# Characterizing the diagnostic sensitivity and specificity of pain biomarkers in cattle using receiver operating characteristic curves

**M. Martin,**[1] ● **M. D. Kleinhenz,**[2]*● **K. S. Schwartzkopf-Genswein,**[3] **D. Melendez,**[3,4] ● **S. Marti,**[3,4] ● **E. A. Pajor,**[5] ●
**E. D. Janzen,**[5] **and J. Coetzee**[1]* ●
[1]Department of Anatomy and Physiology, Kansas State University College of Veterinary Medicine, Manhattan 66506
[2]Department of Clinical Sciences, Kansas State University College of Veterinary Medicine, Manhattan 66506
[3]Agriculture and Agri-Food Canada, Lethbridge Research Centre, Lethbridge, Alberta, Canada T1J 4B1
[4]IRTA, Department of Ruminant Production, Caldes de Montbui, Barcelona, Spain 08140
[5]Department of Production Animal Health, Faculty of Veterinary Medicine, University of Calgary, Canada T2N 1N4

## ABSTRACT

Biomarkers are used to assess pain and analgesic drug efficacy in livestock. However, often the diagnostic sensitivity and specificity of these biomarkers for different painful conditions over time have not been described. Receiver operating characteristic (ROC) curves are graphical plots that illustrate the diagnostic ability of a test as its discrimination threshold is varied. The objective of this analysis was to use area under the curve (AUC) values derived from ROC analysis to characterize the predictive value of potential pain biomarkers at specific time points following a painful stimulus. The biomarkers included in the analysis were plasma cortisol, salivary cortisol, hair cortisol, infrared thermography (IRT), mechanical nociceptive threshold (MNT), substance P, kinematic gait analysis, and a visual analog scale for pain. A total of 7,992 biomarker outcomes collected from 7 pain studies involving pain associated with castration, dehorning, lameness, and abdominal surgery were included in the analysis. Each study consisted of 3 treatments: uncontrolled pain (tissue damage), no pain (handled controls), and analgesic use (tissue damage, administered a nonsteroidal anti-inflammatory drug). Results comparing analgesic effects to uncontrolled pain consistently yielded AUC values >0.7 (95% confidence interval: 0.40 to 0.99) for plasma cortisol (time points: 1.5, 2, 3, 4, 6, and 8 h), hair cortisol (time point: 62 d), and IRT (time point: 72 h). Results comparing analgesic effects to uncontrolled pain consistently yielded AUC values <0.7 (95% confidence interval: 0.28 to 0.90) for salivary cortisol (6, 13, 20, 34, 48, and 62 d); MNT (6, 25, and 49 h); substance P (1, 2, 3, 4, 6, 8, 12, 18, 24, 48, 72, 96, 120, 144, 312, 480, 816, 1,152, and 1,488 h); kinematic gait analysis including area (8, 16, 48, 72, 96, and 120 h), force (8, 16, 24, 48, 72, 96, and 120 h), and pressure (8, 16, 24, 48, 72, 96, and 120 h); and a visual analog scale for pain (1, 2, 3, 4, 5, and 6 d). These results indicate that ROC analysis can be used to characterize the predictive value of pain biomarkers and provide new knowledge on the diagnostic accuracy of pain biomarkers within this data set. This analysis, using data from 7 studies, was a preliminary approach to identify biomarkers and collection time points that could inform additional analytical approaches or meta-analyses with larger sample sizes, which are needed to further validate these hypotheses and conclusions.
**Key words:** analgesia, biomarker, pain, receiver operating characteristic (ROC) curve

## INTRODUCTION

Pain results from mechanical, chemical, or thermal stimulation of nerve endings containing nociceptors (Hudson et al., 2008). Under farm conditions, cattle may undergo elective procedures such as castration and dehorning, which cause pain. They may also experience pain from conditions (such as lameness) or postoperatively (such as after a cesarean section). In a survey of urban citizens, dehorning without pain mitigation was viewed as contentious and not supported (Cardoso et al., 2017), leading to the need for further research into analgesic strategies. Agents that may provide analgesia in cattle include local anesthetics, nonsteroidal anti-inflammatory drugs (**NSAIDs**), opioids, α2-agonists, and *N*-methyl-D-aspartate receptor antagonists (Coetzee, 2013). The American Veterinary Medical Association (AVMA) policy on castration and dehorning states that because these procedures cause pain and discomfort, the use of medications to alleviate pain should be used under the Animal Medicinal Drug Use and Clarification Act (AMDUCA), which allows for extra-label drug

use for pain relief under the oversight of a veterinarian (AVMA, 2019). In the United States, approved analgesic drugs for use in livestock are limited to NSAIDs. Flunixin meglumine, as a transdermal formulation, is the only US Food and Drug Administration (**FDA**)-approved analgesic to specifically control pain in cattle with only an indication for foot rot (FDA, 2017). In the United States, no analgesics are approved to control pain in cattle from castration, dehorning, or surgery. In Canada, meloxicam is approved for cattle to relieve pain and inflammation from castration (Solvet, 2019).

Biomarkers are often used to assess pain and analgesic drug efficacy in livestock. During the drug development process, when submitting data to the FDA for biomarker qualification, one of the components is the characterization of the relationships between the biomarker, the clinical outcomes, and the treatment (Amur et al., 2015). Further characterizing the relationship between analgesic use and biomarker outcomes could be beneficial for future drug approvals. In this study, we chose to characterize biomarkers that were repeatedly collected across 7 pain studies and are commonly used to assess pain in food animals (Stafford and Mellor, 2005a; Heinrich et al., 2010; Stewart et al., 2010; Bustamante et al., 2015; Kleinhenz et al., 2019c). Among the biomarkers currently used to assess pain, the outcomes included in this analysis were cortisol, substance P, infrared thermography (**IRT**), mechanical nociceptive threshold (**MNT**), kinematic gait analysis, and visual analog scale (**VAS**) pain assessment. These biomarkers have been used to quantify an animal's physiologic and behavioral response to a combination of stress, inflammation, and pain; however, some are more invasive than others, some are influenced by restraint, and some can change due to stress rather than pain, which makes it difficult to specifically identify pain. Collecting biomarkers at different time points throughout the stress response following a painful procedure or condition allows us to assess pain over time. However, the nature of biomarker collection can increase stress due to restraint, and the animal may become sensitized or desensitized over time. Thus, determining which time points yield the best diagnostic accuracy would allow for biomarkers to be collected at fewer time points and be less confounded by continued sampling. This brings into question which biomarkers yield acceptable diagnostic accuracy when comparing uncontrolled pain and inflammation to NSAID use at varying time points throughout the stress response following a painful procedure or condition.

Diagnostic accuracy depends on the sensitivity and specificity of pain biomarkers and has not been described for different painful conditions in cattle over time. Receiver operating characteristic (**ROC**) curves



**Receiver Operating Characteristic**
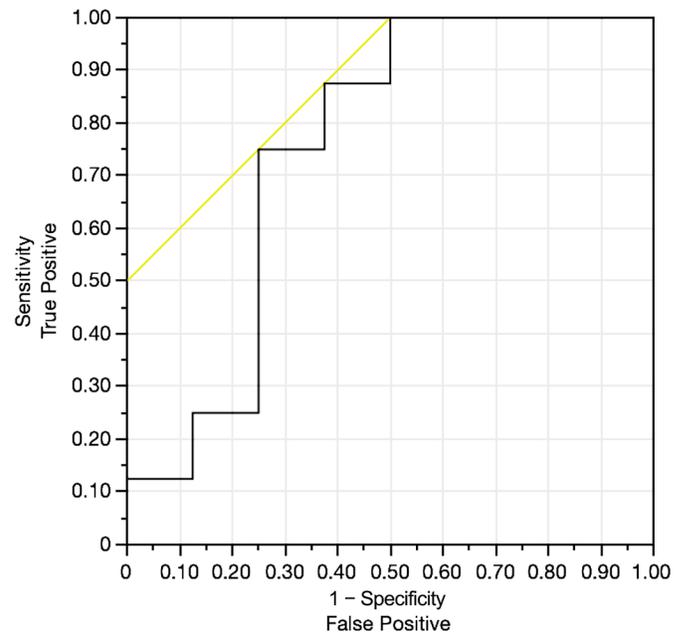
Using Status='Pain' to be the positive level
**AUC**
0.75000

**Figure 1.** Receiver operating characteristic curve example with an area under the curve (AUC) value of 0.75.

are graphical plots that illustrate the diagnostic ability of a test as its discrimination threshold is varied. The plot of true positive (sensitivity) versus false positive (1 − specificity) across possible cut-off values generates a ROC curve (Hajian-Tilaki, 2013). The area under the ROC curve (**AUC**) can be used to measure discriminative ability (the probability that a randomly chosen positive subject is rated or ranked as more likely to be positive than a randomly chosen negative subject; Hajian-Tilaki, 2013; Figure 1). Additionally, AUC values can be compared between ROC curves (Ekelund, 2012). A ROC curve can be constructed for each time point throughout a pain study. The objective of this analysis was to use AUC values derived from ROC analysis to assess the predictive value of pain biomarkers at specific sample collection time points following painful events using already published data from a series of studies. Each study consisted of 3 treatments: (1) uncontrolled pain (tissue damage) and physiologic changes following a painful procedure; (2) no pain (handled controls); and (3) analgesic use mitigating some pain (tissue damage, administered an NSAID) and physiologic changes following a painful procedure. The null hypothesis was that there would be no difference in AUC values across biomarkers or at different sample collection time points.

