ABSTRACT

The objectives of this systematic review were to investigate the association between nonsteroidal anti-inflammatory drug (NSAID) use during the treatment of claw horn lameness in dairy cattle and locomotion score (LS), nociceptive thresholds, and lying times. A total of 229 studies were initially identified and had their title and abstract screened. From this, we screened the full text of 23 articles, identifying 6 articles for inclusion in the systematic review. Of these 6, 5 reported LS, 2 reported nociceptive thresholds, and 1 reported lying times. The quality of evidence was assessed using a Cochrane risk-of-bias tool and CONSORT items reported for each included study. Due to heterogeneity between the studies, data were reported following Cochrane’s Synthesis without meta-analysis guidelines. Identified heterogeneity between the studies included differences in LS systems and statistical analyses, length of time from enrollment to outcome reported, the NSAID used, concomitant treatments administered, and severity and chronicity of lameness. Recommendations are made with respect to consistency of LS reporting and analysis, along with improvements that may be noted with compulsory reporting guidelines. There were at least some concerns over the risk of bias in 4 of the studies, with risks of bias present in missing outcome data between the study groups. Within the 5 studies included with LS outcomes, there were 22 different pairwise comparisons with either NSAID or NSAID + block as the intervention, with measures of association with presence or absence of lameness as the outcome available for 20 of these comparisons. Animals in the NSAID intervention groups had a lower point estimate lameness risk than animals in the comparison groups in 3 of 8 and 9 of 14 analyses for LS outcomes <10 and ≥10 d post-treatment, respectively. However, there was no difference identified between animals in the NSAID intervention groups compared with the animals in the control group in any of these pairwise comparisons with lameness as the outcome. Twelve pairwise comparisons were reported in the 2 studies with nociceptor threshold as an outcome. Animals in the NSAID intervention groups had a greater nociceptor threshold point estimate compared with animals in the comparison groups in 6 of 6 and 1 of 6 analyses for outcomes <10 and ≥10 d post-treatment, respectively. However, no differences were identified between animals in the NSAID intervention groups and those in the comparison groups. All 4 pairwise comparisons reported in the study with lying times as an outcome found no differences between animals in the NSAID groups and those in the comparison groups. Despite the widespread use of NSAID in the treatment of claw horn lameness, there is a lack of studies of NSAID association with LS, nociceptive thresholds, or lying times. The limited evidence is consistent with no association with NSAID use and those parameters, but comparability across studies was limited by heterogeneity.

Key words: locomotion score, systematic review, claw horn lameness, nonsteroidal anti-inflammatory drug
have an increasing importance in the maintenance of the health and welfare of dairy cattle (Laven, 2020), demonstrated by an increase in NSAID research in cattle over the past decade. From 2010 to 2021, there have been 209 articles published with the search terms “NSAID” and “cattle” (Web of Science search engine, accessed January 19, 2022; https://www.webofscience.com/wos/woscc/basic-search). From 1985 to 2010, there were just 107 articles published with the same search terms. From this crude scientometric approach, it can be concluded that approximately two-thirds of all published NSAID research in the English language in cattle has been carried out over the past 11 yr.

However, it is unclear how much of this recent NSAID research in cattle has focused on the effect of NSAID on CH lameness outcomes, as there are few recent reviews. Review articles on the treatment of lameness have previously been conducted (Hirst et al., 2002; Potterton et al., 2012; Huxley et al., 2014), but these were either not specific for NSAID use or did not follow a systematic review process, and thus may be prone to bias. Nevertheless, those reviews have identified a problematic, historical lack of scientific articles on the effect of NSAID on CH lameness. In a scoping review of lameness treatment and prevention of foot lameness in cattle from 2000 to 2012, only 3 papers were identified on the topic of treatment of CH lameness, and none of these related to NSAID use (Potterton et al., 2012). Given the recent increase in NSAID publications, and the comparative lack of NSAID reviews in lame cattle, there is a need for a systematic review of the current literature on the use of NSAIDs for treating lame dairy cattle.

Another motivating factor for conducting this systematic review was that it appeared that there was positive confirmation bias (where individuals search for evidence to confirm their beliefs; Jones and Sugden, 2001) in the cattle lameness field relative to the use of NSAID. This is particularly clear in the citation of 2 papers from the Dairy Herd Health Group based at Nottingham University, United Kingdom. The first paper, Thomas et al. (2015), reported a positive NSAID response in regard to locomotion score (LS), whereas the second paper (Thomas et al., 2016) reported no such benefit. The former paper has been cited more frequently than the latter (48 vs. 31 citations in Web of Science, respectively), despite it being published only 11 mo earlier and, in the authors’ experience, is much more commonly referred to in the on-farm setting (e.g., Healthy Hoof lameness advisory service, DairyNZ, New Zealand). This discrepancy in citation count is despite the principal difference between the 2 studies being that, compared with cows in the original study, cows in Thomas et al. (2016) were treated slightly later and when lameness was judged to be slightly more severe, a situation that probably reflects on-farm reality more than the situation in Thomas et al. (2015). The question of whether confirmation bias exists, or if there are outlier study results, is best addressed using a systematic review.

Lameness can induce a vast array of changes in cattle, ranging from alterations in trace elements and inflammatory markers (Sun et al., 2015), systemic nociceptive thresholds (Laven et al., 2008), and clinical behavioral changes in gait and behavior (Mainau et al., 2022), to production-limiting effects such as reduced longevity (Huxley, 2013). Although these outcomes are important, it was beyond the scope of this manuscript to review all potential lameness-related changes. Instead, the following 3 outcome measures were selected that the authors believed to have the greatest clinical relevance to lameness response to NSAID treatment: LS, nociceptive threshold, and lying times. The most obvious of these outcome measures is LS, as the definition of lameness involves some alteration in LS (Mainau et al., 2022). If NSAID use had an effect on lameness, the most clinically important effect would be on LS. However, LS is a subjective measure, with low sensitivity for detecting early changes and hoof pain (Dyer et al., 2007). To counter this, a more objective measure was desired. Lame cattle often develop hyperalgesia, an exaggerated sensitivity and perception of pain (Laven et al., 2008). A method for indirectly quantitatively assessing this has been developed for use in lame ruminants (Chambers et al., 1994) by measuring nociceptive threshold in response to a stimulus. Nonsteroidal anti-inflammatory drugs have been widely used in humans and other species for their anti-nociceptive effect (Burian and Geisslinger, 2005); therefore, it is logical to test their efficacy in cattle too. Finally, lameness is known to induce changes in the activity and state of an animal, with respect to lying times (reviewed by Mainau et al., 2022). Thus, lying time may be a useful nongait-related behavioral measure for identifying recovery from lameness.

The aim of this systematic review was to review the literature on NSAID use and its association with 3 key targets of clinical lameness treatment (i.e., LS, nociceptor thresholds, and lying times). Thus, in effect, this systematic review will investigate if these outcome variables can predict whether NSAID treatment has been effective. The systematic review question that the authors were investigating was “Does NSAID use at the time of treatment and diagnosis of CH lameness improve LS, nociceptor thresholds, and lying times in dairy cattle?”
METHODS

The systematic review was designed and conducted according to PRISMA 2020 reporting guidelines for systematic reviews (Page et al., 2021), with veterinary adaptations as recommended by Sargeant and O’Connor (2020). As the PRISMA 2020 statement does not fully cover synthesis of data in the absence of meta-analysis, the reporting of the results was conducted according to synthesis without meta-analysis (SWiM) reporting guidelines (Campbell et al., 2020). The protocol was not independently registered.

