Technical Note: Using an electronic drinker to monitor competition in dairy cows

Paige V. McDonald, Marina A. G. von Keyserlingk, and Daniel M. Weary*
Animal Welfare Program, University of British Columbia, 2357 Mall, Vancouver, BC, Canada, V6T 1Z4

ABSTRACT

The objective of this study was to determine whether data from an electronic drinking system that tracks individual animal attendance with ear tag transponders could be used to detect social competition between dairy cows at the drinker. Specifically, we sought to identify the interval between one cow leaving the drinker and another cow taking her place that most accurately identified competitive replacements (when physical contact initiated by one cow causes the other to entirely remove her head from the drinker and the initiator subsequently places her head in the same drinker). The optimal interval to accurately identify replacements at the drinker was determined using a receiver operating characteristic curve and by evaluating the sensitivity (Se), specificity (Sp), and confusion matrix values (true positives, true negatives, false positives, false negatives) of different thresholds identified by 3 metrics: the Youden index, the point closest to (0,1) on the curve, and the point where Se roughly equals Sp. Lactating Holstein dairy cows (n = 20) were monitored for 4 consecutive 24-h periods by video recording and with an electronic drinking system. Two periods were used to identify the optimal interval (the baseline set), and the other 2 periods were used to validate the interval (the validation set). The occurrence of a replacement identified by video was paired with the interval between drinking events of 2 cows at the same electronic drinker to identify the interval that best predicted replacement events. Based on the low prevalence of replacement events compared with nonreplacement events, the interval with the lowest number of false positives was considered optimal. Of the 3 potential metrics, the point on the receiver operating characteristic curve where Se and Sp were roughly equal yielded the fewest false positives. The optimal interval to identify replacements at the drinker was ≤29 s, with 82% Se and 83% Sp in the baseline set. This interval was found to have 85% Se and 89% Sp when applied to the validation set. These results indicate that social competition between dairy cows at the drinker can be accurately measured with an electronic drinking system.

Key words: aggressive behavior, water consumption, automation, accuracy

Technical Note

Monitoring behavior using video analysis can be time consuming and subject to observer error and bias. Technology can be used to automate some types of monitoring. For example, in agricultural systems, radio frequency monitoring systems can accurately record feeding frequency, duration, and intake (see Chizzotti et al., 2015). Data from automatic systems have also been used to make inferences about social interactions (see Guzhva et al., 2016). Social competition is often assessed during feeding by recording displacements, typically defined as when physical contact from 1 cow (actor) results in the recipient cow (reactor) withdrawing from the feed bunk (DeVries et al., 2004). This event can be further defined as a replacement if the actor then takes the place of the previous cow at the feeder. Huzzey et al. (2014) found that replacements could be accurately identified with automated feeders by screening the data for ≤26-s intervals between sequential visits of 2 cows at the same feed bin within the interval. This method has since been used to quantify social interactions in studies on feedbunk stocking density (Crossley et al., 2017) and introduction of postpartum cows into new groups (Jensen and Proudfoot, 2017).

Drinkers may also be useful for monitoring social competition. During periods of elevated ambient temperature when cows increase water intake and decrease feed intake (Kadzere et al., 2002), aggressive interactions are high (Coimbra et al., 2012), resulting in more competition at and close to the water source (Vizzotto et al., 2015). Especially during heat stress, water is an important resource for dairy cows, and competition can result in dominant animals preventing subordinate animals from access (Coimbra et al., 2012). Drinkers are also typically provided at high stocking rates,
likely leading to more displacement events (DeVries et al., 2004). To our knowledge, no work to date has investigated whether data from drinking systems that electronically track individual animal attendance can be used to assess social competition.

The objective of this study was to determine whether the interval between one cow leaving the drinker and another cow taking her place could accurately identify competitive replacements at the drinker. The specific aim was to assess a range of intervals using receiver operating characteristic (ROC) curves and, on this basis, recommend a specific interval that can best identify replacements at the drinker.