**Table 1.** Cattle pain studies included in the receiver operating characteristic curve analysis

| Reference | Procedure or condition | Analgesic regimen | Outcome parameter[1] |
|---|---|---|---|
| Kleinhenz et al. (2017) | Dehorning | Transdermal flunixin meglumine 3.33 mg/kg | Plasma cortisol<br>IRT<br>MNT<br>Substance P |
| Meléndez et al. (2017) | Surgical castration | Subcutaneous meloxicam 0.5 mg/kg | Salivary cortisol<br>Substance P |
| Kleinhenz et al. (2018) | Surgical castration | Transdermal flunixin meglumine 3.33 mg/kg | Plasma cortisol<br>IRT<br>Substance P |
| Marti et al. (2018); Meléndez et al. (2018) | Surgical castration | Subcutaneous meloxicam 0.5 mg/kg | Salivary cortisol<br>Hair cortisol<br>Substance P |
| Kleinhenz et al. (2019a) | Lameness | Transdermal flunixin meglumine 3.33 mg/kg | Plasma cortisol<br>Gait analysis<br>Substance P |
| Kleinhenz et al. (2019b) | Abdominal surgery | Intravenous flunixin meglumine 2.2 mg/kg | IRT<br>Substance P |
| Martin et al. (2020) | Surgical castration | Transdermal flunixin meglumine 3.33 mg/kg | VAS score |

[1]IRT = infrared thermography; MNT = mechanical nociceptive threshold; VAS = visual analog scale.

## MATERIALS AND METHODS

### Study Design

Each of the studies was reviewed and approved by the Institutional Animal Care and Use Committee at Iowa State University (Log # 5-15-8016-B, 6-15-8039-B, 7-16-8314-B, and 8-14-7845-B), at Kansas State University (IACUC# 4002), the University of Calgary (ACC14-0159), and the Lethbridge Research Centre Animal Care Committee (ACC# 1410 and 1428).

A total of 7,992 biomarker outcomes were included in the analysis. These outcomes were collected from 351 animals enrolled in 7 studies using cattle. These studies investigated the use of an NSAID administered alone to one treatment group. Each study collected overlapping biomarkers and used similar biomarker collection methods. The studies included in the analysis quantified pain in cattle associated with typical husbandry procedures and conditions, including castration, dehorning, lameness, and abdominal surgery (Table 1) (Kleinhenz et al., 2017, 2018, 2019a,b; Meléndez et al., 2017, 2018; Martin et al., 2020). The same group of calves was involved in 3 of the studies (Kleinhenz et al., 2017, 2018, 2019b); calves were completely healed from the previous study and rerandomized across treatments for each study. The biomarkers included in the analysis were plasma cortisol, salivary cortisol, hair cortisol, IRT, MNT, substance P, kinematic gait analysis, and VAS. Animals were restrained for collection of all biomarker samples except kinematic gait analysis, in which they were walked across the mat, and VAS scoring. Biomarker outcomes for each of the 7 studies are outlined in Table 1, with collection time points outlined in Table 2. All sample time points collected were analyzed; plasma cortisol and substance P were repeatedly collected multiple times on the first collection day, and all other outcomes were continually collected throughout the days following the painful procedure or condition.

All data were collected between 2016 and 2019. Data from multiple castration studies were combined for the substance P outcome, which was collected and analyzed in the same manner at differing time points. All cattle enrolled in the castration studies were castrated surgically. Cattle were dehorned using electro-cautery in the dehorning study, experimentally induced in the lameness study, and abdominal surgery was performed to evaluate pain postoperatively. Cattle received flunixin transdermally in 1 dehorning, 1 lameness, and 2 surgical castrations studies, and intravenously in 1 abdominal surgery study (Table 1). Cattle received meloxicam subcutaneously in 2 surgical castration studies (Table 1). Each study consisted of 3 treatments: uncontrolled pain (tissue damage), no pain (handled controls), and analgesic use (tissue damage, administered a nonsteroi-

**Table 2.** Collection time points for each biomarker

| Biomarker[1] | Time point |
|---|---|
| Plasma cortisol, h | 0.5, 1, 1.5, 2, 3, 4, 6, 8, 12, 24, 48, 72 |
| Salivary cortisol, d | 6, 13, 20, 34, 48, 62 |
| Hair cortisol, d | 34, 62 |
| IRT, h | 1, 2, 3, 4, 6, 8, 12, 24, 48, 72 |
| MNT, h | 6, 25, 49 |
| Substance P, h | 1, 2, 3, 4, 6, 8, 12, 18, 24, 48, 72, 96, 120, 144, 312, 480, 816, 1,152, 1,488 |
| Kinematic gait analysis, h | 8, 16, 24, 48, 72, 96, 120 |
| VAS, d | 1, 2, 3, 4, 5, 6 |

[1]IRT = infrared thermography; MNT = mechanical nociceptive threshold; VAS = visual analog scale.

dal anti-inflammatory drug). Analgesics administered once at the time of the procedure were NSAIDs: flunixin meglumine (3.3 mg/kg of BW transdermally or 2.2. mg/kg of BW intravenously) or meloxicam (0.5 mg/kg of BW subcutaneously). Baseline time points were not included in the ROC analysis due to the lack of comparison between uncontrolled pain, no pain, and analgesic use, as animals had not experienced the painful procedure or condition. However, baseline values can be referenced from individual studies in their previous publications (Kleinhenz et al., 2017, 2018, 2019a,b; Meléndez et al., 2017, 2018; Martin et al., 2020).

### Physiological and Behavioral Biomarkers

**Plasma Cortisol.** A total of 1,564 plasma cortisol samples from 77 animals made up this data set. The samples were obtained as described by Kleinhenz et al. (2017). Blood was obtained using a 14-gauge jugular catheter or by jugular venipuncture using a 20-mL syringe (Monoject) and 16-gauge 3.8-cm needle (Monoject) while cattle were restrained in headlocks (Kleinhenz et al., 2018, 2019a). The blood was immediately transferred to a tube containing sodium heparin and was centrifuged at $3,000 \times g$ for 10 min. The plasma was pipetted into cryovials, placed on dry ice, and stored at $-80°C$ until analysis. Cortisol concentrations were determined using a commercially available radioimmunoassay (MP Biomedicals) with a detection range of 0.64 to 150 ng/mL.

**Salivary Cortisol.** A total of 523 salivary cortisol samples from 106 animals made up this data. As described in Meléndez et al. (2018), calves were restrained in a hydraulic squeeze chute (Cattlelac Cattle, Reg Cox Feedmixers Ltd.), where saliva samples were collected with a cotton swab, immediately stored in a plastic tube, and frozen at $-20°C$ for further cortisol analysis using an ELISA kit (Salimetrics) with a sensitivity of $<0.007$ μg/dL.

**Hair Cortisol.** A total of 175 hair cortisol samples from 72 animals made up this data set. As described in Meléndez et al. (2018), hair from the forehead was clipped and stored in plastic bags at room temperature for further cortisol analysis. Cortisol was quantified using an ELISA (Salimetrics) with a sensitivity of $<0.007$ μg/dL.

**IRT.** A total of 724 IRT measures from 71 animals made up this data set. Infrared images of the medial canthus of the eye for castration, dehorning, and surgery, and images from lame feet were included in the analysis to quantify changes in inflammation in the foot. Mean, maximum, and temperature differentials between the left and right feet from the lameness study were included in the analysis, as discussed in Alsaaod et

al. (2015). Infrared thermography images were obtained using a research-grade infrared camera (FLIR SC 660; FLIR Systems AB). The IRT camera was calibrated before use with ambient temperature and relative humidity. As described in Kleinhenz et al. (2017, 2018), an image of the lateral aspect of the head was obtained so that the image contained the medial canthus of the eye. As described in Kleinhenz et al. (2019b), images of the foot were obtained at a 45° angle, 1 m from the coronary band; 3 images at each time point were averaged, along with the maximum used for analysis, and the difference between the temperatures of the left and right hind feet (left hind minus right hind) were determined for each time point. Infrared images were analyzed using research-grade computer software by drawing a circle around the medial canthus of the eye or the coronary band of the feet and recording the maximum, minimum, and average temperature provided by the software (FLIR ExaminIR Inc.).

**Mechanical Nociception Threshold.** A total of 1,065 MNT measures from 24 animals made up this data set. Calves were restrained using a halter and blindfolded. As described in Kleinhenz et al. (2017), using a handheld pressure algometer (Wagner Instruments), a force was applied perpendicularly at a rate of approximately 1 kg of force (kgf) per second at 2 locations (lateral and caudal) adjacent to the horn bud. A third control location between the eyes was used to evaluate MNT. A withdrawal response was indicated by an overt movement away from the applied pressure algometer, at which time the investigator immediately removed the algometer. Values were recorded by a second investigator to prevent bias from the first investigator, who did not look at the values. The same investigator applied the algometer and the same investigator recorded the values. Locations were tested 3 times in sequential order, and the values were averaged for statistical analysis.

**Substance P.** A total of 1,402 substance P samples from 207 animals made up this data set. Calves were restrained in a hydraulic squeeze chute (Cattlelac Cattle, Reg Cox Feed Mixers Ltd.) for blood collection. Blood samples were collected via jugular venipuncture into vacuum tubes (BD Vacutainer; Becton Dickinson). As described in Kleinhenz et al. (2017), 200 μg of benzamidine was added to EDTA blood tubes (BD Vacutainer) 48 h before the start of the studies. During sample collection, 6 mL of blood was added to the spiked EDTA tube. The samples were immediately placed on ice, centrifuged at $1,500 \times g$ for 10 min at 4°C within 30 min of collection, and the plasma was placed into cryovials. The cryovials were stored at $-80°C$ until analysis. Substance P levels were determined using the methods described by Van Engen et al. (2014) using

**Table 3.** Ethogram for visual analog scale pain behavior scoring

| Parameter | No pain | Moderate pain | Extreme pain |
|---|---|---|---|
| Depression | Alert, bright, quick to show interest, head up | Dull, slow to show interest, head down | Dull, shows no interest, head down |
| Tail swishing/flicking | No swishing/flicking | Few (<3/min) tail swishes or flicks | Numerous (>3/min) tail swishes or flicks |
| Stance | Normal stance | Legs placed slightly abducted | Legs abducted and caudal to normal posture |
| Head carriage | Held above spine level | | Held below spine level |
| Foot stomping/kicking | None | Few stomps, no kicking | Numerous stomps/kicks at belly |

nonextracted plasma. A similar method was used by Meléndez et al. (2018), but 200 µg of benzamidine was not added to the blood tubes before the start of the study to help prevent substance P breakdown. The limit of detection was 10 pg/mL, and the limit of quantitation was 20 pg/mL.