Search Terms and Databases

The systematic review question was translated into population, intervention, control group, and outcomes (Sargeant and O’Connor, 2020; Table 1). The search terms used were as follows: (Cattle OR Cow*) AND (Lame* OR mobility OR locomotion) AND (NSAID* OR steroidal OR analgesi*). A search of the full Web of Science database (which consists of CAB Abstracts, Medline, Biological Abstracts and Web of Science Core Collection), and the Scopus database (https://www.scopus.com/) was carried out by WM on October 23, 2021, with the search timelines not restricted. Screening of the reference list of review articles (Hirst et al., 2002; Potterton et al., 2012; Huxley et al., 2014; Coetzee et al., 2017) was carried out by WM after databases had been searched and were included at the eligibility phase in Figure 1 to ensure no key references were missed.

Article Selection

De-duplication was conducted in EndNote (Version X9.2, Thomson Reuters) following the methodology proposed by Bramer et al. (2016), and then manually during the screening process.

After duplicate removal, the remaining citations were exported from EndNote to the systematic review software, Rayyan (Rayyan) and screened on titles and abstract by WM and EC. Both reviewers were blinded to the decision of the other reviewer until both had made a decision (include, exclude, or maybe) on 100% of the articles. The screening criteria for titles and abstract were as follows:

1. The title and abstract were written in English or German.
2. The study involved adult dairy cattle.
3. The predominant lameness lesions were described as CH (white-line, sole hemorrhage, sole ulcer, or a combination of these 3 symptoms).
4. The study was an observational or experimental original research paper.
5. Outcome measures included at least 1 of lameness, locomotion, or mobility scores; algometer or nociceptor threshold; or lying or standing time.
6. The NSAID or response to analgesia was mentioned.

Any conflicts between the reviewers for screening decisions were discussed together, with a collective final decision reached. If any doubt existed for any of the screening criteria based on title and abstract alone, or if an agreed decision could not be reached, then the article was not excluded at the screening stage. All articles that had not been excluded underwent full text recovery. For final inclusion, the studies needed to meet all the following criteria:

1. Full text could be obtained.
2. The full text was in English or German.
3. The study was an observational or experimental original research paper.
4. The population of interest contained dairy cattle with CH lameness.
5. Outcome measures included at least 1 of lameness, locomotion, or mobility scores; algometer or nociceptor threshold; or lying or standing time.
6. NSAID were used as a treatment intervention or cohort group.
7. The control group consisted of dairy cattle with CH lameness and did not receive NSAID. Control group could receive concomitant treatments.

Table 1. Population, intervention, control group, and outcome descriptions for the aim and scope of the systematic review

| Criterion       | Description                                                                 |
|-----------------|-----------------------------------------------------------------------------|
| Population      | Clinically lame dairy cattle suffering from claw horn lameness               |
| Intervention    | Nonsteroidal anti-inflammatory drug (NSAID) use at the time of treatment and diagnosis of lame cow, could receive concomitant treatments |
| Control group   | Did not receive NSAID, but could receive other concurrent treatments         |
| Outcome         | Locomotion score, nociceptive threshold, and lying times                     |
These articles were then reviewed by 3 reviewers (WM, EC, and KM), with the exception of German full text articles that were screened by KM only (a native German speaker). Each reviewer was blinded to the other reviewers’ selections until all articles had been scored. Any conflicts between the 3 reviewers were discussed together, with a final tie-breaker decision carried out by WM where necessary.

**Data Extraction, Synthesis, and Reporting**

It was decided a priori that the included articles would not be amenable to meta-analysis, due to distinct heterogeneity between studies with respect to study outcomes. For example, LS reported on different scales (0–3, 1–5, lame vs. not lame, sound vs. not-sound), different time frames (from hours after NSAID treat-
mment to 100 d post-NSAID treatment), varying study population (chronic vs. acute), difference between types and proportion of CH lameness, and varying interventions (e.g., types of NSAID used and concomitant treatments in both treatment and control groups). As a result, synthesis of data from included articles were reported following the methods proposed by Campbell et al. (2020). The 9 SWiM reporting categories are detailed in Table 2, with the approaches used in this study reported.

Along with study primary author, year of publication, country, and whether or not the article was peer reviewed, we extracted data on the intervention and comparison group, the NSAID used, and outcomes measured. Data were grouped into the outcome measurement and length of time from NSAID administration to outcome measurement, and the raw data and statistical outcomes extracted from the text. For articles with LS as an outcome, an attempt was made to synthesize the results as lame or nonlame. If sufficient data were not present in the text to categorize LS, then the corresponding authors were contacted and a request was made for the raw data, or in one instance, data were obtained from a PhD thesis (O’Callaghan, 2003). If the raw data were obtained, the primary author transformed the LS into lame or not (based on the LS system used), and the proportion lame at each period was reported. Contrast analyses were carried out on each combination of intervention versus comparison group in each study. For example, Thomas et al. (2015) contained 4 treatment groups (NSAID, NSAID + block, block, and no treatment). The contrast analyses for this included NSAID versus block, NSAID versus no treatment, NSAID + block versus block, and NSAID + block versus no treatment. If the contrast analyses could not be obtained directly from the manuscript, then Monte-Carlo random simulation was carried out using the normally disturbed characteristics of regression coefficient and standard errors. In the case of Laven et al. (2008), the raw data were reanalyzed following the methodology proposed in Laven et al. (2008). No adjustments for multiple comparisons were used. Further detail on the data extraction for each study is presented in Table 2.

In addition to the above SWiM guidelines, each included article was assessed for risk of bias and for reproducibility and adherence to protocol reporting guidelines for clinical trials. Risk of bias was assessed by WM for each outcome of the included articles according to the methods proposed by Cochrane Risk of Bias tool V2.0 (Sterne et al., 2019). Although these were conducted for each outcome, as recommended by Sterne et al. (2019), the risk of bias for studies that reported more than 1 relevant outcome was the same across all of their outcomes. Thus, an overall risk of bias for each study was presented, rather than for each outcome. A Microsoft Excel macro tool was used to collate the responses for each of the 5 potential bias domains (recovered October 5, 2021, from https://www.riskofbias.info/welcome/rob-2-0-tool/current-version-of-rob-2).

We reviewed all of the included studies for the presence or absence of the 25 items in the CONSORT reporting guidelines checklist (Moher et al., 2010). These were reported for each article, and an overall proportion of articles that included the CONSORT items were also reported.

RESULTS

The PRISMA 2020 flow diagram for the systematic review is presented in Figure 1. A total of 299 articles were identified by the search terms in the databases, 72 of which were identified as duplicates. Thus, 227 articles had their titles and abstracts screened, with 204 excluded at this stage as they did not meet the title and abstract screening criteria. Full text screening was carried out on the remaining 23 articles, with 6 articles identified as meeting all inclusion criteria and included in the systematic review. All 6 articles included in the systematic review were in the English language.

Of the 6 studies, 5 were carried out in the United Kingdom, and 1 was conducted in New Zealand. They were all published after 2000. In all of the studies, the control animals received, at a minimum, a therapeutic trim of the lameness-causing CH lesion. Substantial heterogeneity existed between the included studies, and included duration of lameness (<2 wk duration or ≥2 wk or unknown duration), NSAID investigated (ketoprofen or tolfenamic acid), concomitant treatments administered in animals in both the intervention and control groups (the addition of a hoof block or no further treatments), severity of lameness or lesion, and time from treatment to when outcome was measured (for LS, this ranged from several hours after NSAID administration to 100 d after NSAID administration). Further details of the studies are presented in Table 3.

