g. tie-stalls, parlor, and robotic milking systems) and different herd

* Corresponding author at: Dairy Research and Technology Centre, F-30 Edmonton Research Station, South Campus, University of Alberta, Edmonton, AB T6H 2V5, Canada.
E-mail address: perezmar@ualberta.ca (H.J. Perez Marquez).

https://doi.org/10.1016/j.animal.2021.100205
1751-7311/

sizes where data can be collected consistently to generate real-time estrus alerts.

Introduction

The incidence of falsenegatives and falsepositives in many estrus detection methods contributes to extended artificial insemination intervals, calving delays, poor economic outcomes and decreased longevity in dairy cows (Mayo, 2015; Giordano et al., 2015). The visual observation of cows standingtobemounted is the most reliable estrus detection method and one of the most commonly used (>80%; Denis-Robichaud et al., 2016) due to its low incidence of false-positive estrus detection (Glencross et al., 1981; Sprecher et al., 1995) in North American herds (USDA, 2007). However, estrus detection rates based on visual observation have a low incidence of true-positive estrus alerts per cows in estrus (37–54%; Van Eerdenburg et al., 1996; Sakaguchi, 2010).

Physiologic reasons for reduced estrus detection rates include the reduction of estrus period duration in Holstein cows (from 18 to less than 8 h) over the last 50 year (Reames et al., 2011), negative correlations with higher milk yield (Lopez et al., 2005), reduced estrus behavior (e.g. restlessness and mounting behavior) in extreme ambient temperature (e.g. hot and cold temperatures; Collier et al., 2006) negative energy balance (Grummer and Rastani, 2003) and the differences in estrus behavior (e.g. frequency and duration) between multiparous and primiparous cows (López-Gatius et al., 2005; Chanvallon et al., 2014). The ability to detect estrus using visual observation of cows standing to be mounted is also significantly limited or nonexistent in tie-stall housing environments (Felton et al., 2012) compared with free-stalls or pasture-based herds as a result of cows being tethered while in their stall. Further, 61 percent of dairy herds in Canada are housed in tie-stall barns (Denis-Robichaud et al., 2016).

Good reproductive management relies on accurate monitoring and detection of estrus cues, which are used as indicators of when to inseminate a dairy cow. Research shows that the most cost-efficient time to AI is from 60 to 70 days in milk for multiparous and approximately 105 days in milk for primiparous cows (De Vries, 2006) in order to maintain an optimal calving interval (12–13 months; Stevenson et al., 2014). Extended calving interval leads to an increase of $1.00 USD (primiparous) and $1.80 USD (multiparous) in costs for every extra day a cow remains nonpregnant. Further, these costs increase to $6.00 USD if a cow is open during late lactation ($\geq$160 days in milk; Meadows et al., 2005).

Advances in estrus detection rates (>50%) have been achieved through the use of automated estrus detection devices which continuously monitor physiological and behavioral parameters to detect estrus without additional labor input (Rutten et al., 2014). Automated estrus detection consists of sensors and algorithms that create estrus alerts for proper artificial insemination service. Automated estrus detection devices can be divided into activity monitors (e.g. rumination time, laying bouts, walking, ear movements; Løvendahl and Chagunda, 2010; Aungier et al., 2012), mounting detectors (e.g. mounting counts and mounting duration; Xu et al., 1998; Sauls et al., 2017), body core temperature loggers (e.g. reticulo-rumen, vagina, ear and milk temperature; Fordham et al., 1988; Fisher et al., 2008), and analysis of progesterone concentrations in milk (Delwiche et al., 2001; Adriaens et al., 2017). However, most automated estrus detection devices are designed for application in free-stall situations and often fail to detect estrus or require additional handling such as moving cows to an outside pen if reared in tie-stall housing. Most dairy producers have adopted the combined use of various estrus detection methods, usually estrus detection devices with visual observations. However, no detailed analyses have been performed to describe the effect on accuracy by combining different estrus detection methods (Firk et al., 2002).

Live organisms emit electromagnetic radiation (thermal radiation; Boyd, 1983) some which can be measured using infrared thermography cameras. This energy can be emitted, reflected or transmitted. In particular, animals and humans are susceptible to heat loss (e.g. conduction, convection, radiation, and evaporation) in the environment (Berz, 2007). Changes (e.g. increases or decreases) in the amount of heat loss can indicate different physiological processes. For example, infrared has been used in dairy cattle to measure skin temperature changes to monitor udder health status (Sathiyabarathi et al., 2016), heat stress (Daltro et al., 2017), qualitative differences in cattle lameness (Novotna et al., 2019), and early lactation diseases (e.g. ketosis, metritis, and milk fever; Macmillan et al., 2019). Several studies report increased skin temperature at the vulva associated with the estrus period, which can serve as an estrus alert (Osawa et al., 2004; Talukder et al., 2014; Perez Marquez et al., 2019). Further, estrus detection using infrared cameras have been used to detect estrus and ovulation regardless of housing type in multiparous cows (71% in free-stalls; Talukder et al., 2014, $\geq$50% in tie-stalls; Perez Marquez et al., 2019). Nevertheless, debris present on the animal can potentially influence thermal radiation by masking actual thermal readings (mixed results; Sykes et al., 2012). In addition, it is difficult to standardize the conditions for the use of handheld infrared cameras due to variations in the angle and distance between the camera and target, which can also affect thermal readings (Talukder et al., 2014). Thus, whenever infrared is used, careful consideration must be given to ensure all debris is cleaned away, and conditions are kept standardized.

Biomechanical movements have also been reported as useful biometric parameters to identify different physiological processes in humans (Jain et al., 2004). Similarly, in tie-stall housed dairy cows, changes in restless behavior as measured using <20 mm hip movements (e.g. back – forward and left – right) prior to ovulation have also been demonstrated using 3D-kinematics (Guesgen and Bench, 2018). Infrared technology in beef cattle has also been able to measure an individual animal's behavioral frequencies using an automated RFID-IR platform (Cook et al., 2016). The above research demonstrated that behavioral frequencies could be measured by analyzing changes in thermal distribution within a thermal image by comparing the thermal radiation from a target with a colder background. Based on the above findings, the objective of the present study was to evaluate a combination of thermal and behavioral biometrics as estrus alerts at the estimated estrus period (48–24 h prior ovulation) in naturally cycling dairy cows in a tie-stall housing. We hypothesized that behavioral biometrics using the hip and tail regions combined with infrared metrics from the vulva area would increase the accuracy compared to these same parameters utilized in isolation as indicators of the estimated estrus period.

Material and methods

The current study was conducted from June to October 2016 (summer-fall) at the University of Alberta's, Dairy Research and Technology Centre, a 146-cow tie-stall facility located at Edmonton, Alberta, Canada**.** The study evaluated 18 naturally cycling (not induced by hormone interventions) primiparous Holstein cows following a hypothesis testing: two-sample inference estimation of sample size and power using two means (Rollin, 2016) with an $\alpha = 0.50$ and a power = 0.90. The minimum required number of cow was seven. However, in anticipation of excluding some cows due to abnormal estrous cycles and postpartum disease, 18 cows were assigned to the study. Cows were averaging 43 ± 2 days in

milk (±SD) and producing 27.3 ± 5.63 kg (Mean ± SD) of milk per day at the beginning of the study. During the study period, cows were housed indoors for 31 ± 6 d continuously with no access to an outside pen to avoid any variations in infrared measurements associated with exposure to the outside environment. Cows were milked twice daily (0330–0600 and 1500–1730) in-stall using a pipeline milking system. Free access was given to water and a total mixed ration based on NRC guidelines (National Research Council, 2001) for lactating dairy cows. The main ingredients of the total mixed ration were alfalfa-barley silage, rolled barley-corn, grass hay, and mineral supplements.

Experimental design

The current study followed a split-plot over time experimental design that compared thermal and behavioral biometrics during the proestrus stage (baseline), the expected estrus period, the day of ovulation and two days postovulation for all eighteen cows ($n = 18$). Each cow served as the experimental unit. Cows were assigned to the study if the presence of a *corpus luteum* was confirmed by transrectal ultrasonography (ALOKA SSD-500 scanner fitted with a 7.5 MHz linear array transducer, ALOKA Co., LTD, Tokyo, Japan) by the same technician throughout the study. Ovarian mapping was conducted every other day until *corpus luteum* regression was evident followed by the disappearance of a dominant follicle (**DF**) which indicated the occurrence of ovulation. Once each cow ovulated (Day 0) and the presence of a new *corpus luteum* was confirmed subsequently, ultrasound scanning was resumed every other day (1700) from d 7 to d 13 and daily scans from d 14 until confirmation of subsequent ovulation (d 0) and the appearance of new *corpus luteum* (Fig. 1). Dominant follicles and *corpus luteum* diameters were measured in mm using built-in callipers and recorded for left and right ovaries to determine follicular growth, monitor the presence of DF and *corpus luteum* regression.

Milk sampling and estradiol assay

Milk samples were obtained directly by teat stripping, discarding the first two strips during both milking times (AM and PM) following the same schedule of data collection from thermal and behavioral biometrics (Fig. 1). Milk samples (10 mL) were collected from cows at each milking into 35 mL snap-seal containers (Fisher Scientific Company, Ottawa, ON, Canada). Samples then were transferred to 10 mL plain Vacutainer® tubes (Becton, Dickinson and Company, Franklin Lakes, NJ, USA), centrifuged at 1 940 *g*, at 4 °C for 20 min to remove milk fat, and skim milk samples were stored in two 5 ml plastic tubes (MCT Fisher Scientific, Waltham, MA, USA) at −20 °C until estradiol assays were performed.

Skim milk samples (100-μL) were analyzed using an estradiol ELISA kit (IBL America, MN, USA) in a single assay with duplicate analysis. Grgurevic et al. (2016) and Snoj et al. (2017) previously validated a direct bovine milk sample estradiol analysis in a plasma serum ELISA kit. The estradiol assay kit had a standard range of 3–200 pg/mL with a sensitivity of <1.399 pg/mL and cross-reactivity with the structurally related compounds of estrone (0.2%), estriol (0.05%) and fulvestrant (0.3%). The range in estradiol concentration was 6.85–47.82 pg/mL with an inter-assay coefficient of variation of 9.33% at 65.64 pg/mL and an intra-assay coefficient of variation of 5.57% at 16.15 pg/mL. Estradiol daily means were calculated (AM + PM samples/2) to match ovarian structure data (e.g. *corpus luteum* and DF). However, estradiol concentration peaks were found individually (e.g. each cow) if the estradiol concentration per Sample day were greater than two SD plus the mean.

