Evaluation of two-dimensional accelerometers to monitor behavior of beef calves after castration

Brad J. White, DVM, MS; Johann F. Coetzee, BVSc, PhD; David G. Renter, DVM, PhD; Abram H. Babcock, MS; Daniel U. Thomson, PhD, DVM; Daniel Andresen, PhD

Objective—To determine the accuracy of accelerometers for measuring behavior changes in calves and to determine differences in beef calf behavior from before to after castration.

Animals—3 healthy Holstein calves and 12 healthy beef calves.

Procedures—2-dimensional accelerometers were placed on 3 calves, and data were logged simultaneous to video recording of animal behavior. Resulting data were used to generate and validate predictive models to classify posture (standing or lying) and type of activity (standing in place, walking, eating, getting up, lying awake, or lying sleeping). The algorithms developed were used to conduct a prospective trial to compare calf behavior in the first 24 hours after castration (n = 6) with behavior of noncastrated control calves (6) and with presurgical readings from the same castrated calves.

Results—On the basis of the analysis of the 2-dimensional accelerometer signal, posture was classified with a high degree of accuracy (98.3%) and the specific activity was estimated with a reasonably low misclassification rate (23.5%). Use of the system to compare behavior after castration revealed that castrated calves spent a significantly larger amount of time standing (82.2%), compared with presurgical readings (46.2%).

Objective measures also may be useful to document differences in response to stress-related events. Haley et al. used pedometers to find significant differences in behavioral characteristics of calves that were weaned by use of 2 methods of maternal separation. Accelerometers have been used in other veterinary species to detect lameness and remotely monitor degree of activity.

Prototype systems have been developed to remotely collect cattle physiologic and behavioral data for short periods; however, these measures have not been combined with performance analysis or other measures of pain or stress in cattle in a field research environment. Evaluation of the applicability of remote sensors and individual data analysis techniques to enhance accurate methods of identifying behavioral changes is an important area of research and could lead to a system facilitating collection and analysis of objective behavioral variables.

The purposes of the study reported here were to determine the feasibility and accuracy of the use of 2-dimensional accelerometer measurements to classify postural and activity changes and to use accelerometers for determining behavioral differences in bulls after surgical castration. The overall hypothesis was that accelerometer readings could be used to accurately characterize animal behavior. Further, we believed that quantitative values from accelerometers could be analyzed to assess differences in animal behavior following a painful procedure (i.e., surgical castration).

Materials and Methods

Two studies were performed to address the research objectives. The first project was to collect data (accelerometer and video) and generate a behavior-classification system. The second project used the classification system from the earlier work to determine differences in animal behavior after castration. All experiments were approved by the Animal Care and Use Committee at Kansas State University.

Behavior-classification system—Commercially manufactured remote sensor units were affixed to 3 individually housed Holstein calves (mean ± SD body weight, 74.2 ± 6.2 kg) to develop a posture-and behavior-classification system. Each sensor unit consisted of a mica sensor board including a microelectromechanical system surface micromachined 2-axis, ± 2 g accelerometer, mica wireless microprocessor, and radio transmitter to collect and transfer data. Collected values included sensor identification number, record time (date, hour, minute, and second), packet (tenth of second), and x- and y-axis readings (in millivolts) from the accelerometer.

Data from each unit were collected at 10 Hz and transmitted continuously to a base-station mica microprocessor that logged data to a laptop computer via a universal serial bus connection. Devices were placed on the lateral aspect of the right hind limb just proximal to the metatarsophalangeal joint in a manner such that the y-axis was nearly perpendicular to the ground when the calf was standing (Figure 1). The x-axis was perpendicular to the y-axis (nearly horizontal in a standing position) and parallel to the vertebral column of the calf. Sensors were affixed to calves with a protective plastic housing and canvas straps.

Raw data collected at 10 Hz generated a data set with 10 readings of the x- and y-axis/s. The first step in data analysis was to clean the data and transform the raw readings into comparable, useful information. Accelerometers capture data in the x- and y-axis as a result of movement and the constant pull of gravity. Individual sensor boards were calibrated prior to the experiment in a similar manner as described by other investigators. Each unit was placed on a flat horizontal surface and rotated to generate a signal output corresponding to −1 g and +1 g when the sensitive axis was parallel to the gravitational field. The resulting output was used to transform the raw x- and y-axis readings from each sensor (in millivolts) to a magnitude in terms of acceleration caused by gravity (m/s²) corresponding to the individual sensor.