Lameness Outcomes

A total of 5 of the 6 articles that met the inclusion criteria reported LS outcomes, and these consisted of 22 different contrast analyses with either NSAID or NSAID + block as the intervention. Measures of association with lame or not lame as the outcome were available for 20 of these analyses. None of these analyses identified a difference between groups, with the 95% confidence interval for odds ratio overlapping 1 in all analyses.
Table 2. Synthesis without meta-analysis (SWiM) reporting items and descriptions of the methodology of how each item was addressed; SWiM is intended to complement and be used as an extension to PRISMA (Campbell et al., 2020)

| SWiM reporting item | Description |
|---------------------|-------------|
| Methods             | The studies were first grouped according to their outcome variable, either locomotion scores (LS), nociceptive threshold measurements, or lying times. Studies were then grouped into intervention comparison groups for subgroup synthesis. For example, if a study reported the findings of nonsteroidal anti-inflammatory drug (NSAID) and non-NSAID groups with and without foot block, then each pairwise comparison was reported and grouped separately. Finally, studies were grouped from subgroup synthesis based on the length of time after NSAID treatment for which the outcome variable was reported (<10 or ≥10 d). |
| 1. Grouping studies for synthesis | |
| 2. Describe the standardized metric and transformation methods used | The LS were reported using different systems and scales (i.e., 4 different scales in the 5 articles that reported LS). Due to this, an attempt was made to transform them from the ordinal score to a lame versus not lame binary variable, as defined by each of the outcome scales. If this was not possible, then the data were reported as was in the article. Nociceptive threshold measurements were reported in different units, kPa or N. These were not transformed to a consistent measure (N), as doing this depended on external factors that could not be controlled. Nociceptive threshold data were reported in the original units used in the respective articles. |
| 3. Describe the synthesis methods | An effect size with 95% CI and a direction of association was reported for each pairwise comparison from each included study. Whether the 95% CI overlapped with 1 for binary outcomes, or 0 for continuous outcomes, for each analysis was reported. The direction of the association between the intervention and the comparison group was also reported, regardless of 95% CI. |
| 4. Criteria used to prioritize results for summary and synthesis | As it was expected that there would not be many included articles, all were reported equally. However, those with some concern or high risk of bias were explicitly commented upon. |
| 5. Investigation of heterogeneity in reported effects | There were many sources of heterogeneity and likely too few studies to investigate further. Rather, the important sources of heterogeneity identified were discussed. |
| 6. Certainty of evidence | An overall risk measure, with confidence intervals, was not reported. The proportion of analyses in a given direction of association was reported, with a greater proportion of associations in a given direction likely to be more certain of an association. However, no certainty measurement (confidence interval or probability statistic) could be assigned to this value. The overall risk of bias and the proportion of the studies that reported each of the CONSORT items were discussed, as they had bearing on the certainty of the evidence. |
| 7. Data presentation methods | All results were reported in tabular format, apart from risk of bias, which was presented in graphical format for each domain, and overall risk. The tables included a column for the citation, study type, NSAID and dosing regimen, intervention group, control group, outcome measurement, time period, results, significance of association, direction of association, overall risk of bias, and comments. |
| Results             | As it was expected that outcomes would be in various forms, results were reported in a table with commentary, rather than a forest plot. A PRISMA diagram of the exclusion reasons were presented, as was a description of all the included studies in tabular format. |
| Discussion          | Limitations included, but were not limited to, variation in the methods used for LS, the study population, concomitant treatments, and follow-up or losses to the study. These could be addressed in the methods and results, but could not be accounted for (e.g., in a meta-analysis); therefore, an overall metric was not produced. |
### Table 3

Descriptions of the 6 studies included for the systematic review, including whether the article was peer reviewed or not, country where the work was conducted, study population investigated, the various intervention groups for pairwise comparisons, outcomes and time frames of outcomes, nonsteroidal anti-inflammatory drug (NSAID) used, and the total number enrolled and analyzed.

| Reference                  | Number | Peer reviewed | Country         | Study population                                                                                                                                  | Interventions                                                                                     | Outcomes                                                                                       | NSAID                                                                                      | Total number enrolled (number analyzed if reported) |
|----------------------------|--------|---------------|-----------------|---------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------|--------------------------------------------------|
| Laven et al. (2008)        | 1      | Yes           | New Zealand     | Lame dairy cattle identified by farmer; included if claw horn lameness only, could be lame in more than 1 limb                                       | Six treatment groups, split depending on whether animal required plastic shoe: did not require plastic shoe; (1) NSAID, (2) NSAID + plastic shoe, (3) plastic shoe, (4) no additional treatment; required plastic shoe: (5) NSAID + plastic shoe, (6) plastic shoe | Locomotion score (LS): repeated at d 3, 8, 28, and 100; Nociceptive threshold: repeated at d 3, 8, 28, and 100 | Tolfenamic acid (2 mg/kg BW) single treatment immediately after corrective trim                | 149 (not reported)                              |
| Miguel-Pacheco et al. (2016) | 2      | Yes           | UK              | Dairy cattle with acute (<2 wk duration) lameness identified by trained LS observers; claw horn lameness only included, lame in only 1 limb         | Four treatment groups: (1) no additional treatment, (2) foot block, (3) NSAID, (4) NSAID + foot block | Lying times for 5 days post-treatment                                                                 | Ketoprofen (3 mg/kg BW) SID\(^1\) for 3 d, immediately after corrective trim                  | 44 (40)                                        |
| O’Callaghan-Lowe et al. (2004) | 3      | No            | UK              | Dairy cattle with acute (<2 wk duration) lameness identified by trained LS observers; non-claw horn lameness and claw horn lameness included, exclusion criteria not stated | Three treatment groups for claw horn lameness group: (1) NSAID, (2) foot block, (3) no additional treatment | LS: repeated between 3 and 7 d, 16–20 d and 26–30 d                                                                                      | Ketoprofen (3 mg/kg BW) SID for 3 d, immediately after corrective trim               | 118                                             |
| Thomas et al. (2015)       | 4      | Yes           | UK              | Dairy cattle with acute (<2 wk duration) lameness identified by trained LS observers; claw horn lameness only included, lame in only 1 limb       | Four treatment groups: (1) no additional treatment, (2) foot block, (3) NSAID, (4) NSAID + foot block | LS at 35 d post-treatment                                                                                                                      | Ketoprofen (3 mg/kg BW) SID for 3 d, immediately after corrective trim                  | 183 (167)                                      |
| Thomas et al. (2016)       | 5      | Yes           | UK              | Dairy cattle with chronic (>2 wk duration) lameness identified by trained LS observers; claw horn lameness only included, lame in only 1 limb | Three treatment groups: (1) no additional treatment, (2) foot block, (3) NSAID + foot block      | LS at 42 d post-treatment                                                                                                                    | Ketoprofen (3 mg/kg BW) SID for 3 d, immediately after corrective trim                  | 189 (154)                                      |
| Whay et al. (2005)         | 6      | Yes           | UK              | Dairy cattle with lameness of unknown duration as identified by trained LS observers; infectious and claw horn lameness included, lame in only 1 limb | Two treatment groups: (1) NSAID, (2) sterile saline placebo                                         | LS: repeated at d 1, 3, 8, and 28; nociceptive threshold: repeated at d 1, 3, 8, and 28   | Ketoprofen (3 mg/kg BW) SID for 3 d, first dose 1 h before corrective trim                        | 28 claw horn lameness; 40 total enrolled        |

\(^1\)SID = once daily.
The length of time from treatment to LS measurement varied between studies, with 3 reporting LS outcomes <10 d post-treatment, and all 5 reporting LS ≥10 d post-treatment. Of those LS outcomes ≥10 d post-treatment, animals in the intervention groups had a lower point estimate lameness risk than animals in the comparison groups in 9 of 14 analyses. For the LS outcomes <10 d, only 3 of 8 analyses of contrasts had a lower point estimate lameness risk in favor of animals in the intervention groups compared with the comparison groups.