Infrared thermography

Thermal images were captured at four frames/s for a total of 5 min from a distance of 1 m perpendicular to the caudal-dorsal side of the cows during morning (AM) and afternoon milking (PM). To compare the expected estrus period (48–24 h before ovulation) with the proestrus stage (baseline), thermal images were recorded starting on d 14 of the estrous cycle until two days after ovulation. An A310 thermal camera (FLIR Vision Systems Ltd. Burlington, ON, Canada) was used to record infrared images of 320 × 240 pixels. The thermal sensitivity of the camera was 0.05 °C at 30 °C, with a measurement range of −20 °C to 120 °C, and accuracy of ±2 °C or 2% of the measured temperature. Thermal images were collected using Vacca2 software (Animal Inframetrics Inc. Lacombe, AB, Canada). A laptop computer (ThinkPad, Lenovo Group Limited, Haidian District, Beijing, China), with Vacca2 software and an infrared camera powered by a 12-volt battery with a 1 000-Watt inverter (MotoMaster, Canadian Tire Co. Toronto, ON, Canada) were placed on a wheeled cart that could easily be moved from stall to stall during data collection (Fig. 2). A GLM15 50ft laser measurement tool (Robert Bosch Tool Co. IL, USA.) was used to ensure a consistent distance between the camera and cows. Ambient temperature (°C) and percent relative air humidity (Rh%) were recorded using a hygrometer (Kestrel Nielsen-Kellerman Co. MN, USA.). Emissivity was set to 0.98 following manufacturers recommendations for live tissues (e.g. cow's skin surface).

Thermal images were processed using FLIR ResearchIR software (FLIR Systems Ltd Burlington, ON. Canada) to determine maximum, minimum, and average (SD) skin temperature output of each area of interest from a selected frame per each cow sample collection. Areas in the thermal images that defined the Vulva area with the tail (**Vtail**), Vulva area without the tail (**Vnotail**), and Vulva's external lips (**Vlips**) were predetermined using a standardized ellipse (Vtail and Vnotail) and a free-drawing tool (Vlips) in the FLIR ResearchIR software and was standardized for all images (Fig. 3).

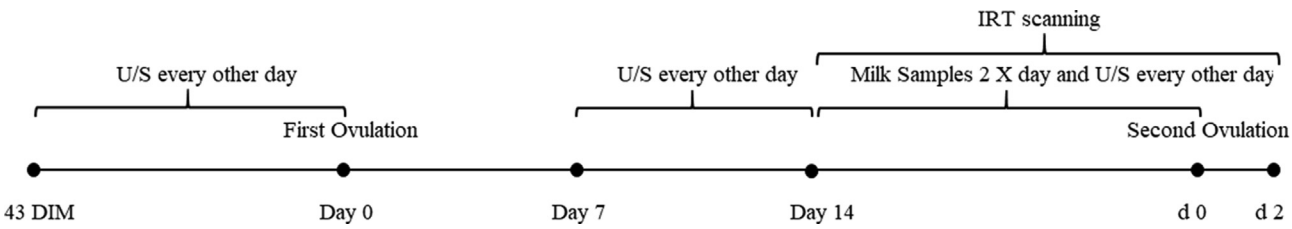

Fig. 1. Experimental timeline. Transrectal ultrasonography (U/S) in cows was performed every other day from 43 ± 2 days in milk (DIM) until the first ovulation (Day 0), which resumed every other day from Day 7 to Day 13 to monitor ovarian dynamics. From Day 14 until two days after the second ovulation (d 0), U/S was performed once daily. Simultaneously, infrared thermography (IRT) was performed and thermal frames were recorded during morning and afternoon (AM-PM) milking to record the maximum skin temperature and the frequency of event for hip and tail behaviors (Events/5min). Additionally, milk samples were collected from Day 14 until d 0 to determine peripheral estradiol concentrations.

Fig. 2. Infrared thermography cart. (A) A310 thermal camera protected in a camera case with a perpendicular angle facing the vulva area with 1 m. (B) Laptop with Vacca2 frame puller software (Animal Inframetrics) connected to the thermal camera via Ethernet cable. (C) Power source (12 volts battery with a 1 000-watt power inverter). (D) Primiparous dairy cow at her stall.

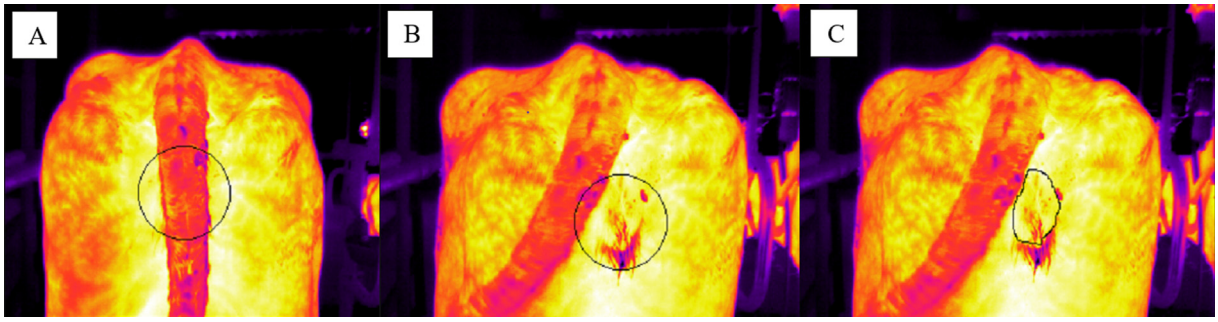

Fig. 3. Vulva area with the cow tail (Vtail; A), Vulva area with vulva exposed (Vnotail; B) and Vulva's external lips (Vlips; C). The ellipses (A and B), and hand draw area (C) were consistently used to record the same number of pixels through all the thermal images to identify the maximum skin temperature for each image.

Behavioral observations

To determine the frequency of behavioral events as ovulation approached, behavioral biometrics were scored using thermal frames at each milking number of Events/5 min at four frames/s. The same person performed behavioral observations to eliminate inter-observer variation. Behavioral frequencies were determined during the same period as infrared scanning (d 14 until two days after ovulation). Behavioral biometric data were categorized into large and small movements from hip and tail. Hip smallmovements (**S-hip**) were defined as any movements side to side (e.g. left–right) within 10 cm from the rest position (standing still), and hip largemovements (**L-hip**) were defined as any movement beyond 10 cm from the rest position. Tail events were categorized as a smalltail movement when the tail movement was within the rear thermal area of the cow, and largetail movement as any tail movement outside the cow's thermal area of each cow (Fig. 4).

Statistical analysis

Behavioral and thermal biometrics were analyzed using SAS software (SAS ver 9.4, Cary, NC, USA). Sample days were standardized (d −5, d −4, d −3, d −2, d −1, d 0, d 1 and d 2) using ovulation as d 0 to compare baseline with pre- (d −1 to d −5) and postovulation (d 1 to d 2) days. Proc Univariate was used to test normality assumptions using a Kolmogorov-Smirnov test ($P > 0.05$). All thermal data complied with the normality assumptions, however, behavioral data did not satisfy normality assumptions and were found to have a Poisson distribution. Models were fitted using a Generalized Linear Mixed Model approach (Proc Glimmix). A Bonferroni separation test was used to present results in Least-Square means, SEM and the statement of ar(1) to account for the lack of independent and homogeneity in the data. Results were considered significant if $P < 0.05$, tendency if $P \geq 0.05$ and <0.10 and P values ≥0.10 were considered not significant. Nonsignificant fixed

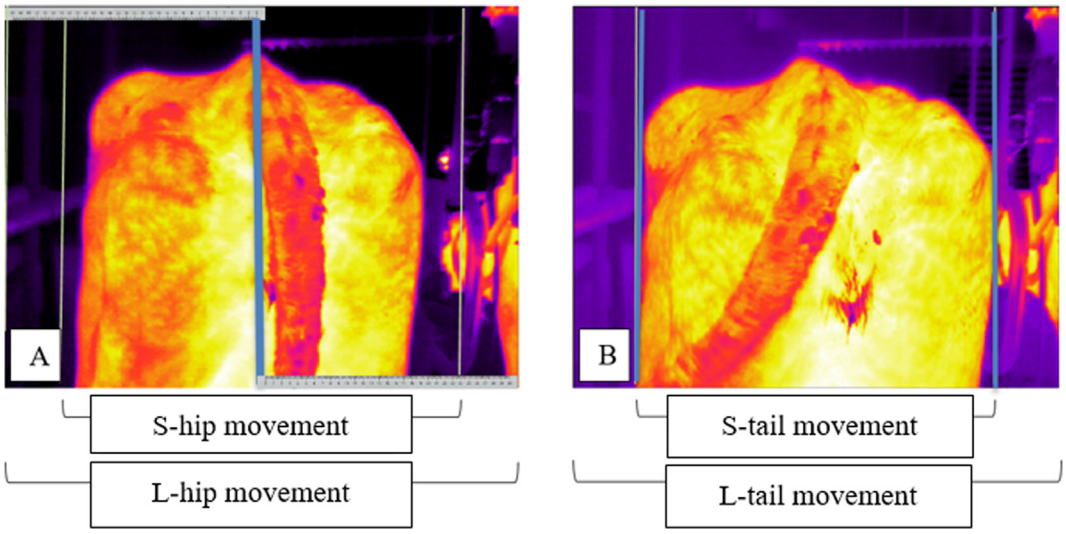

Fig. 4. Hip movement frequency (A) was divided into S-hip (any event within 10 cm side to side using tail head as a middle reference point), and L-hip (all events beyond 10 cm side to side using tail head as a middle reference point). Tail movements (B) were similarly divided into S-tail (tail events inside the thermal shape of the cow within the frame), and L-tail (tail events outside the thermal shape of the cow).

variables were eliminated from subsequent statistical models. Pearson correlation coefficient analysis (Proc Corr) was performed to identify possible associations between estradiol, thermal, and behavioral biometrics on days approaching ovulation.