All procedures were approved by the University of British Columbia Animal Ethics Committee (protocol A14-0040). In September 2013, lactating Holstein dairy cows (n = 20; mean ± SD parity = 3.1 ± 2.2; range of 1–8 lactations) at the University of British Columbia Dairy Education and Research Centre (Agassiz, BC, Canada) were observed for 4 consecutive 24-h periods (mean ± SD ambient temperature = 19.7 ± 4.3°C; range = 12.9–29.3°C). Average DIM at the start of the observation period (mean ± SD) was 11.1 ± 6.9 (range = 1–23), and average milk production over the first 21 d postpartum was 24.8 ± 9.5 kg/d (range = 7.9–41.5 kg/d). Two of the observation periods were assigned as the baseline set (mean ± SD ambient temperature = 18.9 ± 3.8°C; range = 12.9–25.4°C) and used to identify the optimal interval; the 2 other periods were assigned as the validation set (mean ± SD ambient temperature = 20.5 ± 4.7°C; range = 13.9–29.3°C) and thus used to validate the interval. As elevated ambient temperatures affect competitive behavior at the drinker (Vizzotto et al., 2015), the 24-h observation periods were assigned to either the baseline set or the validation set based on mean daily temperature. The baseline set comprised the more moderate 2 d (mean ± SD ambient temperature = 18.3 ± 3.9 and 19.5 ± 3.7°C for baseline set periods 1 and 2, respectively), and the validation set comprised the warmest and coolest days (mean ± SD ambient temperature = 24.0 ± 3.6 and 16.9 ± 2.2°C for validation sets 1 and 2, respectively). In this way, the interval could be assessed in a range of summer conditions. All animals were housed together in a 24-freestall pen equipped with 12 Insentec feed bins and 2 Insentec water bins (Insentec B.V., Marknesse, the Netherlands; for a full description of the facility, see Neave et al., 2017). Animals had ad libitum access to water and were milked and fed twice per day at approximately 0700 and 1700 h and 0800 and 1600 h, respectively. Feed was formulated according to NRC (2001) guidelines. As animals were part of larger studies by Neave et al. (2017, 2018) and Lomb et al. (2018a, b), the sample size was kept to the stocking density (20 cows/pen) required for those studies. Group size and composition were constant during the observations, and all animals present in the pen were included in data collection and analysis. Temperature was recorded continuously at 1-h intervals by the Environment Canada local weather station situated approximately 500 m from the barn.

Each cow was fitted with an ear tag transponder (High Performance ISO Half Duplex Electronic ID Tag; Allflex Canada, Saint-Hyacinthe, QC, Canada) that allowed the electronic drinking system (Rough-age Intake Control system; Insentec B.V.; Figure 1) to record the time each cow entered and left a water bin (for validation of the system, see Chapinal et al., 2007). Social behavior was monitored using video cameras (closed-circuit television camera, model WV-BP330; Panasonic, Osaka, Japan) positioned 6 m above the feed alley of the pen. Cows were identified with unique alphanumeric symbols (1 marking on the back and 1 on the right and left sides) such that it was always possible to identify each cow at the water bin; red lights (100 W) facilitated recording at night. A replacement at a water bin was recorded by the video observer when physical contact initiated by one cow caused the receiving cow to entirely remove her head from the water bin, with the actor subsequently placing her head in the same water bin within 60 s. The same observer rescored one of the periods to assess intraobserver reliability and found this to be high (Cohen’s kappa = 0.95). Earlier work has shown that this definition of a replacement also produces high interobserver reliability (Schirrmann et al., 2011).

All statistical analyses were performed in SAS (version 9.4; SAS Institute Inc., Cary, NC). The interval between drinking events was used as the independent variable, and the occurrence of a replacement or not (as identified from video) was used as the binary outcome variable. To determine the interval between cows drinking at a water bin that best identifies replacements, PROC LOGISTIC was used to construct ROC curves; ROC curves allow for a graphical assessment of the true-positive rate (sensitivity; Se) and false-positive rate (1 – specificity; Sp) for different thresholds of a test (Fawcett, 2006). Sensitivity is defined as the proportion of positive events correctly identified, and Sp is the proportion of negative events correctly identified (Metz, 1978). The results of a binary classification test [true positives, true negatives, false positives (FP), false negatives (FN)] are arranged into a 2 × 2 table called a confusion matrix, from which Se and Sp can be calculated.

Ranging between 0 and 1, the area under the ROC curve (AUC) represents the diagnostic accuracy of the ROC curve as a whole with all possible threshold values (Greiner et al., 2000). An AUC ≤ 0.5 represents a test...
with outcomes worse than chance, and an AUC ≥0.9 is highly accurate.