**Nociceptor Threshold and Lying Times**

Two of the 6 articles that met the inclusion criteria reported nociceptive threshold outcomes (Table 5). No differences were identified in the 12 contrast analyses (either NSAID or NSAID + block as the intervention) from the 2 articles, with the 95% confidence intervals overlapping 0 in all of them. For nociceptive threshold <10 d post-treatment, animals in the intervention groups had a greater nociceptor threshold point estimate in 6 of 6 analyses compared with animals in the comparison groups. For nociceptor thresholds collected ≥10 post-treatment, only 1 of 6 analyses had a greater nociceptor threshold for animals in the intervention group compared with the comparison group.

Only 1 article that met the inclusion criteria reported lying times. All 4 lying time contrast analyses in Miguel-Pacheco et al. (2016) had a point estimate with NSAID-treated animals lying for a longer time per day than animals in the comparison groups, although all of the 95% confidence intervals between intervention groups and comparison groups included 0. The largest increase in lying times was noted in animals treated with NSAID + block compared with block.

**Bias and Reporting Guidelines**

The overall risk of bias varied between studies from low risk to high risk (Figure 2). At least a moderate risk of bias was identified in missing outcome data in 4 of the 6 studies, with a high risk of bias present in O’Callaghan-Lowe et al. (2004). Some risk of bias was identified in Whay et al. (2005) in the randomization process as this was the only study that enrolled animals before diagnosing the lameness-causing lesion. This resulted in differences in the proportion of lesions between intervention groups. Two of the 6 studies had bias concerns with measurement of the outcome, with O’Callaghan-Lowe et al. (2004) only measuring the total number of animals at each LS, not the individual animal scores.

There was a range in the CONSORT compliance between each item in the checklist (Table 6). Perfect compliance for all 6 studies was achieved in the background and objectives, eligibility criteria, and outcome definitions. The most noncompliant CONSORT items included sample size (2 of 6 reporting), the randomization process (2 of 6 reporting), baseline data (2 of 6 reporting), and discussion of limitations (0 of 6 reporting).

**DISCUSSION**

This is the first systematic review investigating the published literature on the use of NSAID for the treatment of CH lameness in dairy cattle. From an initial 227 abstracts and titles screened from the search terms, 23 were progressed to full text screening. Only 6 of these articles met the criteria for inclusion in the review, highlighting the scarcity of evidence in this field. This is despite the recent increase in papers published on NSAID and cattle. This scarcity of evidence, combined with the between-study heterogeneity, means that we lack the data to properly evaluate the effect of NSAID use on LS, nociceptive threshold or lying times, or the factors that influence the response to NSAID use.

The validity and repeatability of research can be assessed with multiple methods. The authors chose the following 2 methods for this review: the Cochrane risk-of-bias tool designed to assess bias in randomized trials (Sterne et al., 2019) and the CONSORT reporting guidelines (Moher et al., 2010), presented as the proportion of included articles that reported each specific CONSORT item. The quality of evidence varied between the 6 studies, from low risk of bias to high risk of bias. A major difference between the 2 studies that had a low risk of bias (Thomas et al., 2015, 2016) compared with the other 4 studies was the use of reporting guidelines. This highlights the scientific benefits of following reporting guidelines to minimize risk of bias in a study and, crucially, to assist with repeatability of the research. The authors commend the *Journal of Dairy Science* on leading the movement to require the use of reporting guidelines.

One of the more challenging aspects of conducting cattle lameness intervention trials is accounting for losses to follow-up. There are many reasons why missing data for the outcome may be related to the treatment, thus creating bias. Two important reasons are that animals may require re-treatment due to welfare reasons (lameness becoming worse or no improvement over time), or if using blocks, they fall off. These reasons clearly could be related to the treatment group. The risk of bias for the 5 LS studies for missing outcome data ranged from
Table 4. Summary of data extracted from the 5 studies included that measured locomotion score (LS) as an outcome [pairwise comparisons were carried out for each intervention and control group, and based on whether LS was collected <10 d from nonsteroidal anti-inflammatory drug (NSAID) administration or ≥10 d from NSAID administration; the odds ratio (OR; 95% CI) and direction of association was reported for each pairwise comparison and an overall risk of bias for each study presented]

| Time frame | Reference                  | Outcome                                                                 | Intervention                  | Comparison group | Raw data                                                                 | Measure of association (OR with 95% CI) | CI include 1 | Direction of association | Risk of bias | Comments                                                                 |
|------------|----------------------------|-------------------------------------------------------------------------|-------------------------------|------------------|--------------------------------------------------------------------------|---------------------------------------|--------------|-------------------------|-------------|--------------------------------------------------------------------------|
| <10 d      | Laven et al. (2008)       | Lame (LS ≥3/5) or not lame (LS <3/5)                                    | NSAID                         | No treatment     | d 3: 0.50 (12/24) vs. 0.35 (6/17) d 8: 0.21 (5/24) vs. 0.24 (4/17)    | 1.32 (0.33–5.20)                      | Y            | Negative                | Some concerns | Raw data re-analyzed to lame or not; OR and 95% unadjusted               |
|            |                            |                                                                         | NSAID                         | Block            | d 3: 0.50 (12/24) vs. 0.50 (9/18) d 8: 0.21 (5/24) vs. 0.18 (3/17)    | 1.05 (0.28–3.98)                      | Y            | Negative                |              |                                                                          |
|            |                            |                                                                         | NSAID and block               | No treatment     | d 3: 0.47 (9/19) vs. 0.35 (6/17) d 8: 0.33 (7/21) vs. 0.24 (4/17)    | 1.6 (0.30–6.57)                       | Y            | Negative                |              |                                                                          |
|            |                            |                                                                         | NSAID and block               | Block            | d 3: 0.47 (9/19) vs. 0.50 (9/18) d 8: 0.33 (7/21) vs. 0.18 (3/17)    | 1.28 (0.20–3.08)                      | Y            | Negative                |              |                                                                          |
|            |                            |                                                                         | Lesion required a block       | Block            | d 3: 0.37 (13/35) vs. 0.33 d 8: 0.20 (7/35) vs. 0.16 (5/31)          | 1.20 (0.39–3.66)                      | Y            | Negative                |              |                                                                          |
| ≥10 d      | Whay et al. (2005)        | Median and quartiles of LS (1–5)                                        | NSAID                         | Sterile saline   | d 1: 3 (3–4) vs. 3.5 (3–5) d 3: 2 (1–3.25) vs. 2.5 (1.75–4) d 8: 2 (1–3) vs. 1 (1–2.25) | NA                                      | NA           | Positive                | Some concerns | Could not access raw data                                               |
|            | O’Callaghan-Lowe et al. (2004) | Lame (LS ≥3/5) or not lame (LS <3/5)                                    | NSAID                         | No treatment     | d 3: 0.7 vs. 0.7 (8/11) vs. 0.7 (9.2/13) d 8: 0.7 vs. 0.5 (5.9/11) | 0.69 (0.3–1.5)                        | Y            | Positive                | Large number of missing animals; data obtained from PhD                |
|            | Thomas et al. (2015)      | Lame (LS ≥2/3) or not lame (LS <2/3)                                    | NSAID and block               | Block            | d 28: 0.11 (2/18) vs. 0.00 (0/15) d 100: 0.00 (1/18) vs. 0.00 (0/13) | 1.0 (0.06–13.6)                       | Y            | Positive                | Low OR and CI could not be calculated, as the control group had zero lame animals |
|            | Thomas et al. (2016)      | Lame (LS ≥2/3) or not lame (LS <2/3)                                    | NSAID and block               | Block            | d 28: 0.13 (2/15) vs. 0.00 (0/15) d 100: 0.00 (0/13) vs. 0.00 (0/13) | 1.0 (0.06–13.6)                       | Y            | Negative                | Low OR and CI calculated from adjusted model data                     |
|            | Whay et al. (2005)        | Median and quartiles of LS (1–5)                                        | NSAID                         | Sterile saline   | d 1: 3 (3–4) vs. 3.5 (3–5) d 3: 2 (1–3.25) vs. 2.5 (1.75–4) d 8: 2 (1–3) vs. 1 (1–2.25) | NA                                      | NA           | Positive                | Some concerns | Could not access raw data                                               |
Some (Whay et al., 2005; Laven et al., 2008), to high concern (O’Callaghan-Lowe et al., 2004). It is of note that the risk of bias was greatest in the article that had not been through journal peer review (O’Callaghan-Lowe et al., 2004).