Infrared thermography analysis

Maximum skin temperature was used in all data analyses to eliminate sources of variation on the surface of the vulva (e.g. min and average temperature). To compensate for the effect of environmental factors on skin temperature, residual skin temperature (**Res IR**) was calculated following Cook et al. (2016) methodology by subtracting the predicted skin temperature from the observed skin temperature (**Raw IR**). To the Raw IR for each cow from Vlips, Vtail, and Vnotail as well Res, IR dependent variables (Vlips, Vtail and Vnotail) were examined using Kolmogorov-Smirnov test (P-value > 0.05). Thermal data were found to follow normality assumptions, and no outliers were identified. Fixed variable was Sample day relative to ovulation (d −5, d −4, d −3, d −2, d −1, d 0, d 1, and d 2) and the model was tested using a Type 3 test with the inverse (ilink) function specified and the statement of ar (1) to account for the lack of independent and homogeneous data.

Behavioral data analysis

To analyze the frequency of behaviors, the fixed variables were Sample day relative to ovulation (d −5, d −4, d −3, d −2, d −1, d 0, d 1, and d 2) and Sample time (AM milking, and PM milking) while Cow was identified as a random statement with a Poisson distribution specified.

Accuracy evaluation

To evaluate the performance of thermal and behavioral biometrics as a potential estrus alert, receiver operating characteristics curve analyses were performed to identify the most optimum reference value (threshold value) for each thermal and behavioral variable. To evaluate each variable, the period between the presence of a DF (>15 mm diameter) and that of a regressing *corpus luteum* (<20 mm diameter; Perry et al., 2017; Burnett et al., 2018), 48–24 h before the disappearance of the DF, was used as indirect indicators of the estrus period retrospectively. To identify the optimum reference value, a balanced proportion of Sensitivity (probability of testing positive when estrus occurred) and Specificity (probability of testing negative in the absence of estrus) were used. To summarize the level of accuracy for each biometrical parameter, a Youden J index (**YJ** = (True positives/(True positives + False negatives) + True negatives/(True negatives – False negatives)) − 1) was used to give equal weight to false- positive and false- negative values ranging from 0 (e.g. worthless test) to 1 (e.g. perfect test). Additional evaluation tools were added to the evaluation of performance such as the positive predictive value (positive predicted value = True positives/(True positives + False positives)) or percentage of cows with a positive test that were in estrus, the negative predictive value (negative predictive value = True negatives/(True negatives + False negatives)) or percentage of cows with a negative test that were non estrus, and the effectiveness or proportion of all test results that were positive results.

Parallel to receiver operating characteristics curve analyses, estrus alerts were also evaluated by calculating the diagnostic odds ratio to identify the odds of a positive test if the cows were in estrus relative to the odds of a test being positive if the cow was not in estrus (diagnostic odds ratio = (True positives/False positives)/(False negatives/True negatives)). The diagnostic odds ratio ranges from 0 to infinity, thus a higher diagnostic odds ratio is indicative of a higher estrus alert test performance. The diagnostic odds ratio analyses also measured the likelihood ratio of the test, the probability of the test to be correct (**LR+ =** sensitivity/1- specificity) vs the probability of the test to be negative result (**LR− =** 1-sensitivity/ specificity) to identify the occurrence of a true positive compared to the true negative test. Efficiency was calculated as the probability that all tests are correct (Efficiency = (True positives + True negatives)/(True positives + True negatives + False Positives + False negatives)).

Raw IR and Res IR from Vtail, Vnotail and Vlips were evaluated for AM and PM milking separately due to the significant differences between skin temperature during AM and PM results found in a previous experiment (Perez Marquez et al., 2019). Similarly, all behavioral biometrics were analyzed for both Sample times (AM and PM). The test of accuracy was performed for all variables individually, and further evaluations were performed with multiple thermal parameters, multiple behavioral biometrics and combined thermal and behavioral parameters. To combine multiple infrared parameters and behavioral biometrics, the parameters were evaluated retrospectively. The 'True Estrus Positive alert' was

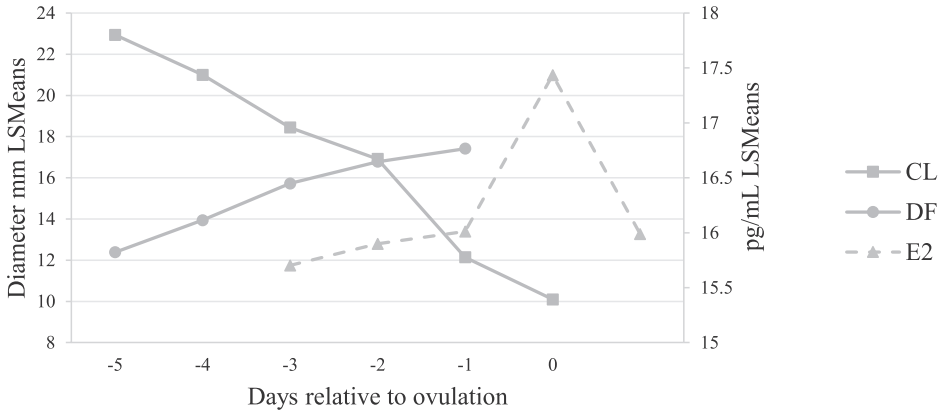

Fig. 5. Diameter in mm (Least-square Means; LSMeans) of ovarian structures and estradiol (E_2) concentrations in skimmed milk as ovulation approaches in cows. *Corpus luteum* (CL) started to regress at d −4 until the lowest diameter during ovulation. Dominant follicle (DF) diameter was at its largest on d −1, however, the peak of E_2 was found until d 0. A weak negative correlation was found between the CL diameter and E_2 concentrations ($P = 0.05$, $r = -0.22$); however, no significant correlations were found between the DF diameter and E_2 concentrations ($P = 0.51$, $r = -0.06$).

determined when all variables were positive (infrared, behavioral, and physiological) 48–24 h before ovulation (d −2 and d −1). If the infrared and behavioral parameters flagged an estrus alert outside the 48–24 h (d −2 and d −1) before ovulation window was defined as a 'False-positive estrus alert'. A True negative estrus alert was when all variables were negative at a nonestrus period (baseline, ovulation and postovulation). The False-Negative alert was declared if only one or multiple parameters did not flag an estrus alert at the expected estrus period (d −2 and d −1). The same rule was applied for all infrared and behavioral biometrics combinations.

Results

Physiological parameters

The average length of estrous cycles was 21.66 ± 3.09 (mean ± SD) days ranging from 17 to 31 days. Changes in the size of the DF and *corpus luteum* and the concentration of estradiol over the pre-estrus, estrus and postovulation periods are shown in Fig. 5. Regression of the *corpus luteum* was confirmed after reduction of *corpus luteum* diameter started at d−4 (d −5: 22.9 ± 1.03 mm to d −4: 21 ± 1.03 mm), and the smallest *corpus luteum* diameter was reported at d −1 (12.13 ± 1.00 mm). Dominant follicle diameter was at its largest measurement during d −1 (17.41 ± 0.64 mm) compared with d −5 (12.38 ± 0.66 mm), which disappeared on the day of ovulation (d 0). Higher estradiol concentrations were found in 12 cows out of 18 used, however, only eight cows had estradiol concentration peaks during the study; three cows at d −2, two cows at d −1, and three cows at d 0. Mean concentrations of estradiol in skimmed milk peaked during the AM Sample time d 0 (17.43 ± 1.76 pg/mL) compared with the PM d 0 (15.98 ± 1.68 pg/mL) and proestrus (d −4 15.70 ± 1.67 pg/mL). However, estradiol concentrations started to increase at d −2 (16.38 ± 1.71 pg/mL) compared with d −4 (15.70 ± 1.67 pg/mL), and d −3 (15.8 ± 1.68 pg/mL see Fig. 5). No correlations (positive or negative) were found between estradiol concentrations and DF diameter ($P = 0.51$, $r = -0.06$). However, a negative correlation was found between *corpus luteum* diameter ($P = 0.05$, $r = -0.22$) and estradiol concentrations.

Changes in skin temperature

The University of Alberta Dairy Research and Technology Centre facility experienced minimal daily variation in ambient temperature and relative air humidity during the study period (temperature; 14.05 ± 3.06 °C, relative air humidity; 68.86% ± 6.94 (Mean ± SD)). The relationship between ambient and animal skin temperature was an average $r = 0.62$ ($P = 0.32$).

Changes in Raw IR at the vulva resulted significant ($P < 0.05$) at PM milking compared to AM milking on days leading to ovulation (see Fig. 6A and B). Significant differences were also observed by Sample day for Res IR results (Fig. 6C and D). Specifically, a significant increase in skin temperature was observed during d −2 PM milking compared to baseline and ovulation day; however, no significant interactions between Sample day and Sample time were found ($P > 0.10$). An increase in skin temperature was also observed in Res IR, with an increase at d −2 of 0.51 ± 0.23 °C compared to d −5 and d 0 PM scan days (Fig. 6D). However, changes in Res IR had less variation between vulva measurements compared with the Raw IR (Fig. 6D). No significant correlations (positive or negative) were found between the peak of skimmed milk estradiol concentration and Raw/Res IR increases ($r > 0.10$).

Behavioral frequencies

The frequency of hip and tail movements did not differ between AM and PM milking times. Further, there were no significant changes in tail movements over the sampling period of d −5 to d 2 (Fig. 7A). However, S-hip and L-hip movements tended to increase over sampling periods but only for the AM milking ($P = 0.07$ and 0.06, respectively). S-hip movements significantly increased ($P < 0.01$) over the sample period during the PM milking (Fig. 7B). No significant interactions were found ($P > 0.10$) between Sample day and Sample time in behavioral parameters.

Accuracy evaluation results

Optimum reference values (e.g. threshold value) with Sensitivity and Specificity level (e.g. highest value of True estrus positives and True estrus negatives) and corresponding YJ index for all thermal and behavioral parameters are presented in Table 1. Residual IR Vtail during PM milking (YJ = 0.34) yielded the highest scores which coincided with the highest diagnostic odds ratio score (diagnostic odds ratio = 4.58), LR+ (1.85), negative predictive value (0.82), Efficiency (0.66), and was the second highest scoring test for positive predicted value (0.50) and Specificity (0.50). Thermal and behavioral biometrics at AM milking did not result in the same diagnostic performance compared with thermal and behavioral biometrics at PM milking (Table 1) with the exception of S-hip dur-

ing the AM milking. Overall, changes in S-hip and L-hip movement frequency was the most important behavioral biometric for use as part of an estrus diagnostic test.