Data were then filtered to generate a 1-second epoch for each sensor with summary calculations for the period. Calculations were performed to generate variables that have been used in previous research to delineate accelerometer data into posture and activity classifications. Created variables for each second in time included mean, median, and SD for readings in the x- and y-axes and the correlation between x- and y-axis readings. A normalized SMA has been described in human literature as a reasonable measure to evaluate differences between types of activities. The following formula modified to account for only 2 axes was used to calculate the SMA for each 1-second period (T) in the data set:

\[
\text{SMA}_T = \frac{1}{T} \left( \sum_{0}^{T} |x(t)| + \sum_{0}^{T} |y(t)| \right)
\]

where x(t) and y(t) refer to the component readings from each axis at each reading, and t refers to the number of readings in each second (T). The SMA also was calculated in a similar manner for each axis independently.

Calves were video recorded for 2-hour intervals while housed in individual pens separated by concrete.
block walls. Calves were not in visual or physical contact with other animals. The video camera time was synched to the time on the laptop (which provided the master clock for the wireless sensors). Video recording was viewed and manually logged by a single person following a hierarchical structure to classify activity (Figure 2). A posture (standing or lying) and activity (standing in place, walking, eating, getting up or down, sleeping, or awake) were assigned for each second of recorded video for each calf.

To minimize subjective interpretation of calf movement captured on video, specific criteria were used for classification. When a calf was lying down, activity would be classified as awake or asleep. For a calf to be classified as sleeping, no head movements were observed for the previous 5-second period. When a calf was standing, activity would be classified as either standing in place, walking, getting up or down, or eating. Walking was defined as a minimum of 2 progressive steps (either forward or backward), whereas eating was defined as the display of normal eating behaviors (foraging, grazing, chewing, or head down). Getting up or down was the short process of transitioning from standing to lying position, and finally, standing in place was defined as lack of other 3 activities. Standing in place did not mean purely static activity and could involve minor limb movements without progressive movement.

Data generated through video logging were then combined with calculated accelerometer data by use of the synched time to join the 2 data sets. The combined data set was imported into a commercial data-mining software package for analysis. The data set was partitioned into a training (70%) and test (30%) on the basis of common data-mining allocation percentages. The training set was then used to build automatically generated decision trees, logistic regressions, and artificial neural networks to classify posture and activity data on the basis of the known outcomes generated by the video log. Classification was performed with 1-second resolution, or rather, each second of data was classified and compared with video observations of actual behavior. Decision trees and artificial neural networks have been used in previous research to classify activity on the basis of accelerometer data. Decision trees were automatically generated by use of a classification tree node. Eight input variables from each second (mean x-axis reading [m/s^2], mean y-axis reading [m/s^2], SD of x-axis reading [m/s^2], SD of y-axis reading [m/s^2], SMA, SMA of x-axis, SMA of y-axis, and x-y correlation) were used as input in the classification tree. Entropy was used as the splitting rule with no pruning techniques used.

An artificial neural network was created with the classification neural network. Resilient back propagation was used as the training algorithm, and 30 epochs were completed for each network. The networks had 1 hidden layer with 10 nodes and the same 8 input variables as used in the decision tree. Outputs of the predictive models were generated on the basis of classifications of video (Figure 2); therefore, there were 2 output variables for posture classification and 5 for activity classification. The iteration with the highest classification level was selected as the result. Logistic regression also was used with the same 8 input variables to predict posture (standing or lying) in each second.

The test data set was then placed in each model to generate a predicted value for the remaining data points. Predicted values for activity and posture were then compared with actual activity and posture values determined by the video analysis. Kappa statistics were calculated by use of standard statistical software to assess the agreement between results beyond that expected as a result of chance. The model with the highest kappa value was considered the most accurate on the basis of agreement between predictive model and actual behaviors as recorded by video analysis.

Castration experimental protocol—A group of 12 mixed-breed beef bull calves (mean ± SD body weight, 278.1 ± 25.6 kg) were used to evaluate behavioral changes after castration. Calves were randomly assigned to 1 of 2 treatments: castration group or control group. Bulls arrived at the facility the day prior to castration between 3:00 PM and 5:00 PM, and accelerometers were placed on all calves in the same manner described in the classification section.

Accelerometer data were collected for the next 16 to 18 hours prior to castration. At 7:00 AM the next day, castration group calves were restrained and surgically castrated by use of a castrating knife and castration tool in accordance with standard industry practices. Local anesthetics or systemic analgesic agents were not administered to castrated or control calves. Calves in the control group were restrained in a comparable manner, and testosterone was palpated to emulate handling procedures in the castration group. Calves were monitored with accelerometers for 20 to 25 hours after castration, resulting in a total data-collection period between 36 and 43 hours for individual calves.