*Kinematic Gait Analysis.* A total of 239 gait analysis readings from 30 animals made up this data set. As described in Kleinhenz et al. (2019a), a commercially available kinematic gait system (MatScan, Tekscan Inc.) was used to record gait and biomechanical parameters from adult cows who walked across the pressure mat at their own pace. The system was calibrated using a known mass daily and before each use of the computer software to ensure accuracy of the measurements at each time point. Video synchronization was used to ensure consistent gait between and within cows at each time point. Using research-specific software (Hugemat Research 5.83, Tekscan Inc.), force, contact pressure, and impulse in the affected feet were assessed, which are the parameters also reported by Schulz et al. (2011).

*Visual Analog Scale.* A total of 2,300 VAS scores from 192 animals made up this data set. As described in Martin et al. (2020), a daily VAS pain assessment was conducted by 2 trained evaluators blinded to treatment allocations on calves of stocker age. The VAS used was a 100-mm (10-cm) line anchored by descriptors of "no pain" on the left (0 cm) and "severe pain" on the right (10 cm). Five parameters were used to assess pain: depression, tail swishing or flicking, stance, head carriage, and foot stomping or kicking (Table 3). "No pain" was characterized by being alert and quick to show interest, no tail swishing, a normal stance, head held above spine level, and absence of foot stomping. "Severe pain" was characterized by being dull and showing no interest, more than 3 tail swishes per minute, legs abducted, head held below spine level, and numerous stomps. The evaluator marked the line between the 2 descriptors to indicate the pain intensity. A millimeter scale was used to measure the score from the zero anchor point to the evaluator's mark. The mean VAS measures of the 2

evaluators were averaged into one score for statistical analysis.

*ROC Curve Determination.* All statistics were performed using statistical software (JMP Pro 14.0, SAS Institute Inc.). Receiver operating characteristic curves were created for each time point, with AUC values comparing uncontrolled pain × no pain (handled controls) and uncontrolled pain × analgesic use, with uncontrolled pain as the positive control. The biomarker outcome was plotted as the $x$, continuous regressor, and status (uncontrolled pain, no pain, analgesic use) was plotted as the $y$, categorical response. Bootstrapping via fractional weights was used to generate confidence intervals for each AUC value; AUC values ≥0.7 were reported due to the ROC rule of thumb used for acceptable discriminative ability described in Yang and Berdine (2017). The maximum (AUC = 1) means that the diagnostic test is perfect in differentiation, AUC = 0.5 means subject discrimination is due to chance, and AUC = 0 means the test incorrectly identifies all subjects (Hajian-Tilaki, 2013). Specific cut-off values were generated through the ROC analysis based upon optimized specificity and sensitivity values by minimizing the square distance between the upper left-hand corner of ROC space and any point on the ROC curve and are presented.

## RESULTS

### Plasma Cortisol

Plasma cortisol AUC values comparing analgesic use versus uncontrolled pain are outlined in Table 4, with rankings in Table 5. Out of 12 time points, the time points that yielded acceptable diagnostic accuracy for surgical castration, dehorning, and lameness when comparing analgesic use to uncontrolled pain are outlined below. Surgical castration study results for plasma cortisol (Figure 2A) yielded acceptable results (AUC >0.7; 95% CI: 0.40 to 0.99) at 1, 1.5, 2, 3, 4, 6, 8, and 12 h, with lower cut-off values following the 1.5-h time point (cut-off values: 17.51, 22.94, 11.70, 5.88, 8.99,

**Table 4.** Area under the curve (AUC)[1] values for analgesia versus pain by collection time point for each biomarker (n = 7,992 outcomes)

| Procedure and biomarker[2] | 1 h | 2 h | 3 h | 4 h | 6 h | 8 h | 12 h | 16 h | 24 h | 48 h | 72 h | 96 h | 120 h | 34 d | 62 d |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Castration | | | | | | | | | | | | | | | |
| Plasma cortisol | **0.75** | **0.98** | **0.91** | **0.90** | **0.72** | **0.72** | **0.87** | | 0.45 | 0.62 | 0.51 | | | | |
| IRT | | 0.620 | | 0.65 | 0.67 | 0.61 | **0.84** | | 0.62 | 0.58 | **0.77** | | | | |
| Substance P | | 0.61 | | 0.68 | 0.61 | 0.59 | 0.60 | | 0.63 | 0.61 | 0.61 | | | | |
| Hair cortisol | | | | | | | | | | | | | | 0.67 | **0.71** |
| Salivary cortisol | | | | | | | | | | | | | | 0.58 | 0.58 |
| VAS | | | | | | | | | 0.63 | 0.58 | 0.54 | 0.62 | 0.56 | | |
| Dehorning | | | | | | | | | | | | | | | |
| Plasma cortisol | 0.67 | **0.79** | **0.75** | **0.74** | **0.71** | **0.72** | 0.67 | | 0.61 | **0.72** | **0.72** | | | | |
| IRT | 0.59 | 0.58 | | 0.58 | 0.68 | | 0.62 | | 0.57 | **0.75** | | | | | |
| MNT | | | | | 0.51 | | | | 0.56 | 0.57 | | | | | |
| Substance P | 0.62 | 0.62 | | 0.62 | | | 0.62 | | 0.61 | 0.61 | 0.62 | | | | |
| Lameness | | | | | | | | | | | | | | | |
| Plasma cortisol | 0.67 | **0.90** | **0.79** | **0.75** | **0.85** | **1.00** | | **0.97** | **1.00** | **0.82** | **0.80** | **0.83** | **0.82** | | |
| Substance P | 0.61 | 0.65 | | | | 0.64 | | | 0.60 | 0.62 | 0.59 | 0.61 | 0.60 | | |
| Pressure mat: | | | | | | | | | | | | | | | |
| Area | | | | | | 0.60 | | 0.63 | 0.70 | 0.64 | 0.61 | 0.58 | 0.59 | | |
| Force | | | | | | 0.62 | | 0.63 | 0.64 | 0.61 | 0.62 | 0.59 | 0.60 | | |
| Impulse | | | | | | 0.64 | | 0.59 | 0.62 | 0.60 | 0.60 | 0.67 | 0.58 | | |
| Pressure | | | | | | 0.58 | | 0.62 | 0.61 | 0.58 | 0.59 | 0.60 | 0.60 | | |
| Abdominal surgery | | | | | | | | | | | | | | | |
| IRT | | | **0.73** | | 0.63 | | 0.57 | | 0.67 | 0.57 | 0.70 | | | | |
| Substance P | | | 0.66 | | 0.64 | | 0.58 | | 0.59 | 0.56 | 0.61 | | | | |

[1]Bold font indicates an AUC value >0.7, which is acceptable discriminative ability for a test (Yang and Berdine, 2017).

[2]IRT = infrared thermography; MNT = mechanical nociceptive threshold; VAS = visual analog scale.

[3]Time points that overlapped across multiple studies are reported in the table; some time points not taken on the hour are not reported in the table.

4.65, 2.52, and 4.58 ng/mL, respectively). Dehorning study results for plasma cortisol (Figure 2B) yielded acceptable results (AUC >0.7; 95% CI: 0.40 to 0.95) at 1.5, 2, 3, 4, 6, 8, 48, and 72 h, with cut-off values fluctuating (cut-off values: 11.32, 9.78, 5.43, 4.12, 5.59, 9.69, 2.60, and 15.57 ng/mL, respectively). Lameness study results for plasma cortisol (Figure 2C) yielded acceptable results (AUC >0.7; 95% CI: 0.42 to 0.99) at

1.5, 2, 3, 4, 6, 8, 16, 24, 32, 40, 48, 56, 64, 72, 80, 96, and 120 h, with lower cut-off values after 40 h (cut-off values: 5.16, 10.16, 8.61, 9.85, 16.47, 11.43, 7.64, 9.92, 6.07, 11.19, 4.43, 4.11, 6.25, 5.33, 7.33, 4.81, and 3.00 ng/mL, respectively).

Plasma cortisol AUC values comparing no pain versus uncontrolled pain are outlined in Table 6, with rankings in Table 7. Out of 12 time points, the time points

**Table 5.** Area under the curve (AUC) rankings >0.7 for analgesia versus pain by collection time point for each biomarker

| Procedure and biomarker[1] | Ranking | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 2 | | 3 | | 4 | | 5 | | 6 | |
| | AUC | Time point, h | AUC | Time point, h | AUC | Time point, h | AUC | Time point, h | AUC | Time point, h | AUC | Time point, h |
| Castration | | | | | | | | | | | | |
| Plasma cortisol | 0.98 | 2 | 0.91 | 3 | 0.90 | 4 | 0.87 | 12 | 0.75 | 1 | 0.72 | 6 |
| IRT | 0.84 | 12 | 0.77 | 72 | | | | | | | | |
| Hair cortisol | 0.71 | 62 | | | | | | | | | | |
| Dehorning | | | | | | | | | | | | |
| Plasma cortisol | 0.79 | 2 | 0.75 | 3 | 0.74 | 4 | 0.72 | 8 | 0.72 | 48 | 0.72 | 72 |
| IRT | 0.75 | 48 | | | | | | | | | | |
| Lameness | | | | | | | | | | | | |
| Plasma cortisol | 1.00 | 24 | 0.97 | 16 | 0.90 | 2 | 0.85 | 6 | 0.83 | 96 | 0.82 | 48 |
| Area | 0.70 | 24 | | | | | | | | | | |
| Abdominal surgery | | | | | | | | | | | | |
| IRT | 0.73 | 3 | 0.70 | 72 | | | | | | | | |