A total of 120 possible combinations between behavioral and thermal parameters were evaluated at different Sample times (AM–PM) using the optimum reference value for each parameter. However, only the 20 combinations with the highest scores are presented in Table 2. The highest YJ score was found for Raw IR Vlips PMRes IR Vnotail PM (YJ = 0.35) with a balanced Sensitivity (0.65) and Specificity (0.70) but a lower diagnostic odds ratio (4.83). The

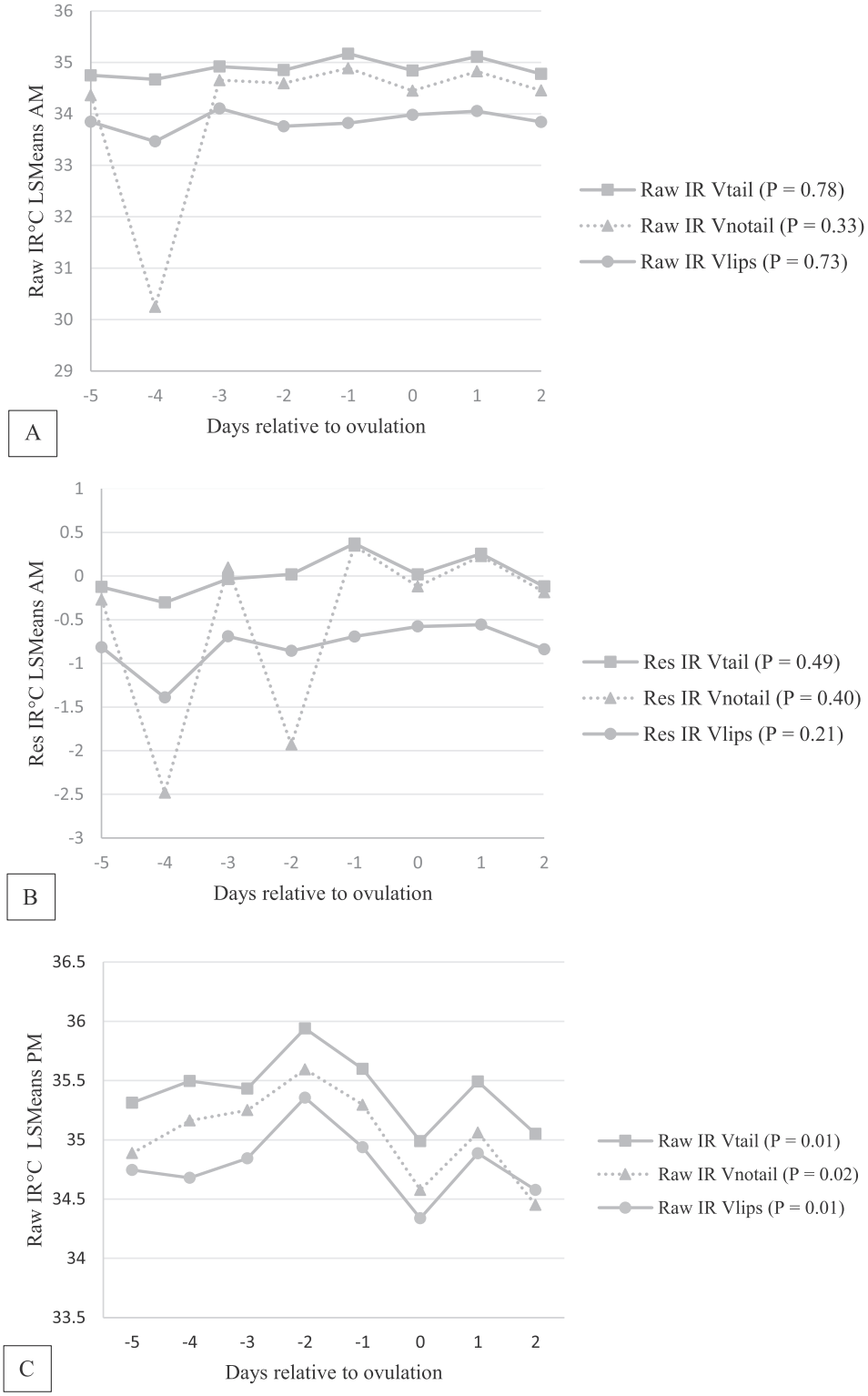

Fig. 6. Raw IR (AM; A, PM; C) and Res IR (AM; B, PM; D) from Vtail, Vnotail, and Vlips (see Table 1 for abbreviations) during milking as ovulation approached in cows. Thermal increases were observed in both infrared thermography (IRT) measurements specifically during d −2 and significant decreases during d 0 which coincided with E_2 concentrations and DF diameter. However, by accounting for ambient temperature (subtracting the predicted skin temperature based on ambient temperature and the observed skin temperature), Res IR data were more consistent compared to Raw IR.

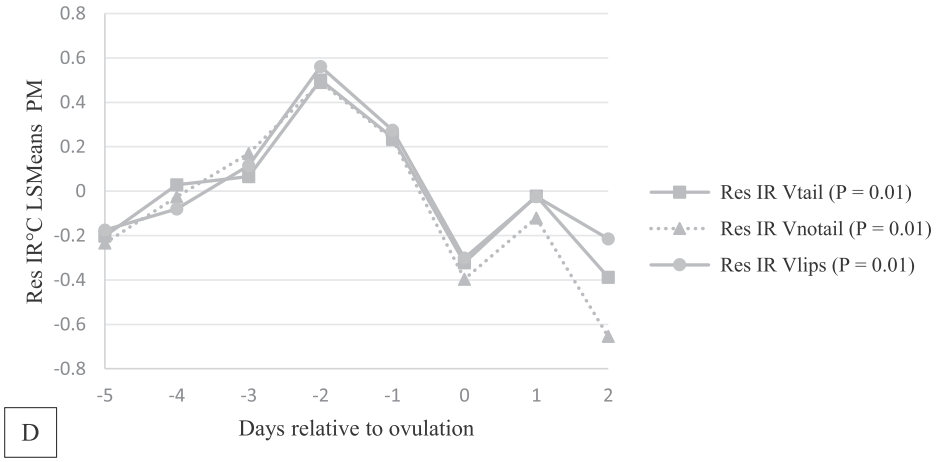

Fig. 6 (*continued*)

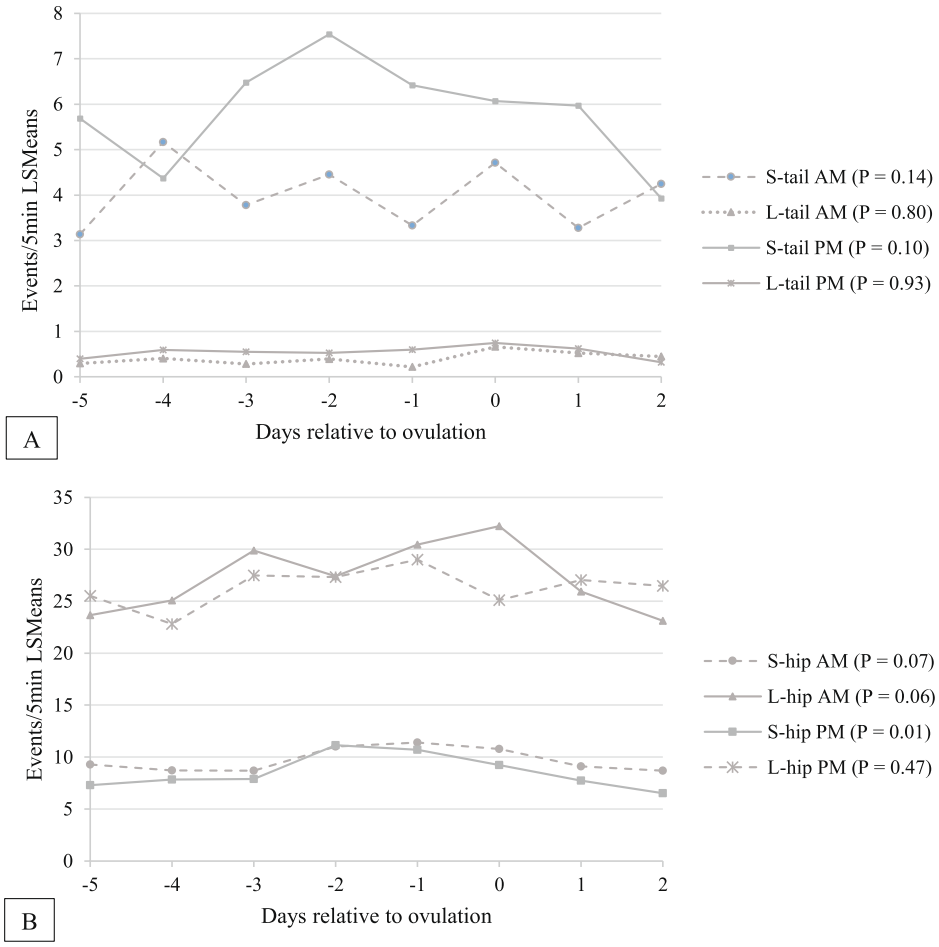

Fig. 7. Behavioral measurements followed a change in frequency of events (Events/5min, Least-square Means; LSMeans) in cows. (A) Increases in d −2 in S-hip during PM milking time ($P = 0.01$) and AM milking ($P = 0.07$) were observed followed by a decrease after ovulation day. Changes in L-hip AM during days relative to ovulation resulted in a tendency (see Table 1 for abbreviations). However, PM milking did not follow a pattern relative to ovulation. (B) Changes in tail movement were observed in S-tail AM-PM compared to L-tail, however, none of the tail frequencies of event were statistically significant ($P \geq 0.10$).

highest diagnostic odds ratio (42.05) was observed for Res IR Vlips, PM S-hip, AM S-hip PM and provided the highest positive predicted value (0.94), Sp (0.99), LR+ (15.35) but a low Sensitivity (0.22) and YJ (0.20). Greater Efficiency (0.77) was found in Res IR Vtail PM + S-hip AM + S-hip PM with an YJ (0.32), diagnostic odds ratio (26.62), LR+ (12.47), Specificity (0.97) and positive predicted value (0.90) compared to other combinations (see Table 2). Additionally, the number of cows flagged in estrus with higher diagnostic odds ratio (Res IR Vlips, PM S-hip, AM S-hip PM) was only four cows (True estrus positive) but 0 False-Positive alerts compared to Res IR Vtail PM + S-hip AM + S-hip PM (diagnostic odds ratio = 26.62; seven cows in estrus with 1 False-Positive alerts). In contrast, the highest YJ (Raw IR Vlips PM + Res IR Vnotail PM = 0.35) found 12 cows in flagged estrus with 20 False-Positive alerts.