Raw data were imported into the data-mining software, and the classification trees selected from the behavioral classification system were used to pre-
dict posture and activity for each second for each calf. Predicted posture and activity values for each second recorded for each calf were aggregated to create a proportion of each hour the calf spent in each posture or activity. The transmission and recording technique was not perfect and resulted in some missing data points. Proportions for each type of activity and posture for each hour were based on the number of data points in each category over the total number of known data points for that hour. Two calves were removed from the analysis because 1 calf was missing all presurgical data and the other had no data recorded after castration as a result of technical difficulties. Hours missing over 50% of the data points also were excluded from the analysis.

Data analysis—Data were exported to a commercial statistical software package for analysis. Hypothesis testing was performed with a value of $P < 0.05$ to identify significant differences. Proportions of time in minutes for each posture and activity were analyzed with logit models to test for associations of effects. Effects evaluated included treatment (castrated or control), record hour, and before versus after treatment. For all regression analyses, a first-order autoregressive correlation structure was defined to account for the repeated measures on calves over time, and type 3 likelihood-ratio statistics were used to test for associations of effects.22

Results

Behavior-classification system—Combining accelerometer readings and the manual video classification resulted in 11,323 seconds (3.15 hours) of known activity and posture behavior from the 3 calves. In this analysis, posture was binary (standing or lying). All 3 classification models performed accurately when predicting posture on the basis of variables collected with accelerometers. Calculated kappa values were 0.94 (95% CI, 0.92 to 0.96), 0.94 (95% CI, 0.92 to 0.96), and 0.96 (95% CI, 0.93 to 0.97) for the artificial neural network, logistic regression, and classification tree, respectively. The decision tree for posture resulted in 11,323 seconds (3.15 hours) of known activity types were divided into 6 categories with 4 (standing in place, walking, eating, and getting up or down) directly linked to the posture of standing. The other 2 activity types (asleep or awake) were linked to the activity of lying down. No data points in the data set from the video log were classified as asleep; thus, predictions were based on the remaining 5 activity types. The automatically generated decision tree resulted in 43 leaves. All available variables except the SMA for each individual axis and x-y correlation were used in the classification tree. The classification tree had a significantly higher overall kappa agreement rate (0.65; 95% CI, 0.63 to 0.67) than did the artificial neural network (0.45; 95% CI, 0.42 to 0.49). The confusion matrix for the classification tree detailed the differences in observed and predicted behaviors (Table 2).

Castration trial—Postural behavior was evaluated by comparing the proportion of time calves spend standing before and after castration. Overall, no significant ($P = 0.32$) difference was found in the proportion of hourly time spent standing prior to castration between the control group and castration group. Time 0 was defined as the time period when all calves were restrained and castrated on the basis of treatment group assignment. After time 0, castrated calves spent numerically more time standing than did control calves, but the difference was not significant ($P = 0.21$).

Because data were collected as repeated measures on individual calves, comparisons could be made within each treatment group between the before and after castration periods. Castrated calves spent significantly ($P = 0.035$) more time standing (82.2%) after castration than they did before castration (37.9%). Control calves spent significantly ($P = 0.047$) more time lying down (96.4%) after castration (16.9%).

Table 1—Comparison and percent agreement of predicted (classification tree) versus observed (video analysis) classifications of posture of 3 calves for each second evaluated.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Predicted posture</th>
<th>Predicted activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Standing</td>
<td>Lying</td>
</tr>
<tr>
<td>Observed posture</td>
<td>3,197</td>
<td>14</td>
</tr>
<tr>
<td>Lying (s)</td>
<td>43</td>
<td>1,146</td>
</tr>
<tr>
<td>Total (s)</td>
<td>2,237</td>
<td>1,189</td>
</tr>
<tr>
<td>Kappa (95% CI)</td>
<td>0.65</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Table 2—Comparison and percent agreement of predicted (classification tree) versus observed (video analysis) classifications of type of activity of 3 calves for each second evaluated.
Remote sensors offer an opportunity to log all animal behavior. Accurate documentation of animal behavior can be achieved by continuous monitoring, but the process is labor intensive. Previous work indicates that evaluation of behavior by sample collection via video recording compares favorably to continuous monitoring, but ideal sample collection intervals are ≤ 15 minutes. Although wireless sensors collect continuous data, 1 obstacle is the ability to create the appropriate process to manage the massive amount of data generated by the system. For collected data to be meaningful, they must accurately classify differences in animal behavior. Researchers in human medicine have used video monitoring to evaluate the ability of accelerometers to predict static (postural) and dynamic (activity) classification and found an overall misclassification rate of 10.7%. In our research, the misclassification rate for the decision tree as measured by kappa statistics was 4% for posture classifications and 35% for activity classification.

