Use of antimicrobials in the treatment of calf diarrhea: a systematic review

C. Bernal-Córdoba¹², R. Branco-Lopes¹², L. Latorre-Segura¹, M. de Barros-Abreu¹, E. D. Fausak³ and N. Silva-del-Río¹²

¹Veterinary Medicine Teaching and Research Center, 18830 Road 112, Tulare, CA 93274, USA; ²Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis, Davis, CA 95616, USA and ³Carlson Health Science Library, University of California, Davis, Davis, CA 95616, USA

Abstract

The objective of this study was to conduct a systematic review of the scientific literature evaluating the efficacy and comparative efficacy of antimicrobials (AMs) for the treatment of diarrhea in calves. Eligible studies were non- and randomized controlled trials evaluating an AM intervention against a positive and negative control, with at least one of the following outcomes: fecal consistency score, fever, dehydration, appetite, attitude, weight gain, and mortality. Four electronic databases were searched. Titles and abstracts (three reviewers) and full texts (two reviewers) were screened. A total of 2899 studies were retrieved; 11 studies met the inclusion criteria. The risk of bias was assessed. Most studies had incomplete reporting of trial design and results. Eight studies compared AMs to a negative control (placebo or no treatment). Among eligible studies, the most common outcomes reported were diarrhea severity (n = 6) and mortality (n = 6). Eligible studies evaluated very different interventions and outcomes; thus, a meta-analysis was not performed. The risk of bias assessment revealed concerns with reporting of key trial features, including disease and outcome definitions. Insufficient evidence is available in the scientific literature to assess the efficacy of AMs in treating calf diarrhea.

Introduction

Rationale

Gastrointestinal disorders are one of the most prevalent diseases of preweaned dairy calves: approximately 21% of dairy calves in US operations are affected and 76% of them receive antimicrobial (AM) treatment (NAHMS-USDA, 2018; Urie et al., 2018). Similarly, diarrhea is the primary reason for AM treatment in beef calf ranches (Waldner et al., 2013). The primary goal of AM therapy in diarrheic calves is to prevent bacteremia and decrease the number of coliform bacteria in the small intestine (Smith, 2015). However, experts recommended that AM treatments should be limited to scouring calves showing clinical signs of systemic illness (Constable et al., 2008).

In the USA, there is a limited number of AMs with a Food and Drug Administration (FDA) approval for the treatment of gastrointestinal diseases in calves (clarithromycin, ampicillin, amoxicillin, oxetacetracycline, tetracycline, and sulfamethazine; FARAD, 2020). Most of the FDA-approved AM drugs belong to the penicillin or tetracycline class, categorized as critically and highly important AMs for human medicine, respectively (WHO, 2019). Although AMs are widely used for prophylaxis, metaphylaxis, and treatment of infectious diseases in calves (Urie et al., 2018), validated evidence on the efficacy of AMs for the treatment of gastrointestinal disorders in calves is lacking (Smith, 2015).

AM use represents a threat to worldwide public health, as it is one of the main drivers of the emergence of antimicrobial resistance (AMR; Van Boeckel et al., 2015; WHO, 2015; FDA CVM, 2018). Therefore, the judicious use of medically important AM drugs in food-producing animals has been proposed as a key strategy to preserve the effectiveness of currently available AM drugs (WHO, 2015; FDA CVM, 2018; OIE, 2018). Accurate and unbiased evidence on the therapeutic efficacy of AMs to treat infectious diseases is necessary to successfully design evidence-based AM stewardship programs (Sargeant et al., 2019a).

The efficacy of AM treatments should be assessed in multi-arm randomized controlled trials (RCTs), but these are rarely available in the scientific literature. So, research synthesis methods of two-armed RCTs can be used to evaluate AM efficacy (O’Connor et al., 2019). Systematic reviews (SRs) and meta-analyses (MA) are powerful tools that can provide scientifically valid information on the scope and conclusions of the existing literature on AM treatments for calf diarrhea. These synthesis methods are needed to design evidence-based...
decision-making guidelines that can be incorporated in AM stewardship programs for livestock.