For a continuous measurement, a threshold must be determined to differentiate between a positive and a negative test result (Gardner and Greiner, 2006). However, Se and Sp are inversely related and the choice of cut-off affects the weight given to each parameter. Methods of determining an optimal cut-off point from an ROC curve are the Youden index (maximizes the sum of Se and Sp; Youden, 1950), distance to (0,1) (the value closest to the “perfect” 100% Se and 100% Sp point; Perkins and Schisterman, 2006), and Se = Sp (the point where Se and Sp are roughly equal; Greiner et al., 1995).

In the baseline set, ROC curves were constructed for 5 cows with the most replacement events as the reactor to examine between-cow variability and then again for the 20 cows pooled as a group. Individual ROC curves were constructed for only these 5 cows as the others engaged in too few replacements as a reactor (range = 1–7) to create reasonable ROC curves. Individual and group (n = 20) ROC curves were analyzed separately to explore how the optimal interval differs at individual and group levels. The ROCPLOT macro was used to generate the 3 metrics [Youden index, distance to (0,1), and Se = Sp] considered in determining the optimal cut-off point. The group ROC was used for further analyses. Sensitivity, Sp, true positives, true negatives, FP, and FN were calculated for each potential cut-off point. The value with the lowest number of FP was chosen as the optimal interval in an a priori decision based on the nature of the data, as the number of nonreplacement events (negative events) greatly outweighed the number of replacement events (positive events). To test the performance of the optimal interval, it was then applied to the validation set using ROC analysis for examination of Se and Sp.

In the baseline set, 5 cows had a total of 174 drinking events with 65 replacements and 109 nonreplacements (see Table 1). All AUC estimates were >0.8, indicating moderate to high test accuracy. Estimates for individual cut-offs ranged from 20 to 36 s; however, regardless of cut-off, most values were associated with high Se and Sp. Using data from all 20 cows, there was a total of 745 drinking events with 124 replacements and 621 nonreplacements. The group ROC curve had an AUC of 0.91 (95% CI: 0.89–0.93), and the optimal cut-off point ranged from 29 to 46 (s) depending on the criterion used (see Figure 2). A cut-off of 29 s was associated with the lowest number of FP (see Table 2), so this value was then applied to the validation set.

In the validation set, the 20 cows had a total of 669 drinking events with 103 replacements and 566 nonreplacements. The group ROC curve had an AUC of 0.94 (95% CI: 0.91–0.96), and the 29-s cut-off point performed well, with an Se of 0.85 (95% CI: 0.77–0.92) and an Sp of 0.89 (95% CI: 0.86–0.91; see Figure 3).

This is the first study to investigate whether social competition at the drinker can be measured with an electronic drinking system. Using a ≤29-s interval between 2 cows sequentially drinking from the same water bin, replacements could be identified with 82 to 85% Se and 83 to 89% Sp. Similar results were found by Huzzey et al. (2014) with respect to electronic feed bins, identifying a 26-s interval as optimal for identifying replacements. The ROC curve analysis differed slightly between the 2 studies, with Huzzey et al. (2014) using the distance to (0,1) metric (point closest to 100% Se and 100% Sp). The current study evaluated 3 metrics [Youden index, distance to (0,1), and Se = Sp], and the point at which Se and Sp are roughly equal was chosen. Comparing 3 commonly used metrics for identifying an optimal cut-off point allowed for more thorough selection of the interval with the best Se, Sp, and confusion matrix values based on the test results and prevalence of replacement events.

![Figure 1](https://example.com/figure1.png)

**Figure 1.** The electronic drinking system (Roughage Intake Control system; Insentec B.V., Marknesse, the Netherlands) used to calculate the interval between sequential drinking events of 2 cows at the same water bin. For this system, each cow was fitted with a unique passive ear tag transponder (pictured) that was detected by the radio frequency identification reader above the water bin to allow access to the water. For each visit to the water bin, the system recorded the cow number, the water bin number, the time the cow entered and left the water bin, the duration of the visit, and the water intake.