[1]IRT = infrared thermography.
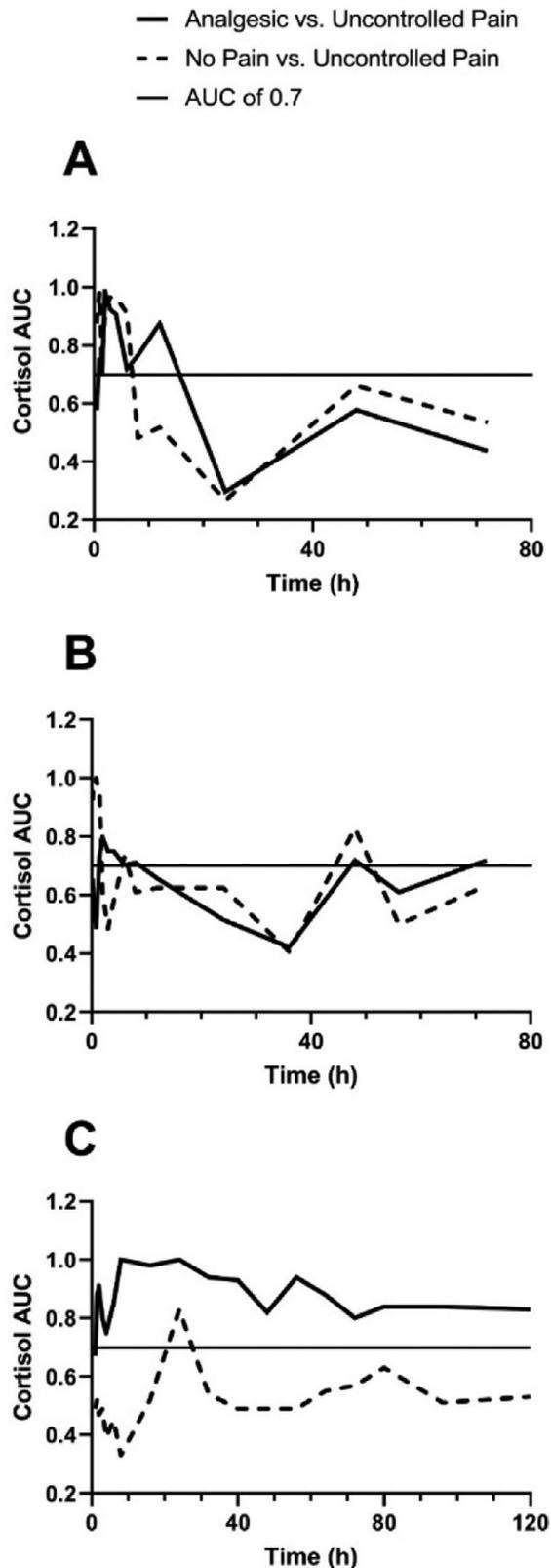
**Figure 2.** Area under the curve (AUC) values for plasma cortisol concentrations following (A) surgical castration, (B) hot iron dehorning, and (C) lameness induction.

that yielded acceptable diagnostic accuracy for surgical castration, dehorning, and lameness when comparing no pain to uncontrolled pain are outlined below. Surgical castration study results for plasma cortisol yielded acceptable results (AUC >0.7; 95% CI: 0.42 to 0.99) at 0.5, 1, 1.5, 2, 3, 4, and 6 h, with cut-off values decreasing from 2 h (cut-off values: 30.34, 17.51, 5.92, 20.04, 14.01, 11.94, and 1.71 ng/mL, respectively). Dehorning study results for plasma cortisol yielded acceptable results (AUC >0.7; 95% CI: 0.42 to 0.99) at 0.17, 0.33, 0.5, 0.67, 0.83, 1, 1.5, 6, and 48 h, with cut-off values decreasing over time (cut-off values: 15.89, 16.36, 16.07, 14.50, 13.62, 9.82, 6.47, 7.83, and 3.78 ng/mL, respectively). Lameness study results for plasma cortisol yielded acceptable results (AUC >0.7; 95% CI: 0.55 to 0.95) at 24 h, with a cut-off value of 12.08 ng/mL. Cut-off values for castration, dehorning, and lameness are outlined in Tables 8, 9, and 10, respectively.

### Salivary Cortisol

Salivary cortisol AUC values comparing analgesic use versus uncontrolled pain are outlined in Table 4. Out of 6 time points, surgical castration study results for plasma cortisol yielded unacceptable diagnostic accuracy (AUC <0.7; 95% CI: 0.35 to 0.82) at all time points examined (6, 13, 20, 34, 48, and 62 d).

Salivary cortisol AUC values comparing no pain versus uncontrolled pain are outlined in Table 6. Out of 6 time points, surgical castration study results for salivary cortisol yielded unacceptable diagnostic accuracy (AUC <0.7; 95% CI: 0.36 to 0.73) at all time points examined (6, 13, 20, 34, 48, and 62 d).

### Hair Cortisol

Hair cortisol AUC values comparing analgesic use versus uncontrolled pain are outlined in Table 4, with rankings in Table 5. Out of 2 time points, surgical castration study results comparing analgesic use versus uncontrolled pain yielded acceptable diagnostic accuracy (AUC >0.7; 95% CI: 0.51 to 0.85) for hair cortisol at 62 d, with a cut-off value of 10.33 pg/mL.

Hair cortisol AUC values comparing no pain versus pain are outlined in Table 6. Out of 2 time points, surgical castration study results comparing no pain versus uncontrolled pain yielded unacceptable diagnostic accuracy (AUC <0.7; 95% CI: 0.44 to 0.71) for hair cortisol at all time points examined.

### Infrared Thermography

Ocular IRT AUC values comparing analgesic use versus uncontrolled pain are outlined in Table 4, with

**Table 6.** Area under the curve (AUC)[1] values for no pain versus pain by collection time point for each biomarker (n = 7,992 outcomes)

| Procedure and biomarker[2] | Time point[3] | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 h | 2 h | 3 h | 4 h | 6 h | 8 h | 12 h | 16 h | 24 h | 48 h | 72 h | 96 h | 120 h | 144 h | 20 d | 34 d | 62 d |
| **Castration** | | | | | | | | | | | | | | | | | |
| Plasma cortisol | **0.98** | **0.92** | **0.97** | **0.97** | **0.91** | 0.62 | 0.62 | | 0.56 | 0.68 | 0.60 | | | | | | |
| IRT | | 0.66 | | 0.64 | **0.73** | **0.79** | 0.62 | | 0.63 | **0.72** | 0.65 | | | | | | |
| Substance P | | 0.66 | | 0.69 | 0.61 | 0.69 | 0.63 | | 0.62 | 0.63 | 0.63 | | | | | | |
| Hair cortisol | | | | | | | | | | | | | | 0.57 | 0.59 | 0.56 | 0.58 |
| Salivary cortisol | | | | | | | | | | | | | | 0.63 | | 0.54 | 0.56 |
| VAS | | | | | | | | | **0.75** | 0.70 | 0.65 | **0.72** | 0.61 | | | | |
| **Dehorning** | | | | | | | | | | | | | | | | | |
| Plasma cortisol | **1.00** | 0.64 | 0.58 | 0.61 | **0.74** | 0.64 | 0.65 | | 0.65 | **0.83** | 0.65 | | | | | | |
| IRT | 0.58 | 0.60 | | 0.60 | 0.59 | | 0.58 | | 0.64 | **0.78** | | | | | | | |
| MNT | 0.63 | 0.58 | | 0.62 | **0.81** | | 0.64 | | **0.88** (25 h) | **0.92** (49 h) | 0.70 | | | | | | |
| Substance P | 0.63 | 0.58 | | 0.62 | | | 0.64 | | 0.62 | 0.62 | 0.70 | | | | | | |
| **Lameness** | | | | | | | | | | | | | | | | | |
| Plasma cortisol | 0.58 | 0.59 | 0.59 | 0.60 | 0.58 | 0.58 | | 0.57 | **0.83** | 0.60 | 0.61 | 0.61 | 0.60 | | | | |
| Substance P | 0.68 | 0.66 | | | | **0.77** | | | **0.83** | 0.68 | 0.67 | 0.68 | **0.71** | | | | |
| Pressure mat: | | | | | | | | | | | | | | | | | |
| Area | | | | | | | | 0.65 | **0.73** | 0.67 | 0.59 | 0.59 | 0.62 | | | | |
| Force | | | | | | | | 0.62 | 0.62 | 0.66 | 0.60 | 0.61 | 0.61 | | | | |
| Impulse | | | | | | | | 0.64 | 0.62 | 0.63 | 0.61 | **0.74** | 0.70 | | | | |
| Pressure | | | | | | | | 0.61 | 0.61 | 0.57 | 0.59 | 0.60 | 0.60 | | | | |
| **Abdominal surgery** | | | | | | | | | | | | | | | | | |
| IRT[1] | | | 0.64 | | 0.64 | | 0.66 | | **0.72** | 0.60 | 0.61 | | | | | | |
| Substance P | | | **0.72** | | 0.70 | | 0.64 | | 0.62 | **0.72** | 0.67 | | | | | | |

[1] Bold font indicates an AUC value >0.7, which is acceptable discriminative ability for a test (Yang and Berdine, 2017).

[2] IRT = infrared thermography; MNT = mechanical nociceptive threshold; VAS = visual analog scale.

[3] Time points that overlapped across multiple studies are reported in the table; some time points not taken on the hour are not reported in the table.