Objectives

The first objective of this study was to conduct an SR to appraise the scientific literature on the efficacy and comparative efficacy of AM treatments for diarrhea in calves under 6 months of age. The second objective was to conduct an MA to evaluate the efficacy of AM drugs compared to the absence of treatment, alternative non-AM treatments, or other AM drugs used to treat diarrhea in calves under 6 months of age.

Methods

Protocol and registration

An a priori review protocol was developed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA-P; Moher et al., 2015) and was archived in the University of California eScholarship online repository (https://escholarship.org/uc/item/0nw528h4). In addition, the protocol was published on the Systematic Reviews for Animals and Food (SYREAF) website (http://www.syreaf.org/protocol). Protocol amendments are described below and include screening questions, risk of bias assessment, and summary measures.

Eligibility criteria

Eligibility criteria, search strategy, and screening questions were designed based on the PICOS question format (Population–Intervention–Comparison–Outcome–Study type; EFSA, 2010; O’Connor et al., 2014a). The population of interest was dairy and beef calves under 6 months of age at the time of study enrollment. The intervention of interest was the administration of oral or injectable AMs (antimicrobials; antibiotics and antiprotozoal drugs) after observing clinical signs of diarrhea or after exposing animals to a diarrhea-causing pathogen (challenge studies). Comparisons of interest were the absence of treatment (e.g. placebo, no-treatment), alternative non-AM treatments (e.g. herbal extracts, probiotics, lactoferrin, oral rehydration solutions), or other AM treatments (e.g. AM used as positive control). Outcomes of interest were limited to mortality, health [e.g. fecal consistency score (FCS), blood in feces, dehydration (DH), appetite, demeanor, or fever], and performance [e.g. average daily gain (ADG) and feed efficiency]. Only studies assessing the efficacy of AMs to treat animals diagnosed with diarrhea based on clinical signs were relevant. Studies exclusively focusing on pathogen fecal shedding were excluded. The SR was limited to primary research including non-, quasi-, and RCTs with at least one AM and one comparator group. Only peer-reviewed publications were retrieved, and ‘gray literature’ (literature not formally published, such as theses and dissertations, conference proceedings, trade articles, research reports, and policy documents) was not included (Dickersin et al., 1994). Eligible studies had to be written in English and publicly available, although not necessarily open access. The searching period was based on database coverage, and no limit on publication date was applied.

Search strategy and information sources

The search strategy was designed by an experienced health and veterinary science academic librarian (E. D. F.), with input and reference citation lists from content experts (C. B. C. and N. S. R.). Relevant articles were identified by the principal investigator (C. B. C.) and keywords and indexing terms were mined through Medline (via PubMed, 1966–2020) and CAB Abstracts (via CAB Direct, 1972–2020). After developing the search strategies in CAB Abstracts and PubMed, the search was translated by E. D. F. to Scopus (via Scopus, 1970–2020) and Biosis (via Web of Science, 1926–2020). Keywords from relevant references were gathered and compared to keywords utilized in the previous search. Yale MeSH Analyzer (http://mesh.med.yale.edu/) was also utilized to compare common Medical Subject Headings across articles. Content experts identified keywords or indexing terms based on key pathogens and relevant AMs. During screening, C. B. C. performed a hand-search of relevant manuscripts and reviews using the snowball method and citation searching (https://libguides.library.uu.nl/PiL). The literature search was conducted from 1st to 2nd July 2019, and a search update was made on 29th June 2020. All studies were exported to Mendeley (Mendeley Ltd., Elsevier), where duplicate citations were deleted. The search strategy used for all databases is described in Supplementary material (SM 1).