However, even in the low-risk group, there was still a large number of animals that were lost to follow-up. Loss rates were similar across the intervention groups in those 2 trials, but an argument could be made that the study design could still result in losses to follow-up that could bias the results. One way to reduce the risk of this type of bias would be to analyze the data as time to event (i.e., time to soundness or time to nonlame). If cattle are required to be re-treated, then the animal can be censored at the time of re-treatment; therefore, not all of the animal time-at-risk would be lost to the analysis. Although a potentially more challenging study to conduct, the authors recommend that future studies using LS as an outcome use survival analysis methods. This will also increase the power of the study and remove the a priori effect of defining a time period that an animal must be sound by, which has little biological relevance to the cow or farmer.

One of the major challenges of comparing studies that report LS as the primary outcome is the variability in the LS systems and scales that are used by different research groups and countries. This was apparent in this review, with 4 different LS systems used in the 5 studies that reported LS outcomes. This issue was exacerbated by all 5 studies using different statistical methods and LS outcomes.

The use of different LS systems has been a consistent issue limiting sharing, comparing, and use of lameness data ever since the development of LS (Schlageter-Tello et al., 2014). Different systems use different criteria for scoring, and simplistic correspondence tables can obscure those differences. For example, in the International Committee for Animal Recording (ICAR) guidelines (ICAR, 2020), they claim that a score of 3 on the 0-to-5 Sprecher scale (ICAR, 2020) is equivalent to a score of 5 on the First-Step system when the criteria for an AHDB score 3 (unable to walk as fast as a brisk human pace and lame leg easy to identify) also match the criteria for a First-Step score of 4 (gait is best described as 1 deliberate step at a time and ability to move freely is obviously diminished). Similar issues with BCS prompted Roche et al. (2004) to examine relationships across differing BCS systems; however, as far as the authors are aware, no such studies have been undertaken for LS. Although the ICAR guidelines (ICAR, 2020) recommend the use of the 1-to-5 Sprecher score, it is unlikely that there will be agreement to use only 1 LS in all countries and all systems. For example, the First-Step system has been...
Table 5. Summary of data extracted from the 2 studies included that measured nociceptor threshold and the 1 study that measure lying times<sup>1</sup>

| Variable | Time frame | Reference | Intervention | Comparison | Outcome | Measure of association | CI overlap | Direction of association | Risk of bias | Comments |
|----------|------------|-----------|--------------|------------|---------|------------------------|------------|------------------------|-------------|---------|
| Nociceptor threshold | <10 d | Laven et al. (2008) | NSAID | No treatment | d 3: 27.1 (1.68) vs. 27.4 (1.72) | 0.76 (−4.93 to 6.45) | Y<sup>2</sup> | Positive | Some concerns | Point estimate and CI from estimated marginal means from adjusted linear regression model; units kPa |
| | | | NSAID | Block | d 3: 27.1 (1.68) vs. 25.3 (2.27) | 2.35 (−2.99 to 7.68) | Y | Positive | |
| | | | NSAID + block | No treatment | d 3: 25.8 (1.92) vs. 27.4 (1.72) | 1.89 (−3.79 to 7.57) | Y | Positive | |
| | | | NSAID + block | Block | d 8: 31.9 (1.31) vs. 28.8 (1.79) | 0.30 (−5.68 to 6.29) | Y | Positive | |
| | | Lesion required a block; NSAID + block | Block | d 8: 31.1 (1.38) vs. 31.3 (1.88) | 1.19 (−3.40 to 5.78) | Y | Positive | |
| | | Whay et al. (2004) | NSAID | Sterile saline | d 1: 9.1 (0.54) vs. 9.1 (0.6) | 0.00 (−1.63 to 1.63)<sup>1</sup> | Y | Positive | Some concerns | Units N; Contained 12 of 40 animals with nonclaw horn lameness |
| | | | NSAID | No treatment | d 3: 10 (0.61) vs. 8.8 (0.57) | 1.20 (−0.49 to 2.89) | Y | Positive | Some concerns |
| | | | NSAID | Block | d 8: 10.3 (0.61) vs. 9.0 (0.42) | 1.30 (−0.20 to 2.80) | Y | Positive | Some concerns |
| | | | NSAID + block | No treatment | d 28: 40.5 (1.75) vs. 42.8 (1.93) | −2.32 (−8.55 to 3.91) | Y | Negative | Some concerns |
| | | | NSAID + block | Block | d 100: 43.7 (1.36) vs. 45.4 (1.44) | −1.86 (−7.70 to 3.98) | Y | Negative | |
| | | | NSAID + block | No treatment | d 28: 38.1 (2.05) vs. 42.8 (1.93) | −1.99 (−8.10 to 4.12) | Y | Negative | |
| | | | NSAID + block | Block | d 100: 45.4 (1.3) vs. 45.4 (1.44) | −2.45 (−8.91 to 4.01) | Y | Negative | |
| | | Lesion required a block; NSAID + block | Block | d 28: 36.6 (1.41) vs. 40.9 (1.28) | −2.99 (−7.36 to 1.38) | Y | Negative | |
| | ≥10 d | Laven et al. (2004) | NSAID | Sterile saline | d 3: 10.6 (0.73) vs. 9.2 (0.53) | 1.40 (−0.42 to 3.22) | Y | Positive | Some concerns | Units N; Contained 12 of 40 animals with nonclaw horn lameness |

<sup>1</sup> Units N; Contained 12 of 40 animals with nonclaw horn lameness
criticized for undue focus on back arching (H. R. Whay, National University of Ireland, Galway, Ireland, personal communication) and in pasture-based herds such as those in New Zealand and Australia. In these types of herds, locomotion scoring takes place after milking (Fabian et al. 2014) with a system that depends on observing whether or not there is a back arch at standing, and walking is of limited value. Furthermore, observing for back arch during walking and standing is incompatible with large-scale, long-term studies, which are required to conduct high quality, adequately powered interventional studies. Thus, recommendations to standardize LS in interventional studies, as have been made for reporting reproductive indices (Lean et al., 2016), are unlikely to be followed. However, almost all manual LS systems have a binary cutpoint for lame versus not lame or sound versus not-sound, and it would be extremely valuable to compare those cutpoints across systems to establish a level of agreement.