**Table 7.** Area under the curve rankings >0.7 for no pain versus pain by collection time point for each biomarker

| Procedure and biomarker[1] | Ranking | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 2 | | 3 | | 4 | | 5 | | 6 | |
| | AUC | Time point, h | AUC | Time point, h | AUC | Time point, h | AUC | Time point, h | AUC | Time point, h | AUC | Time point, h |
| Castration | | | | | | | | | | | | |
| Plasma cortisol | 0.98 | 1 | 0.97 | 3 | 0.97 | 4 | 0.92 | 2 | 0.91 | 6 | | |
| IRT | 0.79 | 8 | 0.73 | 6 | 0.72 | 48 | | | | | | |
| VAS | 0.75 | 24 | 0.72 | 96 | 0.70 | 48 | | | | | | |
| Dehorning | | | | | | | | | | | | |
| Plasma cortisol | 1.00 | 1 | 0.83 | 48 | 0.74 | 6 | | | | | | |
| MNT | 0.92 | 48 | 0.88 | 24 | 0.81 | 6 | | | | | | |
| IRT | 0.78 | 48 | | | | | | | | | | |
| Substance P | 0.70 | 72 | | | | | | | | | | |
| Lameness | | | | | | | | | | | | |
| Plasma cortisol | 0.83 | 24 | | | | | | | | | | |
| Substance P | 0.83 | 24 | 0.77 | 8 | 0.71 | 120 | | | | | | |
| Impulse | 0.74 | 72 | 0.70 | 120 | | | | | | | | |
| Area | 0.73 | 24 | | | | | | | | | | |
| Abdominal surgery | | | | | | | | | | | | |
| IRT | 0.72 | 24 | | | | | | | | | | |
| Substance P | 0.70 | 6 | 0.72 | 2 | 0.72 | 48 | | | | | | |

[2]IRT = infrared thermography; MNT = mechanical nociceptive threshold; VAS = visual analog scale.

rankings in Table 5. Out of 10 time points, the time points that yielded acceptable diagnostic accuracy for surgical castration, dehorning, and abdominal surgery when comparing analgesic use to uncontrolled pain are outlined below. Surgical castration study results for IRT yielded acceptable results (AUC >0.7; 95% CI: 0.40 to 0.97; Figure 3A) at 12 and 72 h, with cut-off values of 38.6 and 37.3°C, respectively. Dehorning study results for IRT yielded acceptable results (AUC >0.7; 95% CI: 0.53 to 0.89) at 48 h, with a cut-off value of 37.5°C. Abdominal surgery study results for IRT (Figure 3B) yielded acceptable results (AUC >0.7; 95% CI: 0.48 to 0.89) at 3 and 72 h, with cut-off values of 33.2 and 35.1°C, respectively.

Ocular IRT AUC values comparing no pain versus uncontrolled pain are outlined in Table 6, with rankings in Table 7. Out of 10 time points, the time points that yielded acceptable diagnostic accuracy for surgical castration, dehorning, and abdominal surgery when comparing no pain to uncontrolled pain are outlined below. Surgical castration study results for IRT yielded acceptable results (AUC >0.7; 95% CI: 0.44 to 0.94) at 6, 8, and 48 h, with cut-off values of 37.7, 38, and 37.5°C, respectively. Dehorning study results for IRT yielded acceptable results (AUC >0.7; 95% CI: 0.58 to 0.91) at 48 h, with a cut-off value of 37.6°C. Abdominal surgery study results for IRT yielded acceptable results (AUC >0.7; 95% CI: 0.47 to 0.92) at 24 h, with a cut-off value of 35.3°C. Cut-off values for castration, dehorning, and abdominal surgery are outlined in Tables 8, 9, and 11, respectively.

### Mechanical Nociception Threshold

Mechanical nociception threshold AUC values comparing analgesic use versus uncontrolled pain are outlined in Table 4. Out of 3 time points, dehorning study results for MNT yielded unacceptable diagnostic accuracy (AUC <0.7; 95% CI: 0.40 to 0.65) at all time points examined (6, 25, and 49 h).

Mechanical nociception threshold AUC values comparing no pain versus uncontrolled pain are outlined in Table 6, with rankings in Table 7. Out of 3 time points, dehorning study results for MNT yielded acceptable diagnostic accuracy (AUC >0.7; 95% CI: 0.75 to 0.95) at time points 6, 25, and 49 h, with cut-off values decreasing over time (cut-off values: 1.05, 0.89, and 0.78 kgf, respectively).

### Substance P

Substance P AUC values comparing analgesic use versus uncontrolled pain are outlined in Table 4. Out of 19 time points, surgical castration, dehorning, lameness, and abdominal surgery study results comparing analgesia versus pain yielded unacceptable diagnostic accuracy (AUC <0.7; 95% CI: 0.28 to 0.88) for substance P at all the time points examined (1, 2, 3, 4, 6, 8, 12, 18, 24, 48, 72, 96, 120, 144, 312, 480, 816, 1,152, and 1,488 h).

Substance P AUC values comparing no pain versus uncontrolled pain are outlined in Table 6, with rankings in Table 7. Out of 19 time points, the time points that

**Table 8.** Cut-off values for no pain versus pain by collection time point for castration

| Biomarker[1] | Time point[2] | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 h | 2 h | 3 h | 4 h | 6 h | 8 h | 12 h | 16 h | 24 h | 48 h | 72 h | 96 h | 120 h | 144 h | 20 d | 34 d | 62 d |
| Plasma cortisol (ng/mL) | 17.51 | 20.04 | 14.01 | 11.94 | 1.71 | 7.89 | 8.30 | | 36.83 | 2.66 | 5.35 | | | | | | |
| IRT (°C) | | 37.6 | | 38.5 | 37.7 | 38.0 | 36.8 | | 38.2 | 37.5 | 37.3 | | | | | | |
| Substance P (pg/mL) | | 82.42 | | 105.26 | 77.08 | 91.63 | 80.16 | | 77.36 | 98.33 | 91.09 | | | | | | |
| Hair cortisol (pg/mL) | | | | | | | | | | | | | | | | 14.87 | 10.33 |
| Salivary cortisol (ng/mL) | | | | | | | | | | | | | | 1.96 | 2.57 | 0.75 | 2.41 |
| VAS (1–10 cm) | | | | | | | | | 7.5 | 5.5 | 7.5 | 3.0 | 3.5 | 4.5 | | | |

[1]IRT = infrared thermography; VAS = visual analog scale.

[2]Time points that overlapped across multiple studies are reported in the table; some time points not taken on the hour are not reported in the table.

yielded acceptable diagnostic accuracy for surgical castration, dehorning, lameness, and abdominal surgery when comparing no pain to uncontrolled pain are outlined below. Surgical castration and dehorning study results for substance P yielded unacceptable diagnostic accuracy (AUC <0.7; 95% CI: 0.20 to 0.88) at all time points examined (1, 2, 3, 4, 6, 8, 12, 18, 24, 48, 144, 312, 480, 816, 1,152, and 1,488 h) except for dehorning at 72 h (0.70; 95% CI: 0.44 to 0.90), with a cut-off value of 118.80 pg/mL. Lameness study results for substance P yielded acceptable results (AUC >0.7; 95% CI: 0.71 to 0.95) at 8, 24, and 120 h, with cut-off values of 77.56, 84.52, and 82.36 pg/mL, respectively. Abdominal surgery study results for substance P yielded acceptable results (AUC >0.7; 95% CI: 0.45 to 0.91) at 3, 6, and 48 h, with cut-off values of 84.17, 93.78, and 79.79 pg/mL, respectively. Cut-off values for castration, dehorning, lameness, and abdominal surgery are outlined in Tables 8, 9, 10, and 11, respectively.

### Kinematic Gait Analysis

Gait AUC values comparing analgesic use versus uncontrolled pain are outlined in Table 4, with rankings in Table 5. Out of 7 time points, the time points that yielded acceptable diagnostic accuracy for lameness when comparing analgesic use to uncontrolled pain are outlined below. Area yielded acceptable results (AUC >0.7; 95% CI: 0.47 to 0.89) at 24 h, with a cut-off value of 0.02 cm$^2$.

Gait AUC values comparing no pain versus uncontrolled pain are outlined in Table 6, with rankings in Table 7. Out of 7 time points, the time points that yielded acceptable diagnostic accuracy for lameness when comparing no pain to uncontrolled pain are outlined below. Area yielded poor results (AUC <0.7; 95% CI: 0.32 to 0.90) at all time points examined (8, 16, 48, 72, 96, and 120 h) except 24 h (0.73; 95% CI: 0.46 to 0.90), with a cut-off value of 0.16 cm$^2$. Force yielded unacceptable results (AUC <0.7; 95% CI: 0.33 to 0.83) at all time points examined (8, 16, 24, 48, 72, 96, and 120 h). Impulse yielded acceptable results (AUC >0.7; 95% CI: 0.50 to 0.90) at 96 h with a cut-off value of 150.28 kg·s. Pressure yielded unacceptable results (AUC <0.7; 95% CI: 0.29 to 0.72) at all time points examined (8, 16, 24, 48, 72, 96, and 120 h).

### Visual Analog Scale

Visual analog scale AUC values comparing analgesic use versus uncontrolled pain are outlined in Table 4. Out of 6 time points, surgical castration study results comparing analgesic versus uncontrolled pain yielded unacceptable diagnostic accuracy (AUC <0.7; 95% CI:

**Table 9.** Cut-off values for no pain versus pain by collection time point for dehorning

| Biomarker[1] | Time point[2] | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 h | 2 h | 3 h | 4 h | 6 h | 8 h | 12 h | 16 h | 24 h | 48 h | 72 h |
| Plasma cortisol (ng/mL) | 9.82 | 4.96 | 8.64 | 13.59 | 7.83 | 18.14 | 8.03 | | 8.67 | 3.78 | 10.51 |
| IRT (°C) | 38.6 | 37.6 | | 38.5 | 40.5 | | 39.1 | | 37.5 | 37.6 | |
| MNT (kgf) | | | | | 1.05 | | | | 0.89 | 0.78 | |
| Substance P (pg/mL) | 100.07 | 109.81 | | 107.35 | | | 84.75 | | 99.47 | 76.19 | 118.80 |

[1]IRT = infrared thermography; MNT = mechanical nociceptive threshold.

[2]Time points that overlapped across multiple studies are reported in the table; some time points not taken on the hour are not reported in the table.

0.43 to 0.74) for VAS scores at all time points examined (1, 2, 3, 4, 5, and 6 d).