Table 1
Accuracy evaluation of thermal and behavioral biometrics as individual parameters for all primiparous cows using a receiver operating characteristics and a diagnostic of odds ratio.

Parameter	Stime	Threshold	Se	Sp	Efficiency	PPV	NPV	YJ	DOR	LR+	LR−
CL	–	<12.06	0.37	0.97	0.84	0.78	0.85	0.34	19.54	12.71	0.65
DF	–	>15.90	0.89	0.61	0.67	0.39	0.95	0.50	13.22	2.29	0.17
E_2	–	>17.05	0.37	0.70	0.63	0.25	0.80	0.06	1.33	1.21	0.91
Vtail	AM	>34.16	0.76	0.27	0.44	0.36	0.7	0.02	1.27	1.03	0.91
Res Vtail	AM	>0.18	0.51	0.59	0.57	0.41	0.71	0.11	1.66	1.26	0.82
Vnotail	AM	>34.80	0.54	0.58	0.57	0.41	0.72	0.12	1.75	1.28	0.8
Res Vnotail	AM	>0.73	0.78	0.28	0.46	0.37	0.73	0.07	1.6	1.09	0.77
Vlips	AM	>34.71	0.24	0.82	0.63	0.43	0.68	0.06	1.62	1.33	0.93
Res Vlips	AM	>−1.67	0.81	0.29	0.46	0.37	0.74	0.09	1.66	1.13	0.68
Vtail	PM	>35.28	0.68	0.52	0.58	0.43	0.77	0.2	2.48	1.41	0.62
Res Vtail	PM	>0.14	0.73	0.61	0.66	0.5	0.82	0.34	4.58	1.85	0.45
Vnotail	PM	>35.00	0.62	0.56	0.59	0.44	0.75	0.19	2.31	1.42	0.67
Res Vnotail	PM	>−0.14	0.76	0.51	0.6	0.45	0.81	0.26	3.55	1.54	0.48
Vlips	PM	>34.10	0.78	0.44	0.56	0.43	0.81	0.22	3.14	1.39	0.5
Res Vlips	PM	>0.23	0.57	0.7	0.67	0.51	0.77	0.27	3.42	1.92	0.61
S-tail	AM	>33	0	0.94	0.62	0	0.65	−0.06	0	0	1.06
L-tail	AM	>49	0	0.94	0.62	0	0.65	−0.06	0	0	1.06
S-tail	PM	>13	0.22	0.83	0.62	0.40	0.67	0.05	1.38	1.3	0.94
L-tail	PM	>44	0.03	0.96	0.64	0.25	0.66	−0.02	0.64	0.65	1.02
S-hip	PM	>6	0.86	0.37	0.54	0.41	0.84	0.23	3.66	1.37	0.37
L-hip	PM	>38	0.44	0.83	0.7	0.57	0.74	0.27	3.87	2.59	0.67
S-hip	AM	>13	0.47	0.79	0.68	0.53	0.74	0.26	3.28	2.2	0.67
L-hip	AM	>29	0.53	0.54	0.54	0.37	0.69	0.07	1.33	1.15	0.87

Abbreviations: Stime = data collection of thermal and behavioral biometrics per day; AM = morning milking (3 AM); PM = afternoon milking (3 PM); Se = Sensitivity; Sp = Specificity; PPV = Positive predicted value; NPV = Negative predicted value; YJ = Youden J index; DOR = Diagnostic odds ratio; LR+ = Positive likelihood ratio; LR− = Negative likelihood ratio; CL = *Corpus luteum*; DF = Dominant follicle; $^{12}E_2$ = Estradiol; Vtail = Skin temperature from the vulva with tail; Res = Residual; Vlips = Skin temperature from the vulva's external lips; Vnotail = Skin temperature from the vulva without tail; S-tail = Small tail movements within the rear thermal area of the cow; L-tail = Large tail movements outside the cow's thermal area of each cow; S-hip = Hip small movements side to side (e.g. left–right) within 10 cm from the rest position (e.g. standing still); L-hip = Hip large movements beyond 10 cm from the rest position.

Table 2
Accuracy evaluation of combined thermal and behavioral biometrics for all primiparous cows using a receiver operating characteristics and a diagnostic of odds ratio. Only significant results (>0.30 YJ or > 1.00 DOR) from combinations within and between thermal and behavioral biometrics.

Combined Parameters	Stime	Se	Sp	Efficiency	PPV	NPV	YJ	DOR	LR+	LR−
Res IR Vtail + Res IR Vlips	PM	0.73	0.56	0.63	0.47	0.81	0.29	3.84	1.67	0.48
Vlips – Res IR Vnotail – Res IR Vlips	PM	0.57	0.75	0.69	0.55	0.78	0.31	4.24	2.24	0.58
Vlips – Res IR Vtail – Res IR Vlips	PM	0.57	0.77	0.71	0.58	0.78	0.34	4.97	2.52	0.56
Vtail – Vlips – Res IR Vtail – Res IR Vlips	PM	0.51	0.77	0.69	0.56	0.76	0.29	3.99	2.28	0.63
Vlips – Res IR Vlips	PM	0.57	0.73	0.69	0.54	0.77	0.30	3.94	2.12	0.59
Vlips – Res IR Vtail	PM	0.65	0.70	0.69	0.54	0.80	0.35	4.83	2.19	0.50
Vlips – Res IR Vlips	PM	0.57	0.73	0.69	0.54	0.77	0.30	3.94	2.12	0.59
S-hip – L-hip	AM	0.97	0.06	0.38	0.35	0.90	0.03	4.94	1.03	0.48
S-hip – S-tail	AM	0.97	0.06	0.38	0.35	0.90	0.03	4.94	1.03	0.48
S-hip – L-tail	AM	0.97	0.07	0.39	0.36	0.92	0.04	6.13	1.05	0.38
L-hip – L-tail	AM	0.97	0.06	0.38	0.35	0.90	0.03	4.94	1.03	0.48
L-hip – S-tail	AM	0.97	0.06	0.38	0.35	0.90	0.03	4.94	1.03	0.48
S-hip – L-hip – L-tail	AM	0.16	0.99	0.71	0.93	0.70	0.15	30.05	11.51	0.85
Raw IR[1]	PM	0.54	0.62	0.60	0.44	0.73	0.16	2.09	1.42	0.74
Res IR[2]	PM	0.57	0.76	0.70	0.57	0.78	0.33	4.58	2.37	0.57
Res IR VtailPM – S-hipAM – S-hipPM	–	0.35	0.97	0.77	0.90	0.75	0.32	26.62	12.47	0.67
Res IR VnotailPM – S-hipAM – S-hipPM	–	0.35	0.96	0.76	0.84	0.74	0.31	15.74	8.32	0.68
Res IR VlipsPM – S-hipAM – S-hipPM	–	0.22	0.99	0.73	0.94	0.71	0.20	42.05	15.35	0.79
VtailPM – S-hipAM-S-hipPM	–	0.32	0.96	0.75	0.83	0.74	0.28	13.98	7.68	0.71
VnotailPM – S-hipAM – S-hipPM	–	0.27	0.97	0.74	0.88	0.72	0.24	18.36	9.59	0.75
Vlips – S-hipAM – S-hipPM	–	0.32	0.93	0.73	0.74	0.73	0.25	7.54	4.61	0.73

Abbreviations: Stime = data collection of thermal and behavioral biometrics per day; AM = morning milking (3 AM); PM = afternoon milking (3 PM); Se = Sensitivity; Sp = Specificity; PPV = Positive predicted value; NPV = Negative predicted value; YJ = Youden J index; DOR = Diagnostic odds ratio; LR+ = Positive likelihood ratio; LR− = Negative likelihood ratio; Raw IR = Raw skin temperature; Res IR = Residual skin temperature See Table 1 for all other abbreviations.
[1] Combination between all the Raw IR parameters in a given Sample time (AM – PM).
[2] Combination between all the Res IR parameters in a given Sample time (AM – PM).

Discussion

Physiological associations with thermal radiation fluctuations

Ambient temperature and percentage of relative air humidity were maintained consistently through the summer–spring season inside the Dairy Research and Technology Centre barn that may explain the nonsignificant impact of ambient parameters in infrared readings. However, this relationship was not observed in all animals, with some animals exhibiting significant relationship while other animals did not. This was most likely due to the environmental monitor being placed in a fixed position in the barn and thus the data recorded were unrepresentative of some of the cow stalls. Furthermore, some animals were in closer proximity to air circulation fans compared to others, which likely would have affected. When ambient temperature and animal skin temperature

were pooled, there was no significant effect giving the impression that ambient temperature and relative air humidity did not affect animal skin temperature ($P > 0.10$). Other studies using infrared technology found air circulation, solar loading, camera distance, emissivity, and percentage of relative air humidity influencing infrared readings (Cook et al., 2016, Perez Marquez et al., 2019).