A pleasing finding was that all articles included in our review had clear objectives, and eligibility criteria and outcome definitions were well defined. These factors have been reported as lacking in a previous veterinary meta-analysis (Muir et al., 2017). In contrast, sample size calculations were only reported for 2 of 6 of studies used in our review. Although this is a nonissue in Thomas et al. (2015), as a difference between groups was identified (NSAID + block associated with greater proportion of cattle with LS < 1 than block or no other treatments), the findings of the other 5 studies cannot be interpreted without information on the power or a discussion on the practical interpretation of not rejecting the null hypothesis. Only 1 of the 6 articles we included in our review (Thomas et al., 2016) discussed type 2 errors and the implications to the inferences of the study. This makes extrapolation of the individual study results difficult. The effect of some of this possible misclassification bias was reduced through the methods used in this systematic review by synthesizing the direction of the association. However, the true association between NSAID and LS, nociceptor thresholds, and lying times remains unknown. This problem of lack of transparency of sample size and power is not only related to cattle lameness research, and it remains an issue within other fields of veterinary science (Muir et al., 2017). The introduction of compulsory reporting guidelines in journals such as the Journal of Dairy Science should go a long way to address some of these issues.

Two of the original articles reported significant differences between NSAID intervention groups and comparison groups in lame cattle (Whay et al., 2005; Thomas et al., 2015). The authors have no reason to believe that the association identified in Thomas et al.
(2015) does not exist. Although the lower bounds of the 95% confidence interval do overlap an odds ratio of 1 when analyzed as lame or not in this current systematic review (LS >1 or LS ≤1 on a 0–3 scale), the study effect size was large for NSAID + block compared with no treatment (odds ratio: 0.31, 95% confidence interval: 0.09–1.11). One of the major differences in the study design of Thomas et al. (2015) compared with the other studies is that they enrolled animals that had been lame for less than 14 d (based on fortnightly LS). However, we cannot ignore the findings of the other studies when considering the effect of NSAID. This is particularly relevant; even though rapid identification and treatment of lameness is strongly recommended (Pedersen and Wilson, 2021), this is unlikely to be what happens in the majority of lameness management on-farm (Leach et al., 2010; Alawneh et al., 2012; Fabian et al., 2014).

The improvement in nociceptor threshold over time in animals treated with ketoprofen reported by Whay et al. (2005) has previously been questioned by Laven et al. (2008), who surmised that the difference between the studies was in the control group, rather than the treatment group. This was confirmed by Whay et al. (2005), who, in the same article, also reported no difference between nociceptor threshold between treatment groups at each measurement day. This was also concluded in the current review from pairwise comparisons of the data presented by Whay et al. (2005). Comparing nociceptive threshold within treatment groups, and then claiming a difference because the comparison in one treatment group was significant but not in the other is a flawed approach (Sainani, 2010). Laven et al. (2008) did not identify any significant difference in the nociceptive threshold between animals treated with NSAID and those not treated with NSAID. Thus, no strong evidence currently exists that NSAID increase nociceptor threshold over time in lame dairy cattle compared with claw trimming alone.

Although there have been other studies investigating the effects of NSAID on lying times in lame cattle (reviewed by Mainau et al., 2022), this current review only identified 1 study that explicitly enrolled cattle with CH lameness (Miguel-Pacheco et al., 2016). The majority of studies investigating lying times and activity measurements in lame cattle compared with nonlame control animals highlight an increase in lying times and a reduction in activity in general (Mainau et al., 2022). Thus, for NSAID to have an effect, it would be expected that there should be a reduction in lying time in lame animals treated with NSAID compared with lame animals not treated with NSAID. However, this has not been consistently found (Mainau et al., 2022). The 1 study of lying times and NSAID use included in the current systematic review (Miguel-Pacheco et al., 2016) did not identify any significant differences in lying times in animals that received NSAID compared with those that did not receive NSAID, although the point estimate for all pairwise comparisons was in favor of an increase in lying times for animals in the NSAID groups, rather than the expected reduction. Therefore, there is no strong evidence currently that NSAID treatment has any effect on lying times in animals with CH lameness.

A critique of the search terms used for this review could be that they were too narrow. An informal sensitivity analysis was conducted to assess this, where “OR Bovine” and “OR meloxicam OR ketoprofen OR
Table 6. The presence or absence of each CONSORT (Moher et al., 2010) checklist item for each of the 6 included studies, as well as the proportion of the 6 studies included that contain each item; the included article number is cross-referenced in Table 3.