Visual analog scale AUC values comparing no pain versus uncontrolled pain are outlined in Table 6, with rankings in Table 7. Out of 6 time points, surgical castration study results comparing no pain versus uncontrolled pain yielded acceptable diagnostic accuracy (AUC >0.7; 95% CI: 0.65 to 0.81) for VAS scores at 1 and 4 d, with cut-off values of 7.5 and 3.0.

## DISCUSSION

It is essential that noninvasive measures of acute and chronic stress be developed for assessment of animal welfare (Stewart et al., 2005). One avenue is the development of robust biomarkers to objectively quantify pain and evaluate analgesic treatment regimen efficacy during routine elective animal husbandry procedures such as castration and dehorning (Coetzee, 2011). Cortisol is a corticosteroid hormone that is commonly used as an indicator of acute stress responses, as well as pain (Glynn et al., 2013). Cortisol levels are dependent on activation of the hypothalamic–pituitary–adrenal axis, reflect the sensory component of pain, and do not require processing by the central nervous system (Ede et al., 2019). Cortisol responses to painful procedures of the type included in the present analysis are often characterized by a rapid increase in concentration following a procedure that peaks, rapidly declines, and then reaches a plateau; however, many limitations exist, and the stress from restraint and invasiveness may affect cortisol response. Studies have shown that plasma cortisol concentrations are influenced by different procedural methods and have high individual variability (Stafford et al., 2002), and some animals have low responses likely due to higher pain thresholds (Stafford and Mellor, 2005b). Difficulties exist in obtaining true baseline measurements, missing rapid response times, the effect of circadian rhythms on hormone levels, differences in breed and temperament, and the nature of blood sampling, which may itself cause a stress response (Coetzee, 2011). Measures that can be taken to attempt to overcome these challenges include assigning animals to a control treatment that experience the stress of restraint, as was done with "no pain" (handled controls) analysis in the present study, collecting samples at the same time each day to account for circadian rhythms, accounting for time of day in the statistical model, and sampling frequently during the acute phase response. The AUC values included in this analysis indicated that plasma cortisol can yield acceptable diagnostic accuracy (AUC >0.7) for characterizing the sensory component of pain for castration, dehorning, and lameness immediately after the procedure and can continue to yield sound results for multiple days following lameness induction in the studies examined. Plasma cortisol yielded acceptable diagnostic accuracy for identifying uncontrolled pain, as well as detecting NSAID analgesic effects. Salivary cortisol data included in the analysis were collected days, rather than hours, after the surgical castration procedure, which could have been a factor leading to poor diagnostic accuracy. Peak plasma cortisol levels have been found to occur 10 min after a stressor, with peak salivary cortisol values lagging 10 min behind (occurring 20 min after a stressor; Hernandez et al., 2014). It would be beneficial for future analyses for salivary cortisol to be collected closer to the time of the procedure. Hair cortisol data included in the analysis yielded promising results at the 62-d time point but was only measured twice following surgical castration due to the nature of hair growth, so data were limited. The most appropriate method to quantify cortisol via plasma, saliva, or hair may be based on duration of response, invasiveness, and sampling capabilities.

A decrease in eye temperature has been observed following castration in calves (Stewart et al., 2010); however, a lack of analgesic effect could be due to an NSAID alone not effectively controlling pain or may suggest that temperature change is more indicative of stress and less accurate as a sole indicator of pain (Glynn et al., 2013). Stimuli that induce fear and anxiety may trigger similar kinds of physiological responses (Weary et al., 2017) and often accompany painful ex-
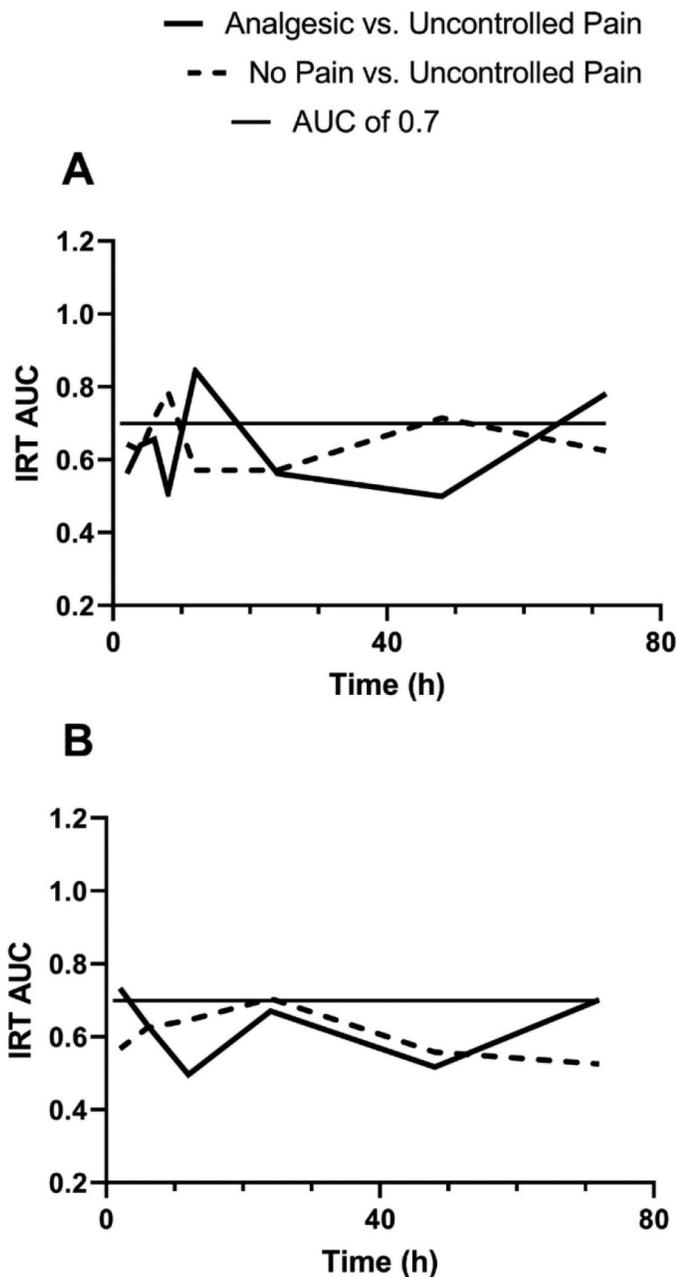
**Figure 3.** Area under the curve (AUC) values for ocular infrared thermography (IRT) temperatures following (A) surgical castration, and (B) abdominal surgery.

periences. The AUC values included in this analysis indicated that IRT may yield acceptable (AUC >0.7) diagnostic accuracy at time points later (castration and surgery: 72 h) than plasma cortisol (castration: 1–12 h) and yield acceptable diagnostic results when identifying pain, as well as analgesic effects.

Determining the MNT via a pressure algometer can establish the minimal amount of pressure that produces a pain response and has been used to evaluate the ef-

**Table 10.** Cut-off values for no pain versus pain by collection time point for lameness

| Biomarker | 1 h | 2 h | 3 h | 4 h | 6 h | 8 h | 12 h | 16 h | 24 h | 48 h | 72 h | 96 h | 120 h |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Plasma cortisol (ng/mL) | 7.55 | 14.72 | 16.66 | 20.48 | 8.83 | 25.55 | | 32.31 | 12.08 | 17.45 | 17.12 | 10.89 | 2.75 |
| Substance P (pg/mL) | 71.98 | 69.23 | | | | 77.56 | | | 84.52 | 66.16 | 76.15 | 75.94 | 82.36 |
| Pressure mat | | | | | | | | | | | | | |
| Area (cm²) | | | | | | 0.02 | | 0.02 | 0.02 | 0.02 | 0.02 | 0.01 | 0.02 |
| Force (kgf) | | | | | | 191.94 | | 149.22 | 162.08 | 173.6 | 250.97 | 248.64 | 174.83 |
| Impulse (kg·s) | | | | | | 120.1 | | 182.81 | 115.09 | 179.63 | 124.56 | 150.28 | 139.57 |
| Pressure (kg/cm²) | | | | | | 4.2 | | 3.7 | 3.6 | 4.1 | 2.8 | 3.8 | 4.0 |

[1]Time points that overlapped across multiple studies are reported in the table; some time points not taken on the hour are not reported in the table.

**Table 11.** Cut-off values for no pain versus pain by collection time point for abdominal surgery

| Biomarker[1] | Time point[2] | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 h | 2 h | 3 h | 4 h | 6 h | 8 h | 12 h | 16 h | 24 h | 48 h | 72 h |
| IRT (°C) | | | 33.7 | | 31.1 | | 31.9 | | 35.3 | 35.4 | 31.2 |
| Substance P (pg/mL) | | | 84.17 | | 93.78 | | 93.40 | | 101.25 | 79.79 | 69.10 |

[1]IRT = infrared thermography.

[2]Time points that overlapped across multiple studies are reported in the table; some time points not taken on the hour are not reported in the table.

fects of analgesia (Heinrich et al., 2009, 2010). Downfalls of MNT determination include a high degree of intra-individual variation, producing an avoidance response, and a lack of a nonpainful control site (Raundal et al., 2014). Interobserver reliability has been shown to increase when MNT values are averaged rather than taking single measurements (Tapper et al., 2013) and averaging of MNT measures was employed for the data included in the present analysis. The AUC values included in this analysis indicated that MNT had better diagnostic accuracy for identifying pain than for analgesic effects due to higher AUC values comparing uncontrolled pain to nonpainful controls rather than for animals administered an analgesic.

Coetzee et al. (2008) suggested that substance P measurement may discriminate between a stressful event, which will cause a cortisol response, and a more specific nociceptive stimulus. Substance P concentrations have been found to be significantly higher in castrated calves compared with controls and in lame cattle with laminitis (Bustamante et al., 2015). Substance P levels may not be altered by NSAID use and may be affected by other factors based on the findings from the present analysis, along with those from Coetzee et al. (2008).