The differences in thermal data found between AM and PM measurements could be attributed to the thermogenic effect of the heat increment of feeding. Feeding took place at 0600 h, after AM milking was finished and infrared images had been recorded, which may explain the lower infrared readings found at the AM Sample time because animals had not yet been offered fresh fed in the morning. Similar temperature changes after feeding intake were found by Montanholi et al. (2010), and Freetly et al. (2006). Another factor potentially affecting infrared readings was the lower air temperature found in the barn during the AM milking compared to PM milking. Despite the lack of correlation between ambient temperature and animal skin temperature, animals thermoregulate to their environment and thus ambient temperatures in the PM might have affected thermoregulation resulting in higher thermal radiation. Another possibility is that circadian rhythms in body core temperature have been widely reported in humans (Costa et al., 2018), sheep (D'Alterio et al., 2011), and dairy cattle (Berry et al., 2003) and this may be the case in dairy cows such that the overall greater activity during the day results in higher skin temperature in PM images. Other studies confirmed that the increase of activity during the day could increase the volume of blood circulating specifically at the skin level (Rahim et al., 2018; Cramer et al., 2019), and higher infrared 48 h prior ovulation were found during PM milking compared to AM milking in a tie-stalls (Perez Marquez et al., 2019).

The fluctuations between d −5 and d 2 of the estrous cycle (proestrus – estrus – ovulation – postovulation) in Raw and Res IR coincided with greater DF diameter (>15 mm), regression of the *corpus luteum* (<20 mm) and estradiol concentrations in skim milk (17.43 pg/mL) at d −1. However, the highest increases in skin temperature observed on d −2 did not match with the larger DF diameter and the peak of estradiol concentrations at d −1. Notwithstanding, the peak in skin temperature coincided with the interaction *corpus luteum* (regression) – DF (development) at d −2 (see Fig. 4). No significant correlations existed between estradiol and skin temperature increases. Potential explanations with the changes in infrared during the presence of larger DF and increases of estradiol have been related with the increase of physical activity at the onset of estrus (Oshi et al., 2006) reported in other estrus detection studies on tie-stall herds (Kennedy and Ingalls, 1995, Guesgen and Bench, 2018). Thermal fluctuations may be related to the changes in endocrine profile during the follicular phase (e.g. gonadotropin releasing hormone, follicle stimulating hormone–luteinizing hormone, estradiol–progesterone interaction, cortisol levels etc.) effect in skin physiology (Frascarolo et al., 1990). Other studies suggested that the thermogenesis is associated to estradiol release during estrus in visceral fat and skeletal muscle through adaptive thermogenesis (Brown et al., 2010; Clarke et al., 2013). The increased activity that precedes the standing to be mounted and the changes in the blood perfusion at the vaginal and vulva area had been reported as major factors that increase temperature in the vulva (Oshi et al., 2006). Similar results to this study have found increases in vagina temperature 24 h prior to ovulation, followed by a decrease in blood flow during ovulation (Hassan et al., 2017).

Thermal areas (Vtail, Vnotail, and Vlips) did not differ significantly when measuring Raw and Res IR. However, even when all the thermal areas follow similar thermal patterns (e.g. increases and decreases), Vlips resulted in a slightly lower skin temperature compared to Vtail and Vnotail. The low Vlips skin temperatures can be attributed to the presence of feces, and urine in the outer lips of the vulva, which creates a moisture environment as the tail does not allow the lips to dry-off. Note: The vulva's outer lips were not clean or dry-off in the current study, as it would not be feasible under barn conditions. Additionally, the maximum skin temperatures were found in pixels around the vulva area, which explains why Vtail and Vnotail were able to identify the changes in skin temperature similarly as Vlips.

Behavioral changes during the expected estrus period

Behavioral data were continuously recorded using an 'all-occurrence sampling' at AM and PM milking times to identify temporal changes in events (Lehner, 1996) in the days leading to ovulation. In the current study, we did not find differences between milking times in the frequency of any of the behavioral events. In a previous study, we found increases in restless behaviors before milking compared to during and after milking (Perez Marquez et al., 2019), which can be an indication of discomfort, or in anticipation to milking (Metz-Stefanowska et al., 1992).

On the other hand, S-hip movements were found to be significantly higher at the d −2 compared to the proestrus period and ovulation day at the PM milking, and there was a tendency for S-hip during AM milking to be higher ($P = 0.07$). Changes in S-hip movements at the PM milking may be related to the increase in activity during the estrus period in dairy cows. Restless behavior (such as number of steps and shifting of weight shifting between legs) has been reported with the potential use to detect differences in standing comfort and as a response to lameness (Chapinal et al., 2009). Similar to the present study, Guesgen and Bench (2018) identified micro-movements in the pelvis 24 h before ovulation in naturally cycling dairy cows in tie-stall housing using 3D kinematic analysis. However, in other studies (Valenza et al., 2012; Burnett et al., 2018), the interval from increased activity to ovulation was approximately 24 h. The different intervals in the increase of activity to ovulation may be due to the differences in data collection periods such as the 24 h window between ultrasonography in the current study vs 12 h window between ultrasonography in Burnett et al. (2018), different behaviors, and housing type (e.g. tie-stalls vs free-stalls). Additionally, behavioral data collection in the present study only occurred during milking compared with other studies in which activity bouts, for example, were assessed using activity monitors on cows housed in free-stalls (during nonmilking time). Other behaviors such as small-tail movement and large-tail movement did not differ as ovulation approached. Some of the factors that may have affected our results were the potential for miscalculation of tail movements due to the velocity of tail movements being faster than could be captured by the frame rate of the camera, (4 frames/s) and the inter- and intra-cow variation. Other factors that may affect the overall tail movements can be related to the presence of flies during the months of summer and early fall at the study location. Frantz et al. (2019) reported changes in tail movement caused by fly population and this was associated with footstep movements, which indicates that the tail frequency of tail movements may not be entirely attributed to the restless behavior at the estrus period. However, the presence of flies was mitigated in the present study by placing flytraps inside the barn.

Individual estrus alerts vs combined estrus alerts

Estrus alerts were constructed using the changes in Raw IR, the Res IR, and changes in the frequencies of hip and tail movements during milking. The evaluation of accuracy for thermal data identified differences between AM and PM Sample times, with a higher score obtained during PM scanning. The observation of significant

changes in skin temperature of the vulva as ovulation approach is consistent with a previous study reporting the same response in synchronized multiparous cows (Perez Marquez et al., 2019). The YJ index resulted in a positive test to identify estrus for thermal data. However, the positive test balances the proportion of false-positive and false-negative occurrence as well (e.g. high true-positive estrus alerts have high false-positive proportions). This means that optimum reference values can be adjusted depending on the objective of the diagnostic test (e.g. cost per artificial insemination relative to pregnancy cost). Furthermore, by comparing Raw IR and Res IR, we observed that Res IR had higher scores (Table 1). Residual IR accounted for thermal regulation to ambient temperature, which may have resulted in a closer approximation to thermal radiated attributed to physiological process such as the estrus period. Further, the combination of infrared parameters increased the YJ index ($\leq$0.30) during PM scanning. One reason for this is that the more parameters used in an estrus alert reduced the error rate attributed to the false-positive and false-negative estrus occurrence. Mainly, Res IR Vtail, Res IR Vlips, Raw IR Vlips PM and Raw IR Vtail PM Res Vlips were found with the highest YJ index for all combinations which means that infrared parameters are particularly better to identify estrus with a balance sensitivity and specificity. However, the balanced proportion of sensitivity and specificity often results in adding a proportion of false negative and false positive that should have to be tested in practical circumstances (e.g. cost per arterial insemination relative to cost of pregnancy). Additionally, the accuracy of the infrared camera (±0.45 °C) could influence the occurrence in false-positive alerts and explain the unknown increase in raw and residual skin temperature during the project; nevertheless, the use of a black body during infrared recording can help to estimate the error rate in a particular scanning session. Regardless of a higher YJ index in infrared combinations, the diagnostic odds ratio analyses did not show higher results (3.84–4.97) probably for the balanced Sensitivity and Specificity, which in estrus detection, the number of true negative estrus has heavier weight since estrus occurred once in the 21-day estrous cycle.

Behavioral biometrics did not show considerable differences between AM and PM evaluation of accuracy results, which may be, explained by the lack of significant differences in behavioral frequencies between milking schedules. Higher accuracy scores were achieved for hip movements during AM and PM milking. The lack of significant differences in tail movements can be related to the lack of tail movement during milking time while Perez Marquez et al. (2019) reported an increase in tail movement as milking time approached, tail movements tended to decrease or be absent during milking. The combination of behavioral biometrics did not improve the YJ score or diagnostic odds ratio with the exception of S-hip, L-hip, large-tail movement with a higher diagnostic odds ratio and positive predicted value. The combination of these three behaviors correctly identified the total of true negative occurrence and false positive, however, the Sensitivity was low.

The combinations between thermal and behavioral biometrics resulted in higher Efficiency, positive predicted value, diagnostic odds ratio and in some cases YJ index. The explanations to these results may be the complementary information as parameters are added (true positives can be confirmed if more than one parameter coincided), for example; the highest combinations consisted of higher scores from individual evaluations (behavioral and thermal) which decrease the error rate. Similar results were found by Hoffmann et al. (2013), by looking at activity monitor estrus detection methods combined with visual observations. The additive effects of using multiple methods in a diagnosis reduces the error rate by eliminating the number of false positives and increasing the identification of true negative occurrence.

The current study objective was to compare and contrast the combined thermal and behavioral biometrics as estrus alerts at an estimated estrus period. Our null hypothesis states that no changes in the accuracy were expected between estrus alerts using thermal and behavioral biometrics individually, compared to combined thermal-behavioral estrus alerts. We rejected the null hypotheses since adding behavioral parameters to thermal estrus alerts reduced the number of both false positive and false negative tests. Additionally, residual thermal measurements were found to be more accurate compared to raw thermal measurements for estrus detection from the vulva area. Infrared thermography from Vtail, Vnotail and Vlips followed the same patterns of fluctuation on the days leading up to ovulation. The resolution in small hip movements was important to distinguish the estrus period from the proestrus stage and ovulation day compared to large hip movements in a tie-stall. The data suggest that the use of multiple parameters has utility as an estrus detection method by combining infrared data from the vulva and smallhip movements during milking in primiparous cows housed in a tie-stall barn.

Ethics approval

This study was approved by the Animal Care and Use Committee: Livestock (University of Alberta, Edmonton Alberta, Canada, approval number: AUP00001652) under the Canadian Council of Animal Care Standards and requirements (2009).

Data and model availability statement

None of the data and model were deposited in an official repository.

Author's ORCID

H.J. Perez Marquez: https://orcid.org/0000-0003-2075-5948.
D.J. Ambrose: https://orcid.org/0000-0002-3313-7199.
C.J. Bench: https://orcid.org/0000-0003-0996-0334.