Journal of Dairy Science Vol. 102 No. 4, 2019
In a binary classification test, determining the optimal cut-off and weight given to Se and Sp depends on event prevalence and priority given to avoiding FP or FN (Greiner et al., 2000). If negative events greatly outnumber positive events and the consequences of FP are determined to be more detrimental than the consequences of FN, more weight should be placed on Sp. In the current study, nonreplacements occurred at 6 times the rate of replacements (and nonreplacements should normally far exceed replacements under most management conditions), so we considered the interval that had the lowest number of FP (identified by the Se = Sp metric) as optimal. The Youden index applies an equal weight to both Se and Sp (Greiner et al., 2000) and, in the case of the current study data set for the group (n = 20) ROC curve, yielded the greatest number of FP. For the ROC curves of individual cows (n = 5), where the balance between positive and negative events was less pronounced, the interval identified by the Youden index had an Sp equal to that of the Se = Sp point in 3 of the cows and a greater Se in all 5 cows. Theoretically, the value closest to the “perfect” (0,1) point should result in the least number of events misclassified as either FP or FN (Greiner et al., 2000) and, in the case of the current study data set for the group (n = 20) ROC curve, yielded the greatest number of FP. For the ROC curves of individual cows (n = 5), where the balance between positive and negative events was less pronounced, the interval identified by the Youden index had an Sp equal to that of the Se = Sp point in 3 of the cows and a greater Se in all 5 cows. Theoretically, the value closest to the “perfect” (0,1) point should result in the least number of events misclassified as either FP or FN (Perkins and Schisterman, 2006), but the current study found that the point where Se and Sp are roughly equal had the lowest number of FP. We refrained from assessing accuracy, which is the fraction of all events classified correctly, given that it weighs FP and FN equally (Weiss and Provost, 2003) and has an inherent bias when the number of negative events greatly outweighs positive events (He and Garcia, 2009). A classifier can be highly accurate in identifying the many negative events correctly while failing to detect the smaller percentage of positive events. A highly specific interval to identify replacements at the drinker is important to avoid FP and the resulting exaggerated estimation of competition, whereas there are many fewer chances to generate FN.

Feed and water are very different resources over which competitive behavior may differ. For example, feed is

<table>
<thead>
<tr>
<th>Cow</th>
<th>No. of intervals</th>
<th>No. of replacements²</th>
<th>AUC (95% CI)³</th>
<th>Cut-off (s)</th>
<th>Se (95% CI)⁴</th>
<th>Sp (95% CI)⁴</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>26</td>
<td>13</td>
<td>0.97 (0.93–1.00)</td>
<td>28</td>
<td>1.00 (0.75–1.00)</td>
<td>0.92 (0.75–0.99)</td>
</tr>
<tr>
<td>2</td>
<td>48</td>
<td>10</td>
<td>0.91 (0.83–1.00)</td>
<td>28</td>
<td>0.92 (0.64–1.00)</td>
<td>0.92 (0.75–0.99)</td>
</tr>
<tr>
<td>3</td>
<td>28</td>
<td>19</td>
<td>0.93 (0.86–1.00)</td>
<td>31</td>
<td>1.00 (0.82–1.00)</td>
<td>0.86 (0.67–0.96)</td>
</tr>
<tr>
<td>4</td>
<td>39</td>
<td>14</td>
<td>0.83 (0.72–0.94)</td>
<td>36</td>
<td>0.84 (0.60–0.97)</td>
<td>0.86 (0.67–0.96)</td>
</tr>
<tr>
<td>5</td>
<td>33</td>
<td>9</td>
<td>0.94 (0.87–1.00)</td>
<td>24</td>
<td>1.00 (0.66–1.00)</td>
<td>0.85 (0.68–0.95)</td>
</tr>
</tbody>
</table>

1For each cow, the cut-off (s) identified by the Youden index and the point closest to 100% Se and 100% Sp [distance to (0,1) metric] was the same and is listed in the first row. The cut-off (s) identified by the point where Se and Sp are roughly equal (Se = Sp metric) is listed underneath.
²As the reactor.
³Wald 95% CI.
⁴Clopper-Pearson 95% CI.
often delivered once or twice per day, but water is typically available throughout the day. Dairy cows spend more time feeding than drinking per day (5 vs. 0.3 h/d; Dado and Allen, 1994), but daily water intake is greater than feed intake (83.6 vs. 20.6 kg/d for lactating cows; Cardot et al., 2008). Dairy barns also typically provide more linear space per cow to eat than to drink. Despite these differences in how the resources are provided and used, the optimal cut-off to identify replacements was similar. One reason for this similarity may be that the physical design of the drinkers used in the current study was identical to that of the feeders used by Huzzey et al. (2014). Future studies should examine the effect of different feeder and drinker designs, as this is likely to affect competition. As data for this study was collected during the summer, it would also be interesting to see whether the optimal cut-off to identify replacements at the drinker changes in cooler weather conditions.