Methods selected to generate predictive models on the basis of actual data included decision classification trees, artificial neural networks, and logistic regression. Decision trees have been used in human research projects to classify types of activities. Artificial neural networks have been used in veterinary medicine to classify gait data in horses to detect potential lameness. Results from predictive model analysis in the present study revealed that all 3 selected methods were accurate at predicting animal posture within each second on the basis of accelerometer readings.

Placement of the device on the lower limb means that the vertical (y) axis is nearly perpendicular to the source of gravity when the animal is standing and will register nearly 1 g. The x-axis is perpendicular to the y-axis and will record a value of approximately 0 g when the animal is standing (Figure 1). When an animal lies down, the axes shift and the resulting y-axis reading is near 0 g, whereas the x-axis is approximately 1 g providing a relatively large distinction between the major postures. In our study, the classification tree had marginally better results in this data set, compared with results for the neural network and logistic regression models. Accuracy of postural orientation in this study (98.3%) was similar to the accuracy reported by Karantonis et al (94.1%) and Foerster et al (95.8%) when similar systems were used to classify human posture.

The classification tree also predicted the specific type of activity (Figure 2) the most accurately, although the overall agreement rate was relatively low, compared with the agreement rate for posture classification. Parkka et al reported classification accuracies for specific activities in people on the basis of wearablesensor readings ranging from 82% to 86%. There were fewer misclassifications in the aforementioned work; however, the study consisted of data collected from multiple devices, in addition to accelerometers. Mathie et al used a single triaxial accelerometer to evaluate human activity and reported a classification rate with 98% sensitivity and 99% specificity for over 1,300 movements.

Our study had 3 limitations that may have contributed to reduced accuracy of activity classification. First, the hardware used in the study consisted of a portable accelerometer that transmitted data wirelessly to a radio receiver. Calves were housed in stalls with large areas of metal plating that interfered with radio transmissions and resulted in missing data. Thus, movements in a continuous, noninvasive manner with an expected minimal intrusion of the technology on natural animal behavior.

Discussion

Accurate documentation of animal behavior can be achieved by continuous monitoring, but the process is labor intensive. Previous work indicates that evaluation of behavior by sample collection via video recording compares favorably to continuous monitoring, but ideal sample collection intervals are ≤ 15 minutes. Remote sensors offer an opportunity to log all animal behaviors.
there were substantial gaps in data streams from some calves at specific times. Second, accelerometers used in this study were 2-dimensional (vertical and horizontal parallel to calf body) and much of the human research to classify activity is based on readings from 3-dimensional (vertical, horizontal parallel to body, and horizontal perpendicular to body) accelerometers. The importance of adding the third dimension is not directly related to the likelihood of abducting or adducting a limb, but rather, the interrelationship between the 3 axes allows a more accurate representation of a limb in 3-dimensional space.

Finally, in the data set used to create the classification algorithms, data was collected from only 3 calves in a confined environment and a high percentage of the data points (outside of lying down) related to eating or standing in place. Pilot data were collected from 3 Holstein calves over short periods, and applicability of the findings to behaviors or interactions in a group setting is limited. The behaviors of walking or getting up and down were infrequently observed, and thus, the training set contained few values in these categories that could be used for classification. Calves spent much of their time engaged in grazing or licking behavior (classified as eating) or standing in place (with minimal movement). These 2 behaviors do not differ significantly in the amount of movement in the hind limb, resulting in similar readings for the 2 major behaviors observed. Classification of time awake but lying down was relatively accurate owing to the ability of the accelerometers to distinguish posture.

The accelerometers provided a novel insight into calf behavior after castration. Our research indicated that accelerometers can measure differences in posture and activity that may be related to pain. In the 24 hours after castration, calves actually were more active and spent more time standing. This contradicts a previous report suggesting that surgically castrated calves are less active than uncastrated calves. However, this finding is in agreement with the results of other studies that used direct observation, as opposed to open-field tests, to record behavior after castration. The observed increase in activity could be a manifestation of transient pain related to castration.