Selection process

Covidence SR management software (Veritas Health Innovation, Melbourne, Australia) was used to manage the screening of the title and abstract of all citations retrieved in the search. Three reviewers with veterinary and animal science backgrounds were trained on PICOS format questions prior to screening the title (C. B. C., L. L. S., and M. B. A.), abstract (C. B. C. and L. L. S.), and full text (C. B. C. and L. L. S.). The screening questions included in the protocol were beta-tested with 40 citations, and afterward modified for clarity if needed. For title and abstract screening questions, the possible answers were ‘no’, ‘maybe/unclear’, and ‘yes’. References moved to the next stage if all title and abstract screening questions were answered ‘yes’ or ‘maybe/unclear’. For full-text screening questions, the possible answers were ‘no’ and ‘yes’. References were included in the SR if all the full-text screening questions were answered ‘yes’. References were excluded if all reviewers responded ‘no’ to one or more questions. Disagreements on manuscript inclusion were resolved by consensus and if necessary, an additional researcher (N. S. R.) was consulted. The final screening questions utilized were:

Title screening
(1) Does the title indicate cattle as the subject of study?
(2) Does the title describe the use of an AM treatment?

Abstract screening
(1) Does the abstract describe a controlled trial?
(2) Does the abstract describe a study of diarrhea in calves?
(3) Does the abstract describe one or more intervention groups of an AM treatment regimen?
(4) Does the abstract report at least one outcome related to clinical cure or performance?
Full-text screening
In this final screening level, the previous six questions and the following questions were used:

1. Is the enrollment age of subject cattle ≤6 months?
2. Are AM given after the diagnosis of diarrhea or the onset of clinical signs?
3. Does the study evaluate clinical outcomes of AM treatment?

Studies evaluating the efficacy of AM use in control (metaphylaxis) and prevention (prophylaxis) of disease, as defined by the American Veterinary Medical Association (AVMA, 2020), were excluded. Studies where AM treatments were given as growth promoters, and studies with unclear or no reporting primary data were not considered. The reasons for manuscript exclusion were recorded at this level.

Data collection process
The data extraction process was completed following the guidelines by Sargeant and O’Connor (2014). Two reviewers (C. B. C. and L. L. S.) independently used pre-designed spreadsheets to collect data (Excel 2010, Microsoft Corp., Redmond, WA). Data extraction disagreements were resolved by discussion until a consensus was reached; if needed, a third reviewer (N. S. R.) was consulted. Study-level data included population, interventions, comparators, and outcomes for each independent study. Population data included: breed, sex, enrollment age, housing, inclusion criteria, and sample size. Intervention and comparator-level data were extracted and included: randomization process, group size, treatment features (active ingredient, dose, route, length, and frequency), complementary treatments (e.g., fluid therapy and anti-inflammatory drugs), and features of personnel who delivered treatments (e.g., training or blindness). Additionally, pathogens (e.g., genus and species) and infection type (e.g., challenge study or natural infection) were extracted. Outcome data extracted included: type, evaluation features (e.g., assessment methods, evaluation period, and frequency of measurements), and features of personnel who assessed clinical outcomes (e.g., training or blindness). Treatment failure and success definitions, when these were available, were extracted – without modifications – from the original manuscripts. The summary effects of the outcomes were extracted from either adjusted (if available) or unadjusted data as well as their corresponding measures of variability. Moreover, the significance and variability of the reported outcome were recorded when available [e.g., standard deviation, standard error, odds ratio, relative risk, confidence intervals (CIs), and P-value].