Approval for the use of transdermal flunixin for control of pain due to foot rot was achieved through gait analysis using a floor-based pressure mat system (FDA, 2017). One downfall of the system is that if an animal stops or slows down while walking across the pressure mat, the measurement is often lost (Maertens et al., 2011). The AUC values included in this analysis indicated that kinematic gait analysis may not yield acceptable (AUC >0.7) diagnostic accuracy for experimentally induced lameness at the time points examined. The use of this type of gait analysis in naturally occurring clinical lameness has not been well described in the literature.

Visual analog scale assessment is a method of evaluating pain intensity based on a subjective integration of behavioral parameters. For the data included in the present analysis, the parameters used to assess pain were depression, tail swishing, stance, head carriage, and foot stomping (Martin et al., 2020). Some disadvantages exist when using subjective visual assessment; behavioral observations may lack sensitivity because of individual animal variation or because many behaviors are socially facilitated, and cattle behavior can be influenced by outside factors such as predation, social interactions, and their environment (Kluever et al., 2008). The AUC values included in this analysis indicated that VAS assessment was not affected by NSAID use.

The data in the present analysis were combined from studies that consisted of 3 treatments: (1) cattle experiencing uncontrolled pain, tissue damage, and physiologic changes following a painful procedure; (2) cattle that were controls who were handled and restrained but not subjected to a painful procedure; and (3) cattle who experienced tissue damage and physiologic changes following a painful procedure but were administered an NSAID to mitigate some pain. The data were from studies where only NSAIDs were used for analgesia. Results would likely have differed if a local anesthetic were included in the protocol, as they have been shown to decrease acute pain immediately following the procedure (Winder et al., 2018). However, less than 15% of US producers report always using local anesthesia or analgesia at the time of disbudding in calves <2 mo of age, and less than 25% of producers report always using local anesthesia or analgesia for surgical castration in calves <12 mo of age (Johnstone et al., 2021). Nonsteroidal anti-inflammatory drugs prevent inflammation by inhibiting cyclooxygenase enzymes that produce prostaglandins. Use of an NSAID alone does not effectively reduce the acute stress or pain associated with many elective procedures but provides analgesic and anti-inflammatory effects during the postoperative period (Coetzee, 2011). The mean half-life has been observed to be 6.42 h for topical flunixin, 4.99 h for intravenous flunixin meglumine administration (Kleinhenz et al., 2016), and approximately 26 h for meloxicam (Heinrich et al., 2010). Differences in duration of analgesic effect along with drug absorption may have affected the diagnostic accuracy of analgesia versus pain comparisons, with AUC values likely decreasing as the analgesic effect wore off.

By including the raw data from 7 studies conducted with similar methods using only an NSAID to mitigate pain, our goal was to limit confounders such as different collection and analysis methods, as well as the use of different classes of analgesic compounds with different durations and mechanisms of action. The purpose of the analysis was to evaluate the sensitivity and specificity of a similar group of data across different time points and painful procedures and conditions to preserve the internal validity of the analysis, which limits the scope of this analysis to a specific study type where NSAIDs were used. Animals from the 7 studies were not all the same age group or breed, which could have introduced additional variation to the biomarker results. Due to the limited number of studies included in the ROC analysis, confidence intervals for AUC values and cut-off values varied. Area under the curve values ≥0.7 are reported to have acceptable diagnostic accuracy (Yang and Berdine, 2017) but should be interpreted along with the confidence intervals, cut-off values, and patterns across different collection time points. Thus, AUC values can be valuable for comparing the discriminative ability of a biomarker at one time point compared with another time point, or for comparing the collection of 2 biomarkers at the same time point.

In future ROC analyses, examining the diagnostic accuracy of pain biomarkers following administration of a local anesthetic or combination of systemic analgesia and a local anesthetic would help further characterize the predictive value of pain biomarkers. Examining additional biomarkers along with analgesics with differing durations of action would be beneficial to researchers designing future pain studies. This ROC analysis was a novel approach that did not include all pain biomarkers or collection time points that currently exist in the literature. Additional analyses including a larger number of studies to increase sample size are needed to further validate the diagnostic accuracy of the biomarkers used in the present analysis.

When assessing pain, choosing which biomarkers would be of interest and the appropriate time points for sample collection is reliant upon many factors. First, calves respond to both the pain of the procedure (such as castration or dehorning) and the physical restraint (Faulkner and Weary, 2000). Following a procedure such as dehorning or castration, an acute painful response is observed followed by a period of inflammatory pain (Stock et al., 2013). Acute pain is capable of producing an acute stress response by activating the sympathetic nervous system and secretion of glucocorticoids (Anderson and Muir, 2005). Duration of the stress response for different procedures and conditions likely varies, along with the amount of pain the animal is experiencing at different time points. The effect of analgesia on this response to restore homeostasis depends on whether the analgesia is administered before the procedure, the mechanism of action and duration of the analgesic regimen, and the duration of the stress response. Many of the biomarkers described may be influenced by the animal's physiologic response to the painful procedure, even if pain is mitigated. Finally, biomarker selection that will result in acceptable diagnostic accuracy should consider whether the objective is to identify physiologic effects of a painful procedure or quantify the effects of analgesia.

## CONCLUSIONS

Results from the present study comparing NSAID analgesic effects to uncontrolled pain consistently yielded acceptable diagnostic accuracy for plasma cortisol, hair cortisol, and IRT. The objectives of the study, length of the stress response, and duration of the analgesic regimen administered should all be considered when selecting biomarkers to assess painful procedures and conditions. This analysis was a preliminary approach to identify biomarkers and collection time points that could inform additional analyses using a larger number of studies to further validate the accuracy of these biomarkers. These results indicate that ROC analysis can be used to characterize the predictive value of pain biomarkers, and we have provided new knowledge on the diagnostic accuracy of pain biomarkers within this data set. These results can be used to guide refinement of future research regarding painful procedures and analgesic efficacy.

## ACKNOWLEDGMENTS

## REFERENCES

Alsaaod, M., A. L. Schaefer, W. Büscher, and A. Steiner. 2015. The role of infrared thermography as a non-invasive tool for the detection of lameness in cattle. Sensors (Basel) 15:14513–14525. https://doi.org/10.3390/s150614513.

Amur, S., L. LaVange, I. Zineh, S. Buckman-Garner, and J. Woodcock. 2015. Biomarker qualification: Toward a multiple stakeholder framework for biomarker development, regulatory acceptance, and utilization. Clin. Pharmacol. Ther. 98:34–46. https://doi.org/10.1002/cpt.136.

Anderson, D. E., and W. W. Muir. 2005. Pain management in cattle. Vet. Clin. North Am. Food Anim. Pract. 21:623–635. https://doi.org/10.1016/j.cvfa.2005.07.002.

AVMA. 2019. Castration and Dehorning of Cattle. Accessed August 10, 2021. https://www.avma.org/KB/Policies/Pages/Castration-and-Dehorning-of-Cattle.aspx.

Bustamante, H. A., A. R. Rodríguez, D. E. Herzberg, and M. P. Werner. 2015. Stress and pain response after oligofructose induced-lameness in dairy heifers. J. Vet. Sci. 16:405–411. https://doi.org/10.4142/jvs.2015.16.4.405.

Cardoso, C. S., M. A. von Keyserlingk, and M. J. Hötzel. 2017. Brazilian citizens: Expectations regarding dairy cattle welfare and awareness of contentious practices. Animals (Basel) 7:89. https://doi.org/10.3390/ani7120089.

Coetzee, J. F. 2011. A review of pain assessment techniques and pharmacological approaches to pain relief after bovine castration: Practical implications for cattle production within the United States. Appl. Anim. Behav. Sci. 135:192–213. https://doi.org/10.1016/j.applanim.2011.10.016.

Coetzee, J. F. 2013. A review of analgesic compounds used in food animals in the United States. Vet. Clin. North Am. Food Anim. Pract. 29:11–28. https://doi.org/10.1016/j.cvfa.2012.11.008.

Coetzee, J. F., B. V. Lubbers, S. E. Toerber, R. Gehring, D. U. Thomson, B. J. White, and M. D. Apley. 2008. Plasma concentrations of substance P and cortisol in beef calves after castration or simulated castration. Am. J. Vet. Res. 69:751–762. https://doi.org/10.2460/ajvr.69.6.751.

Ede, T., M. A. von Keyserlingk, and D. M. Weary. 2019. Assessing the affective component of pain, and the efficacy of pain control, using conditioned place aversion in calves. Biol. Lett. 15:20190642. https://doi.org/10.1098/rsbl.2019.0642.

Ekelund, S. 2012. ROC curves—What are they and how are they used? Point Care 11:16–21. https://doi.org/10.1097/POC.0b013e318246a642.

Faulkner, P. M., and D. Weary. 2000. Reducing pain after dehorning in dairy calves. J. Dairy Sci. 83:2037–2041. https://doi.org/10.3168/jds.S0022-0302(00)75084-3.

FDA. 2017. Freedom of Information Summary: Original New Animal Drug Application NADA 141–450 Banamine® Transdermal flunixin transdermal solution. https://animaldrugsatfda.fda.gov/adafda/app/search/public/document/downloadFoi/1944.

Glynn, H. D., J. Coetzee, L. Edwards-Callaway, J. Dockweiler, K. Allen, B. Lubbers, M. Jones, E. Fraccaro, L. Bergamasco, and B. KuKanich. 2013. The pharmacokinetics and effects of meloxicam, gabapentin, and flunixin in postweaning dairy calves following dehorning with local anesthesia. J. Vet. Pharmacol. Ther. 36:550–561. https://doi.org/10.1111/jvp.12042.

Hajian-Tilaki, K. 2013. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. Caspian J. Intern. Med. 4:627–635.

Heinrich, A., T. Duffield, K. Lissemore, and S. Millman. 2010. The effect of meloxicam on behavior and pain sensitivity of dairy calves following cautery dehorning with a local anesthetic. J. Dairy Sci. 93:2450–2457. https://doi.org/10.3168/jds.2009-2813.

Heinrich, A., T. Duffield, K. Lissemore, E. Squires, and S. Millman. 2009. The impact of meloxicam on postsurgical stress associated with cautery dehorning. J. Dairy Sci. 92:540–547. https://doi.org/10.3168/jds.2008-1424.