There are some limitations with the automatic detection of behavior as described in this study. Twenty-nine replacements identified by video were not paired with electronic data, as the electronic drinking system did not register the identity of 1 of the 2 cows involved in the event. Chapinal et al. (2007) noted similar occurrences when validating the water recording system of the same manufacturer. As can be seen in the ROC curve analysis of individual cows (n = 5) in the baseline set (see Table 1), the automatic detection of replacements at the drinker with a ≤29-s interval between sequential visits of 2 cows at the same drinker may work better for some cows compared with others. Binary classification tests will also inevitably result in some misclassifications (Metz, 1978).

In conclusion, this study provides the first evidence that social competition at the drinker can be measured using an electronic drinking system. An interval of ≤29 s between 2 cows drinking from the same water bin sequentially identified competitive replacements with high Se and Sp. Automated measures of drinking behavior may be particularly useful for studies on competition during heat stress. Automatically detecting competition is a novel use of an electronic drinking system.

**ACKNOWLEDGMENTS**

We thank Julie Huzzey, Julia Lomb [University of British Columbia (UBC) Animal Welfare program; Vancouver, BC, Canada], and the staff and students at the UBC Dairy Education and Research Centre for their help in collecting the data used for this analysis. We also thank Annabelle Beaver (UBC Animal Welfare

![Figure 3](https://example.com/figure3.jpg)  
**Figure 3.** The group (n = 20) receiver operating characteristic curve of different interval cut-off points (s) between 2 cows drinking at the same water bin sequentially to identify replacements at the drinker in the validation set. Points are labeled by different interval cut-offs (s). The optimal cut-off (s) identified in the baseline set is in bold.

### Table 2. Optimal cut-off interval (s) between 2 cows drinking at the same water bin to identify replacements at the drinker as identified by the Youden index, the point closest to 100% sensitivity (Se) and 100% specificity (Sp) [distance to (0,1) metric], and the point where Se and Sp are roughly equal (Se = Sp metric)

<table>
<thead>
<tr>
<th>Metric</th>
<th>Cut-off (s)</th>
<th>Se (95% CI)</th>
<th>Sp (95% CI)</th>
<th>TP</th>
<th>TN</th>
<th>FP</th>
<th>FN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Youden index</td>
<td>46</td>
<td>0.95 (0.90-0.98)</td>
<td>0.73 (0.70-0.77)</td>
<td>118</td>
<td>455</td>
<td>166</td>
<td>6</td>
</tr>
<tr>
<td>Distance to (0,1)</td>
<td>33</td>
<td>0.88 (0.81-0.93)</td>
<td>0.79 (0.76-0.82)</td>
<td>109</td>
<td>491</td>
<td>130</td>
<td>15</td>
</tr>
<tr>
<td>Se = Sp</td>
<td>29</td>
<td>0.82 (0.74-0.89)</td>
<td>0.83 (0.80-0.86)</td>
<td>102</td>
<td>515</td>
<td>106</td>
<td>22</td>
</tr>
</tbody>
</table>

1Data are from the baseline set of observations on 20 cows over two 24-h periods. For each cut-off (s), we show the Se, Sp, and confusion matrix values [true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN)]. The interval chosen as the optimal cut-off (s) is in bold.

2Clopper-Pearson 95% CI.

3Number of events classified as TP or FN out of a total of 124 positive events.

4Number of events classified as TN or FP out of a total of 621 negative events.
Program) for her comments on an earlier version of this manuscript. This research was supported by Canada's Natural Sciences and Engineering Research Council (NSERC; Ottawa, ON, Canada) Industrial Research Chair Program with industry contributions from the Dairy Farmers of Canada (Ottawa, ON, Canada), British Columbia Dairy Association (Burnaby, BC, Canada), Westgen Endowment Fund (Milner, BC, Canada), Intervet Canada Corporation (Kirkland, QC, Canada), Novus International Inc. (Oakville, ON, Canada), Zoetis (Kirkland, QC, Canada), BC Cattle Industry Development Fund (Kamloops, BC, Canada), Alberta Milk (Edmonton, AB, Canada), Valacta (St. Anne-de-Bellevue, QC, Canada), and CanWest DHI (Guelph, ON, Canada).

REFERENCES