Risk of bias assessment
The risk of bias at the outcome level was independently assessed by three reviewers (C. B. C., R. B. L., and L. L. S.) using the Cochrane Risk of Bias Tool for Randomized Trials (Sterne et al., 2019). Five commonly used domains (bias arising from the randomization process, bias due to deviations from intended interventions, bias due to missing outcome data, bias in the measurement of the outcome, and bias in the selection of reported results) and a novel domain (bias related to disease definition; SM 2) were assessed. As described below, signaling questions were modified following the approach described by Sargeant et al. (2019a, 2019b) in prior livestock synthesis studies. In the randomization process domain, the question ‘was the allocation sequence random?’ was modified to ‘was the study randomized?’. The answers to this question were modified to ‘probably no’ if the study did not report data on sequence generation, ‘probably yes’ if the study reported random sequence allocation but not the randomization process, and ‘yes’ when the study reported a random component in the sequence generation process (e.g., computer random number generator). Also, the allocation sequence concealment question was not included as it is unlikely that a farmworker, producer, or researcher would have a treatment preference for any given calf. In the domain regarding deviations from intended interventions, the question ‘were participants aware of their assigned intervention during the trial?’ was always answered as ‘no’, as the ‘participants’ in all studies were calves. This domain also inquires about the blinding of study personnel; for the purposes of this SR, the animal caregivers and/or people responsible for delivering treatment were the relevant study personnel. The risk of bias tool was tested with three studies to ensure consistency across reviewers (Sargeant and O’Connor, 2014). Reviewers were trained on the risk of bias tool, and disagreements between reviewers were resolved by consensus to adjudicate the final judgment. The outcome chosen for bias assessment was the severity of FCS or diarrhea, but if not reported, diarrhea duration was used instead.

Synthesis of results
As described in the study protocol, the goal of this SR was to conduct an MA to evaluate the efficacy of AMs in the treatment of calf diarrhea. Our SR identified few eligible manuscripts; there was wide variability in interventions and outcomes across studies. Scarcity of the scientific literature and heterogeneity among studies made it unfeasible to address the review question. Thus, no quantitative synthesis could be performed, and heterogeneity was not formally assessed. Following the PRISMA guidelines, study results were summarized in forest plots for visualization purposes.

Summary measures
The effect size [risk ratio (RR) or mean difference] was calculated for the most common outcomes reported at the group level: diarrhea (or fecal score) severity and calf mortality. For categorical data, mean difference and 95% CIs were calculated using the OpenEpi online tool (https://www.openepi.com/Mean/t_testMean.htm); pooled standard error reported in each original manuscript was used in these calculations. For binary data, the RR and 95% CI were calculated using MedCalc Statistical Software version 20.0.5 (MedCalc Software, Ostend, Belgium). For RR calculation, calves that received the intervention were considered exposed, and calves that were given the comparator were considered unexposed. For FCS or diarrhea, RR was calculated using mild to severe diarrhea cases in exposed and unexposed calves. Post-hoc analysis was not necessary when manuscripts reported the effect size as RR.

Results
Study characteristics
Results below correspond to SR results only, as an MA could not be conducted due to scarcity of studies and differences in interventions and outcomes among selected studies. The search retrieved 2899 publications from which 102 full-text manuscripts were assessed for eligibility (Fig. 1). In total, 11 manuscripts
Fig. 1. Flow diagram illustrating the selection of eligible studies for the systematic review of the efficacy of antimicrobials in the treatment of calf diarrhea (Adapted from the PRISMA guidelines). *102 full-text manuscripts which contained 126 independent studies.

Table 1. Characteristics of eligible trials investigating the efficacy of AMs in the treatment of calf diarrhea