| Section and topic     | Item no | Checklist item                                                                 | Included article number | Percentage with item (number of studies with relevant item) |
|-----------------------|---------|---------------------------------------------------------------------------------|-------------------------|----------------------------------------------------------|
| Title and abstract    | 1a      | Identification as a randomized trial in the title                              | N                       | N N N Y Y Y N 33% (2/6)                                  |
|                       | 1b      | Structured summary of trial design, methods, results, and conclusions           | Y                       | Y Y Y Y Y Y 100% (6/6)                                  |
| Introduction          | 2a      | Scientific background and explanation of rationale                             | Y Y Y Y Y Y 100% (6/6)  |
|                       | 2b      | Specific objectives or hypotheses                                               | Y Y N Y Y Y 83% (5/6)   |
| Methods               | 3a      | Description of trial design (such as parallel, factorial) including allocation  | N                       | N Y N Y Y Y 66% (4/6)                                  |
|                       | 3b      | Important changes to methods after trial commencement (such as eligibility    | NA                      | NA NA NA NA NA NA 0% (0/6)                              |
|                       |         | criteria), with reasons                                                        |                         |                                                          |
| Participants          | 4a      | Eligibility criteria for participants                                           | Y Y Y Y Y Y 100% (6/6)  |
|                       | 4b      | Settings and locations where the data were collected                           | Y Y Y Y Y Y 100% (6/6)  |
| Interventions         | 5       | The interventions for each group, with sufficient details to allow replication,| N                       | N N N Y Y Y 50% (3/6)                                  |
|                       |         | including how and when they were administered                                  |                         |                                                          |
| Outcomes              | 6a      | Completely defined prespecified primary and secondary outcome measures,        | Y Y Y Y Y Y 100% (6/6)  |
|                       | 6b      | including how and when they were assessed                                      | N                       | N N N N N N 0% (0/6)                                  |
| Sample size           | 7a      | How sample size was determined                                                  | N                       | N N N Y Y Y 33% (2/6)                                  |
|                       | 7b      | When applicable, explanation of any interim analyses and stopping guidelines   | NA                      | NA NA NA NA NA NA 0% (0/6)                              |
| Randomization         | 8a      | Method used to generate the random allocation sequence                          | Y Y N Y Y Y N 83% (5/6)  |
| Sequence generation   | 8b      | Type of randomization; details of any restriction (such as blocking and block  | N                       | N Y N Y Y Y 50% (3/6)                                  |
| Allocation concealment| 9       | Mechanism used to implement the random allocation sequence (such as sequentially| N                       | N Y N Y Y Y 50% (3/6)                                  |
| mechanism             |         | numbered containers), describing any steps taken to conceal the sequence until  |                         |                                                          |
|                       |         | interventions were assigned                                                    |                         |                                                          |
| Implementation        | 10      | Who generated the random allocation sequence, who enrolled participants, and    | N                       | N Y N Y N N 33% (2/6)                                  |
|                       |         | who assigned participants to interventions                                      |                         |                                                          |
| Blinding              | 11a     | If done, who was blinded after assignment to interventions (for example,      | N                       | N N N N N Y 17% (1/6)                                  |
|                       |         | participants, care providers, those assessing outcomes and how                  |                         |                                                          |
| Statistical methods   | 11b     | If relevant, description of the similarity of interventions                     | NA                      | NA NA NA NA NA NA 0% (0/6)                              |
|                       | 12a     | Statistical methods used to compare groups for primary and secondary outcomes   | Y                      | Y Y N Y N Y N 66% (4/6)                                  |
|                       | 12b     | Methods for additional analyses, such as subgroup analyses and adjusted         | Y                      | NA N Y NA N 50% (2/4)                                    |
| Results               | 13a     | For each group, the numbers of participants who were randomly assigned,        | N                       | Y N N Y Y Y N 50% (3/6)                                  |
|                       |         | received intended treatment, and were analyzed for the primary outcome         |                         |                                                          |
|                        | 13b     | For each group, losses and exclusions after randomization, together with       | N                       | Y N N Y Y Y N 50% (3/6)                                  |
|                       |         | reasons                                                                        |                         |                                                          |
| Recruitment           | 14a     | Dates defining the periods of recruitment and follow-up                         | Y                       | Y Y Y Y Y Y 100% (6/6)                                  |
|                       | 14b     | Why the trial ended or was stopped                                             | NA                      | NA NA NA NA NA NA 0% (0/6)                              |

Continued
Table 6 (Continued). The presence or absence of each CONSORT (Moher et al., 2010) checklist item for each of the 6 included studies, as well as the proportion of the 6 studies included that contain each item; the included article number is cross-referenced in Table 31

| Section and topic | Item no | Checklist item                                                                 | Included article number | Percentage with item (number of studies with relevant item) |
|-------------------|---------|--------------------------------------------------------------------------------|-------------------------|-----------------------------------------------------------|
| Baseline data     | 15      | A table showing baseline demographic and clinical characteristics for each group | Y N N Y N N             | 33% (2/6)                                                 |
| Numbers analyzed  | 16      | For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups | N Y N Y Y N             | 50% (3/6)                                                 |
| Outcomes and estimation | 17a   | For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval) | N Y Y Y N N             | 50% (3/6)                                                 |
|                   | 17b     | For binary outcomes, presentation of both absolute and relative effect sizes is recommended | N NA NA Y Y N           | 50% (2/4)                                                 |
| Ancillary analyses | 18      | Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory | Y NA NA NA NA Y         | 100% (2/2)                                                |
| Harms             | 19      | All important harms or unintended effects in each group (for specific guidance see CONSORT for harms) | N N N Y Y N             | 33% (2/6)                                                 |
| Discussion        | 20      | Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses | N N N N N N             | 0% (0/6)                                                  |
| Limitations       | 21      | Generalizability (external validity, applicability) of the trial findings         | Y Y N Y Y Y             | 83% (5/6)                                                 |
| Generalizability  | 22      | Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence | Y Y Y Y Y Y             | 83% (5/6)                                                 |
| Interpretation    |         |                                                                                   |                         |                                                           |
| Other information | Registration | Registration number and name of trial registry                                    | N N N N N N             | 0% (0/6)                                                  |
| Protocol          | 24      | Where the full trial protocol can be accessed, if available                         | NA NA NA NA NA NA       |                                                           |
| Funding           | 25      | Sources of funding and other support (such as supply of drugs), role of funders    | Y Y Y Y Y Y             | 100% (6/6)                                                |

1NA = not available; Y = yes; N = no.
“flunixin” were added to the search term and produced 30 further records. All 30 of these articles would have been excluded at the title and abstract screening phase. Furthermore, no new articles were identified from the review articles that had not already been identified through database searching, giving confidence that the search terms had a high sensitivity for identifying relevant articles. It needs to be acknowledged that the literature on NSAID use for cattle lameness is more extensive than what is presented here. The authors were principally interested in studies that mimic “real-life” on-farm situations, and on the common, more chronic lameness conditions (i.e., CH lameness). As such, any mention of acute infectious lameness (e.g., footrot or digital dermatitis), or experimentally induced lameness, were not addressed in this review. Although these studies may have identified benefits from using NSAID, this review did not attempt to address all pain-mitigating associations that NSAID have with distal limb pain.

The NSAID regimen was the same in 5 of the articles (ketoprofen at 3 mg/kg BW daily for 3 d), with only 1 other regimen used (tolfenamic acid at 2 mg/kg once; Laven et al., 2008). This potentially reduced the external validity as we did not have knowledge on whether other NSAID compounds, or various dosing regimens of ketoprofen or tolfenamic acid, would have resulted in a different outcome. The distinct lack of studies involving meloxicam, the worldwide market-leading NSAID product with respect to volume sales (Metcam; CEESA – Executive animal health study center; company data, Boehringer-Ingelheim, Auckland, New Zealand) is of note and further highlights the lack of research in this field. However, given the duration of action of all known NSAID that are used in cattle, and the timeframes of the outcomes reported, we believe this unlikely to bias the inferences.

Some positive trends have been seen worldwide in the use of NSAID in production animals, with an increase in usage being reported (Laven, 2020). One possible unseen consequence of this trend is the confirmation bias from more welfare-conscious animal health advisors and farmers. Nevertheless, from the small number of studies conducted in this area, there was no definitive evidence from this systematic review that the use of NSAID improved LS and nociceptive threshold over an extended period of time post-treatment. This is particularly true in the clinical setting, whereby the majority of cattle have been lame for an extended period of time before being diagnosed and treated.

Nonetheless, the authors strongly support the use of NSAID in even the mildest clinical lame case, as the anti-inflammatory effects are likely to have benefits during the acute and chronic phases (Stock and Coetzee, 2015). Furthermore, the act of trimming lame cows is a painful act (Chapinal et al., 2010) and should warrant NSAID treatment. If the limited published data do reflect the lack of effect of NSAID on the LS, nociceptive threshold, and lying times of lame cows, it is because these are insensitive measures of assessing whether NSAID use is beneficial for the cow rather than accurate assessments of NSAID efficacy.

**ACKNOWLEDGMENTS**

Financial assistance was received from Kirk Family Postgraduate Scholarship through Massey University, Palmerston North, New Zealand. We thank the staff at VetEnt (Waikato, New Zealand) for their support during the production of this manuscript, and the fellow lameness researchers and enthusiasts that inspired this systematic review. The authors have not stated any conflicts of interest.