Monitoring calf behavior after castration has been used to compare types of procedures and effectiveness of analgesic protocols. Ting et al. evaluated standing behavior after castration (via a castration clamp) with or without the addition of analgesics and found that castrated calves spent a larger amount of time standing than uncastrated or medicated calves. The authors concluded that, combined with their other findings, the posture was indicative of the degree of pain felt by the calves. Research in lambs suggests that specific behaviors are useful to distinguish between degrees of pain, but the total percentage of time lambs spend lying does not correlate with differences in treatment. The authors in the aforementioned study only documented behavioral differences for 3 hours after castration, and the current research followed the calves for almost 24 hours after castration with continuous monitoring.

One important aspect of the present study was to evaluate the accelerometers as a potential tool for quantifying behavioral differences in future research trials. Interestingly, behavior in castrated calves was not significantly different from noncastrated controls, and this may have been the result of relatively large amounts of variability in behavior patterns over the period of study. However, when data were analyzed comparing the presurgical and postsurgical readings from the castration group, significant differences were identified. Individuals have unique behavior patterns that impact their responses to environmental stimuli. Therefore, differences could not be determined between treatment groups as a result of high variability between calves; yet, evaluation of calves before and after the castration event by use of this technology allowed identification of significant differences in the percentage of time standing.

These findings support the hypothesis that accelerometers can be used to gauge animal response to painful stimuli and also suggest that establishing an accurate baseline measurement for each animal is a critical component of accurate analysis. Evaluating data from each calf before and after castration accounts for intracalf behavioral differences and provides insight into the relationship between behavior and castration. This result is important because future trials should be structured to facilitate collection of baseline data for comparison.

Assessing pain by monitoring animal behavior is difficult because common measures range from somewhat subjective (posture) to especially subjective (attitude and depression scores). In addition to pain, other nonrelated events, such as excitement or anxiety as a result of the presence of human evaluators, can cause similar changes in the outcome variables. One potential advantage of the remote sensors is the ability to capture data in the natural environment of an animal without human presence; therefore, recorded information should represent the behavior of the animal with minimal outside influence.

Lying behavior is used as a gauge of cow comfort in dairy operations, but previous research illustrates that the timing of data collection is important because dairy cow lying behavior follows daily patterns related to rest and activity. An advantage of the accelerometers is that data are collected continuously and can be aggregated or subdivided to evaluate specific periods. Data were aggregated on an hourly basis for final analysis to allow for identification of daily circadian patterns. Activity and standing behaviors vary throughout the day related to normal daily events (eating, resting, and sleeping). In the observation of data by record hour, control calves appear to follow a similar pattern for standing in the presurgical and postsurgical phases. Castrated calves spent more time standing after castration, but sparse data limited the analysis by each record hour. Visual evaluation revealed specific times of day where the discrepancy between before and after castration appeared larger. Discrepancies occurred during periods (late afternoon and early morning) that were associated with normal resting patterns in previous research. Calves spent more total time standing after castration, yet at times when calves had higher amounts of daily activity, the differences were more difficult to detect. This finding emphasizes the importance of continuous behavioral monitoring to quantify changes in animal
behavior. The accelerometers provided the ability to remotely and continuously monitor animal behavior in a natural setting and identify specific differences.

Monitoring changes in the amount of time cattle spend eating has been used to help identify metritis in dairy cows\(^8\) and bovine respiratory disease in beef calves.\(^7\) The accelerometers measured behavioral patterns that were classified as eating; however, actual calf intake was not directly evaluated. Further research is necessary to determine associations between behavior classified as eating and actual health or well-being.

On the basis of this research, 2-dimensional accelerometers were accurate in classifying animal posture as standing or lying. The accelerometers also had reasonable accuracy at determining movements when calves were active. The limitations of this research were the result of the confined environment, the ability to only record 2 dimensions of data, and the movement of the devices on the limbs of calves. Further research should be conducted to determine whether 3-dimensional accelerometers can improve overall classification rates.

Use of accelerometers in this study provided a novel insight into animal behavior. Results of this study indicate that accelerometers have the potential to be a valuable tool in providing quantitative measures of animal activity and time standing. This research revealed that as a result of high individual variability in behavior patterns, the most appropriate manner to analyze these data is through collection of baseline and postevent data on the same calf for comparison. These tools provided a means to objectively measure differences in calf behavior after a painful event. The identification of a measurable response to pain may facilitate the study of pain mitigation procedures. Future studies are required to validate and standardize these findings with concurrent video camera surveillance on larger numbers of calves in more natural group environments.

References