Hernandez, C. E., T. Thierfelder, K. Svennersten-Sjaunja, C. Berg, A. Orihuela, and L. Lidfors. 2014. Time lag between peak concentrations of plasma and salivary cortisol following a stressful procedure in dairy cattle. Acta Vet. Scand. 56:61. https://doi.org/10.1186/s13028-014-0061-3.

Hudson, C., H. Whay, and J. Huxley. 2008. Recognition and management of pain in cattle. In Pract. 30:126–134. https://doi.org/10.1136/inpract.30.3.126.

Johnstone, E. C., J. F. Coetzee, P. J. Pinedo, and L. Edwards-Callaway. 2021. Current attitudes of veterinarians and producers regarding the use of local and systemic analgesia in beef and dairy cattle in the United States. J. Am. Vet. Med. Assoc. 258:197–209. https://doi.org/10.2460/javma.258.2.197.

Kleinhenz, M. D., P. Gorden, J. Smith, J. Schleining, K. Kleinhenz, J. Juarez, D. Rea, and J. Coetzee. 2019a. Effects of transdermal flunixin meglumine on experimentally induced lameness in adult dairy cattle. J. Dairy Sci. 102:6418–6430. https://doi.org/10.3168/jds.2018-15091.

Kleinhenz, M. D., J. S. Smith, N. Jackson, P. J. Gorden, K. E. Kleinhenz, and J. A. Schleining. 2019b. Evaluation of flunixin meglumine and suture selection on post-operative pain following abdominal surgery in cattle. Page 394 in Proc. Am. Assoc. Bovine Practitioners. https://doi.org/10.21423/aabppro20197295.

Kleinhenz, M. D., N. Van Engen, P. Gorden, J. Ji, P. Walsh, and J. Coetzee. 2017. Effects of transdermal flunixin meglumine on pain biomarkers at dehorning in calves. J. Anim. Sci. 95:1993–2000. https://doi.org/10.2527/jas2016.1138.

Kleinhenz, M., N. Van Engen, P. Gorden, B. KuKanich, S. M. Rajewski, P. Walsh, and J. F. Coetzee. 2016. The pharmacokinetics of transdermal flunixin meglumine in Holstein calves. J. Vet. Pharmacol. Ther. 39:612–615. https://doi.org/10.1111/jvp.12314.

Kleinhenz, M., N. Van Engen, J. Smith, P. Gorden, J. Ji, C. Wang, S. Perkins, and J. Coetzee. 2018. The impact of transdermal flunixin meglumine on biomarkers of pain in calves when administered at the time of surgical castration without local anesthesia. Livest. Sci. 212:1–6. https://doi.org/10.1016/j.livsci.2018.03.016.

Kleinhenz, M. D., P. J. Gorden, M. Burchard, J. A. Ydstie, and J. F. Coetzee. 2019c. Rapid Communication: Use of pressure mat gait analysis in measuring pain following normal parturition in dairy cows. J. Anim. Sci. 97:846–850. https://doi.org/10.1093/jas/sky450.

Kluever, B. M., S. W. Breck, L. D. Howery, P. R. Krausman, and D. L. Bergman. 2008. Vigilance in cattle: The influence of predation, social interactions, and environmental factors. Rangeland Ecol. Manag. 61:321–328. https://doi.org/10.2111/07-087.1.

Maertens, W., J. Vangeyte, J. Baert, A. Jantuan, K. C. Mertens, S. De Campeneere, A. Pluk, G. Opsomer, S. Van Weyenberg, and A. Van Nuffel. 2011. Development of a real time cow gait tracking and analysing tool to assess lameness using a pressure sensitive walkway: The GAITWISE system. Biosyst. Eng. 110:29–39. https://doi.org/10.1016/j.biosystemseng.2011.06.003.

Marti, S., D. M. Meléndez, E. A. Pajor, D. Moya, D. Gellatly, E. D. Janzen, and K. S. Schwartzkopf-Genswein. 2018. Effect of a single dose of subcutaneous meloxicam prior to band or knife castration in 1-wk-old beef calves: II. Inflammatory response and healing. J. Anim. Sci. 96:4136–4148. https://doi.org/10.1093/jas/sky291.

Martin, M., M. Kleinhenz, K. Kleinhenz, E. Reppert, D. Blasi, T. Parks, A. Baysinger, J. Hutcheson, and J. Coetzee. 2020. Comparison of the effect of tildipirosin administered alone or in combination with transdermal flunixin on the performance, health, activity and well-being of transported feedlot calves on arrival at the feedlot. Transl. Anim. Sci. 4:452–459. https://doi.org/10.1093/tas/txaa005.

Meléndez, D. M., S. Marti, E. Pajor, D. Moya, D. Gellatly, E. Janzen, and K. Schwartzkopf-Genswein. 2017. Effect of timing of subcutaneous meloxicam administration on indicators of pain after knife castration of weaned calves. J. Anim. Sci. 95:5218–5229. https://doi.org/10.2527/jas2017.1978.

Meléndez, D. M., S. Marti, E. Pajor, D. Moya, D. Gellatly, E. Janzen, and K. Schwartzkopf-Genswein. 2018. Effect of a single dose of meloxicam prior to band or knife castration in 1-wk-old beef calves: I. Acute pain. J. Anim. Sci. 96:1268–1280. https://doi.org/10.1093/jas/sky034.

Raundal, P. M., P. H. Andersen, N. Toft, B. Forkman, L. Munksgaard, and M. S. Herskin. 2014. Handheld mechanical nociceptive threshold testing in dairy cows–intra-individual variation, inter-observer agreement and variation over time. Vet. Anaesth. Analg. 41:660–669. https://doi.org/10.1111/vaa.12159.

Schulz, K. L., D. E. Anderson, J. F. Coetzee, B. J. White, and M. D. Miesner. 2011. Effect of flunixin meglumine on the amelioration of lameness in dairy steers with amphotericin B-induced transient synovitis-arthritis. Am. J. Vet. Res. 72:1431–1438. https://doi.org/10.2460/ajvr.72.11.1431.

Solvet. 2019. Meloxicam Oral Suspension, USP. https://solvet.ca/wp-content/uploads/2019/11/SOL-Meloxicam-Oral-Suspension-USP-CVP.pdf.

Stafford, K. J., and D. Mellor. 2005a. Dehorning and disbudding distress and its alleviation in calves. Vet. J. 169:337–349. https://doi.org/10.1016/j.tvjl.2004.02.005.

Stafford, K. J., and D. Mellor. 2005b. The welfare significance of the castration of cattle: A review. N. Z. Vet. J. 53:271–278. https://doi.org/10.1080/00480169.2005.36560.

Stafford, K. J., D. Mellor, S. Todd, R. Bruce, and R. Ward. 2002. Effects of local anaesthesia or local anaesthesia plus a non-steroidal anti-inflammatory drug on the acute cortisol response of calves to five different methods of castration. Res. Vet. Sci. 73:61–70. https://doi.org/10.1016/S0034-5288(02)00045-0.

Stewart, M., G. Verkerk, K. Stafford, A. Schaefer, and J. Webster. 2010. Noninvasive assessment of autonomic activity for evaluation of pain in calves, using surgical castration as a model. J. Dairy Sci. 93:3602–3609. https://doi.org/10.3168/jds.2010-3114.

Stewart, M., J. Webster, A. Schaefer, N. Cook, and S. Scott. 2005. Infrared thermography as a non-invasive tool to study animal welfare. Anim. Welf. 14:319–325.

Stock, M. L., S. L. Baldridge, D. Griffin, and J. F. Coetzee. 2013. Bovine dehorning: Assessing pain and providing analgesic management. Vet. Clin. North Am. Food Anim. Pract. 29:103–133. https://doi.org/10.1016/j.cvfa.2012.11.001.

Tapper, K., A. Johnson, L. Karriker, K. Stalder, R. Parsons, C. Wang, and S. Millman. 2013. Pressure algometry and thermal sensitivity for assessing pain sensitivity and effects of flunixin meglumine and sodium salicylate in a transient lameness model in sows. Livest. Sci. 157:245–253. https://doi.org/10.1016/j.livsci.2013.07.017.

Van Engen, N. K., M. L. Stock, T. Engelken, R. Vann, L. W. Wulf, L. A. Karriker, W. Busby, J. Lakritz, A. Carpenter, B. Bradford, W. H. Hsu, C. Wang, and J. F. Coetzee. 2014. Impact of oral meloxicam on circulating physiological biomarkers of stress and inflammation in beef steers after long-distance transportation. J. Anim. Sci. 92:498–510. https://doi.org/10.2527/jas.2013-6857.

Weary, D. M., P. Droege, and V. A. Braithwaite. 2017. Behavioural evidence of felt emotions: Approaches, inferences and refinements. Adv. Stud. Behav. 49:27–48. https://doi.org/10.1016/bs.asb.2017.02.002.

Winder, C. B., C. L. Miltenburg, J. M. Sargeant, S. J. LeBlanc, D. B. Haley, K. D. Lissemore, M. A. Godkin, and T. F. Duffield. 2018. Effects of local anesthetic or systemic analgesia on pain associated with cautery disbudding in calves: A systematic review and meta-analysis. J. Dairy Sci. 101:5411–5427. https://doi.org/10.3168/jds.2017-14092.

Yang, S., and G. Berdine. 2017. The receiver operating characteristic (ROC) curve. Southwest Respir. Crit. Care Chron. 5:34–36. https://doi.org/10.12746/swrccc.v5i19.391.

## ORCIDS

M. Martin ● https://orcid.org/0000-0002-4746-5011

M. D. Kleinhenz ● https://orcid.org/0000-0001-9453-3657

D. Melendez ● https://orcid.org/0000-0002-0153-9708

S. Marti ● https://orcid.org/0000-0002-8115-0582

E. A. Pajor ● https://orcid.org/0000-0003-0747-3997

J. Coetzee ● https://orcid.org/0000-0003-1802-3991