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Housing (farms)</th>
<th>n</th>
<th>Agea</th>
<th>Breed type</th>
<th>Main etiological agent</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Challenge</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bywater (1977)</td>
<td>England</td>
<td>NR</td>
<td>42</td>
<td>5–10 days</td>
<td>NR</td>
<td><em>E. coli</em></td>
</tr>
<tr>
<td>Fecteau et al. (2003)</td>
<td>The United States</td>
<td>University/research</td>
<td>29</td>
<td>At birth</td>
<td>Dairy</td>
<td><em>Salmonella typhimurium</em></td>
</tr>
<tr>
<td>Loftstedt et al. (1996)</td>
<td>Canada</td>
<td>University/research</td>
<td>30</td>
<td>≤6 h</td>
<td>Dairy</td>
<td><em>E. coli</em></td>
</tr>
<tr>
<td>Ollivett et al. (2009)</td>
<td>The United States</td>
<td>University/research</td>
<td>23</td>
<td>At birth</td>
<td>Dairy</td>
<td><em>Cryptosporidium parvum</em></td>
</tr>
<tr>
<td>Schnyder et al. (2009)</td>
<td>Switzerland</td>
<td>University/research</td>
<td>6</td>
<td>1–3 days</td>
<td>NR</td>
<td><em>C. parvum</em></td>
</tr>
<tr>
<td>Silva et al. (2010)</td>
<td>Brazil</td>
<td>NR</td>
<td>12</td>
<td>10–15 days</td>
<td>NR</td>
<td><em>Salmonella Dublin</em></td>
</tr>
<tr>
<td>White et al. (1998)</td>
<td>England</td>
<td>NR</td>
<td>38</td>
<td>1–2 w</td>
<td>NR</td>
<td><em>E. coli</em></td>
</tr>
<tr>
<td><strong>Natural infection</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grandemange et al. (2002)</td>
<td>France and Belgium</td>
<td>Commercial</td>
<td>184</td>
<td>≤5 days</td>
<td>Beef</td>
<td><em>E. coli</em></td>
</tr>
<tr>
<td>Grimshaw et al. (1987)</td>
<td>England, France, and Germany</td>
<td>University/research</td>
<td>452</td>
<td>3–10 days</td>
<td>Dairy</td>
<td><em>E. coli</em></td>
</tr>
<tr>
<td>Sheldon (1997)</td>
<td>United Kingdom</td>
<td>Commercial</td>
<td>65</td>
<td>1–30 days</td>
<td>Dairy</td>
<td><em>E. coli</em></td>
</tr>
<tr>
<td>Sunderland et al. (2003)</td>
<td>France and Germany</td>
<td>Commercial</td>
<td>402</td>
<td>7–90 days</td>
<td>Both</td>
<td><em>E. coli</em></td>
</tr>
</tbody>
</table>

NR, not reported.
aAge at enrollment.
containing 11 unique studies met all the inclusion criteria and were included in the SR. The main characteristics of the 11 selected studies are described in Table 1. Four studies reported the funding source [private (Lofstdet et al., 1996; Fecteau et al., 2003), public (Silva et al., 2010), or mixed (Ollivett et al., 2009)] whereas seven did not. No study provided a sample-size calculation, and randomization was unclear in one study (Bywater, 1977).

**Intervention and comparator features**

Intervention (treatment) and comparator (control) groups are described in Table 2. No study provided information about the training of personnel administering treatments, and only two studies reported blinding of personnel (Sheldon, 1977; Ollivett et al., 2009). Due to irrelevant interventions, one or more groups (arms) were not considered: (1) unchallenged calves treated or non-treated (Fecteau et al., 2003; Schnyder et al., 2009; Silva et al., 2010); (2) AMs in combination with other treatments (Bywater, 1977; Silva et al., 2010); and (3) AMs given as prophylactic intervention (Schnyder et al., 2009).

**Outcomes and definitions**

**Clinical outcomes**

The most common clinical outcomes evaluated (FCS, temperature, DH, appetite, and attitude) are described in Table 3. Other clinical variables evaluated included eye position (Lofstdet et al., 1996) and blood in feces, tenesmus, and sucking reflex (Grandemange et al., 2002). In three studies, outcome assessors were reported to be blinded and identified as veterinarians (Sunderland et al., 2003), vet students (Ollivett et al., 2009), or researchers (White et al., 1998); but eight studies did not provide this information.

**Performance outcomes**

One study assessed ADG (Ollivett et al., 2009). Bodyweight gain was evaluated in 6 studies (Bywater, 1977; Grimshaw et al., 1987; White et al., 1998; Fecteau et al., 2003; Schnyder et al., 2009; Silva et al., 2010), however, one study did not report results (Schnyder et al., 2009). Only one manuscript reported the tool or method used to weigh calves (digital scale; Ollivett et al., 2009).