Author contribution

H. J. Perez Marquez contributed with the data collection, statistical analysis and manuscript drafts. D. J. Ambrose was involved in experimental design, provided advice regarding the reproductive physiology aspects, assisted with ultrasonography data collection and manuscript reviews. A. L. Schaefer contributed with expertise in the IR thermography outputs and interpretation. N. J. Cook assisted with data analysis, estradiol analysis and manuscript edits. C. J. Bench was involved in experimental design, objectives and hypothesis formulation as well as manuscript revisions.

Declaration of interest

The authors declare no conflict of interest.

Acknowledgements

The authors thank the Alberta Livestock and Meat Agency and Alberta Milk (Grant; 2015F042R) for providing funding for this research. In addition, the authors acknowledge the Department of Agricultural, Food and Nutritional Science of the University of Alberta, Animal Inframetrics, and InterAg New Zealand for providing in-kind support. We also acknowledge Dr Mirjam Guesgen for assistance with infrared scanning and behavioral coding, PhD student Gobikrushanth Mohanathas for assistance with ultrasound

scanning, and the Dairy Research and Technology Centre staff who provided assistance with animal care and handling.

Financial support statement

Financial support for this research was provided by Alberta Livestock and Meat Agency and Alberta Milk (Grant; 2015F042R) and in-kind contributions were provided by the Department of Agricultural, Food and Nutritional Science at the University of Alberta, Animal Inframetrics, and InterAg New Zealand.

References

Adriaens, I., Huybrechts, T., Geerinckx, K., Daems, D., Lammertyn, J., De Ketelaere, B., Saeys, W., Aernouts, B., 2017. Mathematical characterization of the milk progesterone profile as a leg up to individualized monitoring of reproduction status in dairy cows. Theriogenology 103, 44–51. https://doi.org/10.1016/j.theriogenology.2017.07.040.

Aungier, S.P.M., Roche, J.F., Sheehy, M., Crowe, M.A., 2012. Effects of management and health on the use of activity monitoring for estrus detection in dairy cows. Journal of Dairy Science 95, 2452–2466. https://doi.org/10.3168/jds.2011-4653.

Berry, R.J., Kennedy, A.D., Scott, S.L., Kyle, B.L., Schaefer, A.L., 2003. Daily variation in the udder surface temperature of dairy cows measured by infrared thermography: potential for mastitis detection. Canadian Journal of Animal Science 83, 687–693. https://doi.org/10.4141/A03-012.

Berz, R., 2007. The medical use of infrared-thermography history and recent applications. In Proceedings of the Thermografic-Kolloquium, 27–28 September 2007, Hessen, Germany, pp. 1–12.

Boyd, R., 1983. Radiometry and the Detection of Optical Radiation. Wiley-Interscience, New York, NY, USA.

Brown, L.M., Gent, L., Davis, K., Clegg, D.J., 2010. Metabolic impact of sex hormones on obesity. Brain Research 1350, 77–85. https://doi.org/10.1016/j.brainres.2010.04.056.

Burnett, T.A., Polsky, L., Kaur, M., Cerri, R.L.A., 2018. Effect of estrous expression on timing and failure of ovulation of Holstein dairy cows using automated activity monitors. Journal of Dairy Science 101, 11310–11320. https://doi.org/10.3168/jds.2018-15151.

Canadian Council on Animal Care, 2009. CCAC guidelines on: the care and use of farm animals in research, teaching and testing. Retrieved on 8 June 2017 from www.ccac.ca.

Chanvallon, A., Coyral-Castel, S., Gatien, J., Lamy, J.M., Ribaud, D., Allain, C., Clément, P., Salvetti, P., 2014. Comparison of three devices for the automated detection of estrus in dairy cows. Theriogenology 82, 734–741. https://doi.org/10.1016/j.theriogenology.2014.06.010.

Chapinal, N., de Passillé, A.M., Weary, D.M., von Keyserlingk, M.A.G., Rushen, J., 2009. Using gait score, walking speed, and lying behavior to detect hoof lesions in dairy cows. Journal of Dairy Science 92, 4365–4374. https://doi.org/10.3168/jds.2009-2115.

Clarke, S.D., Clarke, I.J., Rao, A., Evans, R.G., Henry, B.A., 2013. Differential effects of acute and chronic estrogen treatment on thermogenic and metabolic pathways in ovariectomized sheep. Journal of Endocrinology 154, 184–192. https://doi.org/10.1210/en.2012-1758.

Collier, R.J., Dahl, G.E., VanBaale, M.J., 2006. Major advances associated with environmental effects on dairy cattle. Journal of Dairy Science 89, 1244–1253. https://doi.org/10.3168/jds.S0022-0302(06)72193-2.

Cook, N.J., Timsit, E., Colyn, J., Chabot, B., Liu, T., Wolfger, B., Schaefer, A.L., 2016. Assessing consistency in radiated thermal output of beef steers by infrared thermography. Journal of Imaging 2, 21–38. https://doi.org/10.3390/jimaging2030021.

Costa, C.M.A., Gomes-Moreirab, D., Sillero-Quintanac, M., Britod, C.J., Azambuja-Pussieldie, G., Fernandesb, A., Piñonosa-Canoc, S., Bouzas-Marins, J.C., 2018. Daily rhythm of skin temperature of women evaluated by infrared thermal imaging. Journal of Thermal Biology 72, 1–9. https://doi.org/10.1016/j.jtherbio.2017.12.002.

Cramer, M.N., Moralez, G., Huang, M.U., Crandall, C.G., 2019. No thermoregulatory impairment in skin graft donor sites during exercise-heat stress. Medicine and Science in Sports & Exercise 51, 868–873. https://doi.org/10.1249/MSS.0000000000001883.

D'Alterio, G., Casella, S., Gatto, M., Gianesella, M., Piccione, G., Morgante, M., 2011. Circadian rhythm of foot temperature assessed using infrared thermography in sheep. Czech Journal of Animal Science 56, 293–300. https://doi.org/10.17221/CJAS10.17221/CJAS-710.17221/1294-CJAS.

Daltro, D., Fischer, V., Alfonzo, E.P.N., Dalcin, V.C., Stumpf, M.T., Kolling, G.J., Barbosa da Silva, M.V.G., McManus, C., 2017. Infrared thermography as a method for evaluating the heat tolerance in dairy cows. Brazilian Journal of Animal Science 5, 374–383. https://doi.org/10.3390/s150717258.

De Vries, A., 2006. Economic value of pregnancy in dairy cattle. Journal of Dairy Science 89, 3876–3885. https://doi.org/10.3168/jds.S0022-0302(06)72430-4.

Delwiche, M., Tang, X., BonDurant, R., Munro, C., 2001. Improved biosensor for measurement of progesterone in bovine milk. Transactions of the American Society of Agricultural Engineers 44, 1997–2002. https://doi.org/10.13031/2013.6976.

Denis-Robichaud, J., Cerri, R.L.A., Jones-Bitton, A., LeBlanc, S.J., 2016. Survey of reproduction management on Canadian dairy farms. Journal of Dairy Science 99 (11), 9339–9351. https://doi.org/10.3168/jds.2016-11445.

Felton, C.A., Colazo, M.G., Ponce-Barajas, P., Bench, C.J., Ambrose, D.J., 2012. Dairy cows continuously-housed in tie-stalls failed to manifest activity changes during estrus. Canadian Journal of Animal Science 92, 189–196. https://doi.org/10.4141/cjas2011-134.

Firk, R., Stamer, E., Junge, W., Krieter, J., 2002. Automation of oestrus detection in dairy cows: a review. Livestock Production Science 75, 219–232. https://doi.org/10.1016/S0301-6226(01)00323-2.

Fisher, A.D., Roberts, N., Bluett, S.J., Verkerk, G.A., Matthews, L.R., 2008. Effects of shade provision on the behaviour, body temperature and milk production of grazing dairy cows during a New Zealand summer. New Zealand Journal of Agriculture Research 51, 99–105. https://doi.org/10.1080/00288230809510439.

Fordham, D.P., Rowlinson, P., McCarthy, T.T., 1988. Oestrus detection in dairy cows by milk temperature measurement. Research of Veterinary Science 44, 366–374. PMID: 3406533.

Frantz, L.M., Morabito, E.A., Dolecheck, K.A., Bewley, J.M., 2019. Short communication: A comparison of cow cleanliness, fly population, and fly avoidance behaviors among docked, switch-trimmed, and switch-intact dairy cows in 3 commercial dairy herds. Journal of Dairy Science 102, 1584–1588. https://doi.org/10.3168/jds.2018-14921.

Frascarolo, P., Schutz, Y., Jequier, E., 1990. Decreased thermal conductance during the luteal phase of the menstrual cycle in women. Journal of Applied Physiology 69, 2029–2033. https://doi.org/10.1152/jappl.1990.69.6.2029.

Freetly, H.C., Nienaber, J.A., Brown-Brandl, T., 2006. Changes in heat production by mature cows after changes in feeding level. Journal of Animal Science 84, 1429–1438. PMID: 16699100.

Giordano, J.O., Stangaferro, L., Wijma, R., Chandler, W.C., Watters, R.D., 2015. Reproductive performance of dairy cows managed with a program aimed at increasing insemination of cows in estrus based on increased physical activity and fertility of timed artificial inseminations. Journal of Dairy Science 98, 2488–2501. https://doi.org/10.3168/jds.2014-896.

Glencross, R.G., Esslemont, R.J., Bryant, M.J., Pope, G.S., 1981. Relationships between the incidence of the pre-ovulatory behaviour and the concentrations of oestradiol-17 β and progesterone in bovine plasma. Applied Animal Ethology 7, 141–148.

Grgurevic, N., Koracin, J., Majdic, G., Snoj, T., 2016. Effect of dietary estrogens from bovine milk on blood hormone levels and reproductive organs in mice. Journal of Dairy Science 99, 6005–6013. https://doi.org/10.3168/jds.2015-10610.

Grummer, R.R., Rastani, R.R., 2003. When should lactating dairy cows reach positive energy balance. Professional Animal Scientist 19, 197–203. https://doi.org/10.15232/S1080-7446(15)31405-4.