**REFERENCES**

Alawneh, J. I., M. A. Stevenson, N. B. Williamson, N. Lopez-Villalobos, and T. Otley. 2012. The effect of clinical lameness on live-weight in a seasonally calving, pasture-fed dairy herd. J. Dairy Sci. 95:663–669. https://doi.org/10.3168/jds.2011-4505.

Bramer, W. M., D. Giustini, G. B. de Jonge, L. Holland, and T. Bekhuis. 2016. De-duplication of database searches results for systematic reviews in EndNote. J. Med. Libr. Assoc. 104:240–243. https://doi.org/10.3163/1536-0506.104.3.014.

Burian, M., and G. Geisslinger. 2005. COX-dependent mechanisms involved in the antinociceptive action of NSAIDs at central and peripheral sites. Pharmacol. Ther. 107:139–154. https://doi.org/10.1016/j.pharmthera.2005.02.004.

Campbell, M., J. E. McKenzie, A. Sowden, S. V. Katiukreddi, S. E. Brennan, S. Ellis, J. Hartmann-Boyce, R. Ryan, S. Shepperd, J. Thomas, V. Welch, and H. Thomson. 2020. Synthesis without meta-analysis (SWiM) in systematic reviews: Reporting guidelines. BMJ 368:l6890. https://doi.org/10.1136/bmj.l6890.

Chambers, J. P., A. E. Waterman, and A. Livingston. 1994. Further development of equipment to measure nociceptive thresholds in large animals. Vet. Anaesth. Analg. 21:66–72. https://doi.org/10.1111/1467-2959.1994.tb00489.x.

Chapinal, N., A. M. de Passillé, and J. Rushen. 2010. Correlated changes in behavioral indicators of lameness in dairy cows following hoof trimming. J. Dairy Sci. 93:5758–5763. https://doi.org/10.3168/jds.2010-3426.

Coetzee, J. F., J. K. Shearer, M. L. Stock, M. D. Kleinhenz, and S. R. van Amstel. 2017. An update on the assessment and management of pain associated with lameness in cattle. Vet. Clin. North Am. Food Anim. Pract. 33:389–411. https://doi.org/10.1016/j.cvfa.2017.02.009.

Dyer, R. M., N. K. Neerchal, U. Tasch, Y. Wu, P. Dyer, and P. G. Rajkondawar. 2007. Objective determination of claw pain and its relationship to limb locomotion score in dairy cattle. J. Dairy Sci. 90:4592–4602. https://doi.org/10.3168/jds.2007-0006.

Fabian, J., R. A. Laven, and H. R. Whay. 2014. The prevalence of lameness on New Zealand dairy farms: A comparison of farmer estimate and locomotion scoring. Vet. J. 201:31–38. https://doi.org/10.1016/j.tvjl.2014.05.011.

Hirst, W. M., A. M. Le Fevre, D. N. Logue, J. E. Offer, S. J. Chapin, R. D. Murray, W. R. Ward, and N. P. French. 2002. A systematic compilation and classification of the literature on lameness in cattle. Vet. J. 164:7–19. https://doi.org/10.1053/tvjl.2002.0729.
Huxley, J. N. 2013. Impact of lameness and claw lesions in cows on health and production. Livest. Sci. 156:64-70. https://doi.org/10.1016/j.livsci.2013.06.012.

Huxley, J. N., S. C. Archer, O. C. D. Atkinson, N. J. Bell, E. A. Berry, N. Bollard, R. S. Dean, C. Mason, O. Maxwell, P. G. Miguel, S. L. Potterton, J. Remnant, P. Sleeman, H. J. Thomas, and H. R. Whay. 2014. An update on the treatment of claw horn lesions in cattle. Pract. Vet. 22:189–192.

ICAR. 2020. Section 7- Guidelines for Lameness in Bovine. Accessed between 2000 and 2011. Vet. J. 193:612–616. https://doi.org/10.1016/j.tvjl.2012.06.040.

Roche, J. R., P. G. Dillon, C. R. Stockdale, L. H. Baumgard, and M. J. VanBaale. 2004. Relationships among international body condition scoring systems. J. Dairy Sci. 87:3076–3079. https://doi.org/10.3168/jds.s0022-0302(04)73441-4.

Sainani, K. 2010. Misleading comparisons: The fallacy of comparing statistical significance. PLoS Med. 7:559–562. https://doi.org/10.1016/j.panj.2010.04.016.

Sargeant, J. M., and A. M. O’Connor. 2020. Scoping reviews, systematic reviews, and meta-analysis: Applications in veterinary medicine. Front. Vet. Sci. 7:111. https://doi.org/10.3389/fvets.2020.00101.

Schlageter-Tello, A., E. A. M. Bokkers, P. W. G. Koerkamp, T. Van Hertem, S. Viazzii, C. E. B. Romanini, I. Halachmi, C. Bahr, D. Berckmans, and K. Lokhorst. 2014. Manual and automatic locomotion scoring systems in dairy cows: A review. Prev. Vet. Med. 116:12–25. https://doi.org/10.1016/j.prevetmed.2014.06.006.

Sterne, J. A. C., J. Savović, M. J. Page, R. G. Elbers, N. S. Blencowe, I. Boutron, C. J. Cates, H.-Y. Cheng, M. S. Corbett, S. M. Elrod, J. R. Emerson, M. A. Hernán, S. Hopewell, A. Hrobjartsson, D. R. Junqueira, P. Juni, J. J. Kirkham, T. Lasserson, T. Li, A. McAleenan, B. C. Reeves, S. Shepperd, I. Shrier, L. A. Stewart, K. Tilling, I. R. White, P. F. Whiting, and J. P. T. Higgins. 2019. RoB 2: A revised tool for assessing risk of bias in randomised trials. BMJ 366:l4898. https://doi.org/10.1136/bmj.l4898.

Stock, M. L., and J. F. Coetzee. 2015. Clinical pharmacology of analgesic drugs in cattle. Vet. Clin. North Am. Food Anim. Pract. 31:113-138. https://doi.org/10.1016/j.cvfa.2014.11.002.

Sun, D., C. Li, C. Gu, J. Chen, Y. Qu, X. Wang, J. Gao, S. Wei, J. Wang, R. Wu, and D. Guo. 2015. Analysis of mineral elements, metabolism, and inflammation indexes in the plasma of dairy cows suffering from different degrees of lameness. Biol. Trace Elem. Res. 168:372–379. https://doi.org/10.1007/s12011-015-0380-x.

Thomas, H. J., G. G. Miguel-Pacheco, N. J. Bollard, S. C. Archer, N. J. Bell, C. Mason, O. J. R. Maxwell, J. G. Remnant, P. Sleeman, H. R. Whay, and J. N. Huxley. 2015. Evaluation of treatments for claw horn lesions in dairy cows in a randomized controlled trial. J. Dairy Sci. 98:4477–4486. https://doi.org/10.3362/jds.2014-8982.

Thomas, H. J., J. G. Remnant, N. J. Bollard, A. Burrows, H. R. Whay, N. J. Bell, C. Mason, and J. N. Huxley. 2016. Recovery of chronically lame dairy cows following treatment for claw horn lesions: A randomised controlled trial. Vet. Rec. 178:116. https://doi.org/10.1136/vr.103394.

Whay, H. R., A. J. F. Webster, and A. E. Waterman-Pearson. 2005. Role of ketoprofen in the modulation of hyperalgesia associated with lameness in dairy cattle. Vet. Rec. 157:729–733. https://doi.org/10.1136/vr.157.25.729.