**Health definitions**

Five studies reported a definition of diarrhea (Bywater, 1977; Grimshaw et al., 1987; Grandemange et al., 2002; Sunderland et al., 2003; Ollivett et al., 2009); it was exclusively based on FCS but its description and score-point system highly varied across studies. Five studies used the term ‘diarrhea’ but provided no definition (Lofstdet et al., 1996; White et al., 1998; Fecteau et al., 2003; Schnyder et al., 2009; Silva et al., 2010), and one study defined health events based on abnormal FCS without using the term ‘diarrhea’ (Sheldon, 1997). Two studies reported treatment failure and success (Sheldon, 1997; Grandemange et al., 2002), but the term ‘failure’ was not defined in Sheldon (1997).

**Results of specific outcomes**

**Mortality**

Calf mortality was reported in six studies; the calculated RR for each study was presented as a forest plot (Fig. 2). The RR for three of the comparisons (Amoxicillin vs No treatment; Sulbactam: Ampicillin vs Placebo; Ampicillin vs. Placebo) favored intervention relative to control (the CI did not include 1).

**Severity of diarrhea**

Seven studies reported the severity of diarrhea (Lofstdet et al., 1996; Sheldon, 1997; White et al., 1998; Grandemange et al., 2002; Sunderland et al., 2003; Ollivett et al., 2009; Silva et al., 2010). However, one study (Grandemange et al., 2002) was not considered due to incomplete reporting. Diarrhea severity was reported as RR in one manuscript (Ollivett et al., 2009) and calculated in five manuscripts [mean difference (n = 2); RR (n = 3); Fig. 3]. No comparison favored intervention relative to the comparator.

**Additional results**

A summary of all the statistically significant treatment effects reported in each of the 11 studies is provided in SM 3. Three studies reported assessment of adverse effects after the intervention; two studies found an absence of adverse effects (Lofstdet et al., 1996; Sunderland et al., 2003); and one study observed an increase in diarrhea severity after AM treatment (Schnyder et al., 2009). One study informed of relapse in clinical signs after completing the AM and positive control interventions (Grandemange et al., 2002).

**Risk of bias assessment**

The risk of bias at the outcome level was based on the severity of diarrhea (or FCS; Lofstdet et al., 1996; Sheldon, 1997; White et al., 1998; Grandemange et al., 2002; Sunderland et al., 2003; Ollivett et al., 2009; Silva et al., 2010) or diarrhea duration (Bywater, 1977; Grimshaw et al., 1987; Fecteau et al., 2003; Schnyder et al., 2009). Results of the risk of bias assessment for each domain are shown at the study level (Fig. 4) and as the proportion across all included studies (Fig. 5).

**Discussion**

Dairy and beef calves are often affected with gastrointestinal disorders and treated with AMs (Waldner et al., 2013; NAHMS-USDA, 2018); however, it is unclear if AMs are effective for the treatment of calf gastrointestinal disorders (Smith, 2015). The present work aimed to support the development of calf AM use guidelines by appraising the scientific literature on the efficacy and comparative efficacy of different AM treatments for diarrhea in calves under 6 months of age. Although diarrhea in calves is most common during the first 2 months of life (Preweaning period), we chose an inclusive age criterion because weaning time and age at diarrhea events may vary with management (e.g. production system, breed, and country). Our SR identified 11 relevant studies; nevertheless, the limited number of studies and the differences in interventions (AM class and type of pathogenic agent) prevented us from pursuing a MA evaluation (Valentine et al., 2010). Overall, the eligible studies indicated that diarrhea severity (n = 4, challenge) and mortality (n = 3, challenge; n = 3, natural infection) were numerically inferior after AM intervention; but only three of the aforementioned studies showed significant statistical differences for diarrhea severity (n = 1) and mortality [challenge (n = 1) and natural infection (n = 1)].