Guesgen, M., Bench, C.J., 2018. Using kinematics to detect micro-behavioural changes relative to estrus in naturally cycling tie-stall dairy heifers. Livestock Science 207, 126–132. https://doi.org/10.1016/j.livsci.2017.11.019.

Hassan, M., Sattar, A., Bilal, M., Avais, M., Ahmad, N., 2017. Evaluation of changes in blood flow of the uterine artery by Doppler ultrasonography during the estrous cycle in lactating Bos indicus cows. Animal Reproduction Science 184, 78–85. https://doi.org/10.1016/j.anireprosci.2017.07.001.

Hoffmann, G., Schmidt, M., Ammon, C., Rose-Meierhöfer, S., Burfeind, O., Heuwieser, W., Berg, W., 2013. Monitoring the body temperature of cows and calves using video recordings from an infrared thermography camera. Veterinarian Research Communication 37, 91–99. https://doi.org/10.1007/s11259-012-9549-3.

Jain, A.K., Ross, A., Prabhakar, S., 2004. An introduction to biometric recognition. IEEE Transactions on Circuits and Systems for Video Technology 14, 4–20. https://doi.org/10.1109/TCSVT.2003.818349.

Kennedy, A.D., Ingalls, J.R., 1995. Estrus detection with activity tags in dairy cows housed in tie-stalls. Short communication. Canadian Journal of Animal Science 75, 633–636. https://doi.org/10.4141/cjas95-094.

Lehner, P.A., 1996. Data Collection Methods. In Handbook of Ethological Methods. (ed. PA Lehner). Cambridge University Press, New York, NY, USA, pp. 197–201.

Lopez, H., Caraviello, D.Z., Satter, L.D., Fricke, P.M., Wiltbank, M.C., 2005. Relationship between level of milk production and multiple ovulations in lactating dairy cows. Journal of Dairy Science 88, 2783–2793. https://doi.org/10.3168/jds.S0022-0302(05)72958-1.

López-Gatius, F., Santolaria, P., Mundet, I., Yániz, J.L., 2005. Walking activity at estrus and subsequent fertility in dairy cows. Theriogenology 63, 1419–1429. https://doi.org/10.1016/j.theriogenology.2004.07.007.

Løvendahl, P., Chagunda, M.G.G., 2010. On the use of physical activity monitoring for estrus detection in dairy cows. Journal of Dairy Science 93, 249–259. https://doi.org/10.3168/jds.2008-1721.

Macmillan, K., Colazo, M.G., Cook, N.J., 2019. Evaluation of infrared thermography compared to rectal temperature to identify illness in early postpartum dairy cows. Research Veterinary Science 125, 315–322. https://doi.org/10.1016/j.rvsc.2019.07.017.

Mayo, L.M., 2015. Assessing the efficacy of automated detection of estrus in dairy cattle. Master thesis, Animal and Food Sciences, University of Kentucky, Lexington, KY, USA.

Meadows, C., Rajala-Schultz, P.J., Frazer, G.S., 2005. A spreadsheet-based model demonstrating the nonuniform economic effects of varying reproductive performance in Ohio dairy herds. Journal of Dairy Science 88, 1244–1254. https://doi.org/10.3168/jds.S0022-0302(05)72791-0.

Metz-Stefanowska, J., Huijsmans, P.J.M., Hogewerf, P.H., Ipema, A.H., Keen, E., 1992. Behaviour of cows before, during and after milking with an automatic milking system. Applied Animal Behaviour Science 65, 278–288. https://doi.org/10.1016/S0168-1591(98)00229-9.

Montanholi, Y.R., Swanson, K.C., Palme, R., Schenkel, F.S., McBride, B.W., Lu, D., Miller, S.P., 2010. Assessing feed efficiency in beef steers through feeding behavior, infrared thermography and glucocorticoids. Animal 4, 692–701. https://doi.org/10.1017/S1751731109991522.

National Research Council, 2001. Nutrient Requirements of Dairy Cattle, 7th Edition. National Academy Press, Washington, DC, USA. https://doi.org/10.17226/9825.

Novotna, I., Langova, L., Havlicek, Z., 2019. Risk factors and detection of lameness using infrared thermography in dairy cows – a review. Annals of Animal Science 3, 563–578. https://doi.org/10.2478/aoas-2019-0008.

Osawa, T., Tanaka, M., Morimatsu, K., Syuto, B., 2004. Used of infrared thermography to detect the change in the body surface temperature with estrus in the cow. Proceedings from the society for theriogenology annual conference. Journal of Animal Science and Technology 59, 21–49. https://doi.org/10.1186/s40781-017-0146-4.

Oshi, M., Dochi, O., Koyama, H., 2006. The observation of estrus behavior at night in dairy cattle on a free stall barn. Journal of Rakuno Gakuen University 30, 229–234. ISSN: 0388–001X.

Perez Marquez, H.J., Ambrose, D.J., Schaefer, A.L., Cook, N.J., Bench, C.J., 2019. Infrared thermography and behavioral biometrics associated with estrus indicators and ovulation in estrus-synchronized dairy cows housed in tiestalls. Journal of Dairy Science 102, 4427–4440. https://doi.org/10.3168/jds.2018-15221.

Perry, G.A., Swansona, O.L., Larimorea, E.L., Perry, B.L., Djirab, G.D., Cushmanc, R.A., 2017. Relationship of follicle size and concentrations of estradiol among cows exhibiting or not exhibiting estrus during a fixed-time AI protocol. Theriogenology 81, 925–935. https://doi.org/10.1016/j.domaniend.2014.02.001.

Rahim, A., Paray, A.R., Khosla, H., Danish, Z., 2018. Infrared thermography as non-invasive technique for early detection of mastitis in dairy animals-a review. Asian Journal of Dairy & Food Research 37, 1–6. https://doi.org/10.18805/ajdfr.R-1746.

Reames, P.S., Hatler, T.B., Hayes, S.H., Ray, D.L., Silvia, W.J., 2011. Differential regulation of estrous behavior and luteinizing hormone secretion by estradiol-17ß in ovariectomized dairy cows. Theriogenology 75, 233–240. https://doi.org/10.1016/j.theriogenology.2010.08.009.

Rollin, B., 2016. Inference for Means: Comparing Two Independent Samples. Retrieved on 1 May 2016 from https://www.stat.ubc.ca/~rollin/stats/ssize/n2.html.

Rutten, C.J., Steeneveld, W., Inchaisri, C., Hogeveen, H., 2014. An ex ante analysis on the use of activity meters for automated estrus detection: To invest or not to invest? Journal of Dairy Science 97 (11), 6869–6887. https://doi.org/10.3168/jds.2014-7948.

Sakaguchi, M., 2010. Oestrous expression and relapse back into anoestrus at early postpartum ovulations in fertile dairy cows. Veterinary Record 167, 446–450. https://doi.org/10.1136/vr.c3853.

Sathiyabarathi, M., Jeyakumar, S., Manimaran, A., Jayaprakash, G., Pushpadass, H.A., Sivaram, M., Ramesha, K.P., Das, D.N., Kataktalware, M.A., Prakash, M.A., Kumar, R.D., 2016. Infrared thermography: a potential noninvasive tool to monitor udder health status in dairy cows. Veterinary World 9, 1075–1081. https://doi.org/10.14202/vetworld.10.14202/vetworld.2016.1075-1081.

Sauls, J.A., Voelz, B.E., Hill, S.L., Mendonça, L.G.D., Stevenson, J.S., 2017. Increasing estrus expression in the lactating dairy cows. Journal of Dairy Science 100, 807–820. https://doi.org/10.3168/jds.2016-11519.

Snoj, T., Zuzek, M.C., Cebulj-Kadunc, N., Majdic, G., 2017. Short communication: Heat treatment and souring do not affect milk estrone and 17B-estradiol concentrations. Journal Dairy Science 101, 61–65. https://doi.org/10.3168/jds.2017-13205.

Sprecher, D.J., Farmer, J.A., Nebel, R.L., Mather, E.C., 1995. The educational implications of reproductive problems identified during investigations at Michigan dairy farms. Theriogenology 43, 373–380. https://doi.org/10.1016/0093-691X(94)00030-X.

Stevenson, J.S., Hill, S.L., Nebel, R.L., DeJarnette, J.M., 2014. Ovulation timing and conception risk after automated activity monitoring in lactating dairy cows. Journal of Dairy Science 97, 4296–4308. https://doi.org/10.3168/jds.2013-7873.

Sykes, D.J., Couvillion, J.S., Cromiak, A., Bowers, S., Schenck, E., Crenshaw, M., Ryan, P.L., 2012. The use of digital infrared thermal imaging to detect estrus in gilts. Theriogenology 78, 147–152. https://doi.org/10.1016/j.theriogenology.2012.01.030.

Talukder, S., Kerrisk, K.L., Ingenhoff, L., Thomson, P.C., Garcia, S.C., Celi, P., 2014. Infrared technology for estrus detection and as a predictor of time of ovulation in dairy cows in a pasture-based system. Theriogenology 81, 925–935. https://doi.org/10.1016/j.theriogenology.2014.01.009.

USDA, 2007. Dairy 2007, Reproductive Practices on U.S. Dairy Operations. Retrieved on 11 December 2019 from https://www.aphis.usda.gov/animal_health/nahms/dairy/downloads/dairy07/Dairy07_is_ReprodPrac_1.pdf.

Valenza, A., Giordano, J.O., Lopes, G., Vincenti, L., Amundson, M.C., Fricke, P.M., 2012. Assessment of an accelerometer system for detection of estrus and treatment with gonadotropin-releasing hormone at the time of insemination in lactating dairy cows. Journal of Dairy Science 95, 7115–7127. https://doi.org/10.3168/jds.2012-5639.

Van Eerdenburg, F.J.C.M., Loeffler, H.S.H., van Vliet, J.H., 1996. Detection of oestrus in dairy cows: a new approach to an old problem. Veterinary Quarterly 18, 52–54. https://doi.org/10.1080/01652176.1996.9694615.

Xu, Z.Z., McKnight, D.J., Vishwanath, R., Pitt, C.J., Burton, L.J., 1998. Estrus detection using radiotelemetry or visual observation and tail painting for dairy cows on pasture. Journal of Dairy Science 81, 2890–2896. https://doi.org/10.3168/jds.S0022-0302(98)75849-7.