Prior SRS evaluating the efficacy of AMs in livestock were also unable to complete a MA due to the heterogeneity of the
<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention</th>
<th>Comparator</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Active ingredient</td>
<td>Dose&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Challenge</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bywater (1977)</td>
<td>Amoxicillin</td>
<td>400 mg TD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fecteau et al. (2003)</td>
<td>Ceftiofur</td>
<td>5 mg kg&lt;sup&gt;−1&lt;/sup&gt; IM</td>
</tr>
<tr>
<td>Lofstedt et al. (1996)</td>
<td>Sulbactam:ampicillin</td>
<td>3.3:6.6 mg kg&lt;sup&gt;−1&lt;/sup&gt; IM</td>
</tr>
<tr>
<td></td>
<td>Ampicillin</td>
<td>6 mg kg&lt;sup&gt;−1&lt;/sup&gt; IM</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ollivett et al. (2009)</td>
<td>Nitazoxanide</td>
<td>1504 mg TD</td>
</tr>
<tr>
<td>Schnyder et al. (2009)</td>
<td>Nitazoxanide</td>
<td>15 mg kg&lt;sup&gt;−1&lt;/sup&gt; PO</td>
</tr>
<tr>
<td>Silva et al. (2010)</td>
<td>Florfenicol</td>
<td>20 mg kg&lt;sup&gt;−1&lt;/sup&gt; IM</td>
</tr>
<tr>
<td>White et al. (1998)</td>
<td>Danofloxacin</td>
<td>1.25 mg kg&lt;sup&gt;−1&lt;/sup&gt; IM</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Natural infection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grandemange et al. (2002)</td>
<td>Marbofloxacin</td>
<td>1 mg kg&lt;sup&gt;−1&lt;/sup&gt; PO</td>
</tr>
<tr>
<td>Grimshaw et al. (1987)</td>
<td>Sulbactam:ampicillin</td>
<td>3.3:6.6 mg kg&lt;sup&gt;−1&lt;/sup&gt; IM</td>
</tr>
<tr>
<td></td>
<td>Ampicillin</td>
<td>6.6 mg kg&lt;sup&gt;−1&lt;/sup&gt; IM</td>
</tr>
<tr>
<td>Sheldon (1997)</td>
<td>Florfenicol</td>
<td>20 mg kg&lt;sup&gt;−1&lt;/sup&gt; IM</td>
</tr>
<tr>
<td>Sunderland et al. (2003)</td>
<td>Danofloxacin</td>
<td>6 mg kg&lt;sup&gt;−1&lt;/sup&gt; SC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PO, oral; IM, intramuscular; SC, subcutaneous; q24h: every 24 h; q12h: every 12 h; q48h: every 48 h.

<sup>a</sup>Doses are reported as mg kg<sup>−1</sup> (milligrams per kilogram of body weight) or as mg or ml TD (total dose in milligrams or milliliters).

<sup>b</sup>Ranges in the treatment length are related to time to cure.

<sup>c</sup>Frequencies are indicative of the frequency reported by the original manuscript; these could briefly differ from the original value.
interventions across primary studies (O’Connor et al., 2006; Sargeant et al., 2019a; 2019b). Even though very few studies were identified in our SR, it is plausible that additional valid research data exist but have not yet been published in peer-review publications, especially if data were generated to support drug label claims or if the study results refuted the initial hypothesis (Constable, 2004; Wellman and O’Connor, 2007). It should be noted that a large number of studies were excluded because they evaluated the efficacy of AMs following a prophylaxis or metaphylaxis treatment approach, or because they defined ‘diarrhea’ based on fecal pathogen shedding instead of clinical signs.

In livestock clinical trials, the accuracy of the outcome measured has raised concerns due to incomplete reporting of methods and study design (Burns and O’Connor, 2008; Sargeant et al., 2009; Winder et al., 2019). However, in the future, the standards of quality clinical trials may improve, as some relevant journals are now requesting authors to use the REFLECT statement (O’Connor et al., 2010), a guide for standardized design and reporting, prior to considering a manuscript for publication. In our SR, most studies were designed as challenge experiments. However, there are limitations associated with challenge studies, as they tend to result in exaggerated treatment effects and do not provide a high level of evidence for the effectiveness of an intervention in a commercial setting (Sargeant et al., 2009, 2019a).

Based on the current FDA indications, most studies included in the SR used AM outside label claims. Marbofloxacin (broad-spectrum fluoroquinolone for dogs and cats) and nitazoxanide (human cryptosporidiosis) are not labeled in the USA for use in cattle or calves, and ceftiofur, danofloxacin, and florfenicol are labeled for calf treatments but for ailments other than diarrhea. Ampicillin and amoxicillin were the only treatments with FDA approval for the treatment of Escherichia coli enteritis in calves. However, the treatment length of ampicillin in these studies was outside label recommendations. The extra-label use of fluoroquinolones (e.g. danofloxacin and marbofloxacin) and cephalosporins (e.g. ceftiofur) is totally prohibited in food animals due to the high risk of AMR emergence based on the ‘Animal Medicinal Drug Use Clarification Act of 1994’ and ‘21 Code Federal Regulations 530’ (FDA, 2021). Also, the route of administration differed across studies, orally (amoxicillin, marbofloxacin, ceftiofur, florfenicol, and danofloxacin). Differences in route of administration may have also contributed to differences in treatment response; oral administration of AMs may induce changes in the microbiome and aggravate diarrhea presentation (Smith, 2015).

Furthermore, our SR revealed that few relevant studies included, as an intervention, the most common AM chosen to treat calf diarrhea in California commercial operations (sulfa- mides as the first choice; ceftiofur products as the second choice; Okello et al., 2021). Although the knowledge that SR provides about AM efficacy and effectiveness is important, it is clearly not the only metric of importance in AM selection. The work of veterinarians and practitioners is key to improving AM use in livestock; other relevant factors that guide veterinarians in AM treatment selection are treatment algorithms and protocols, AM stewardship guidelines, local AM prescribing policies, label recommendations, sensitivity testing results for target animals, and cost–benefit analysis (O’Connor et al., 2019).

Consistent with previous SR in livestock, issues with the risk of bias assessment were observed related to incomplete reporting of the randomization process and the blindness of personnel who delivered treatments and outcome assessors (Francoz et al., 2017; Sargeant et al., 2019b). Randomization was classified as a high-risk grade based on unclear allocation to the treatment as the randomization process was not described, or randomization was outside label recommendations. Marbofloxacin (broad-spectrum fluoroquinolone for dogs and cats) and nitazoxanide (human cryptosporidiosis) are not labeled in the USA for use in cattle or calves, and ceftiofur, danofloxacin, and florfenicol are labeled for calf treatments but for ailments other than diarrhea. Ampicillin and amoxicillin were the only treatments with FDA approval for the treatment of Escherichia coli enteritis in calves. However, the treatment length of ampicillin in these studies was outside label recommendations. The extra-label use of fluoroquinolones (e.g. danofloxacin and marbofloxacin) and cephalosporins (e.g. ceftiofur) is totally prohibited in food animals due to the high risk of AMR emergence based on the ‘Animal Medicinal Drug Use Clarification Act of 1994’ and ‘21 Code Federal Regulations 530’ (FDA, 2021). Also, the route of administration differed across studies, orally (amoxicillin, marbofloxacin, ceftiofur, florfenicol, and danofloxacin). Differences in route of administration may have also contributed to differences in treatment response; oral administration of AMs may induce changes in the microbiome and aggravate diarrhea presentation (Smith, 2015).

Furthermore, our SR revealed that few relevant studies included, as an intervention, the most common AM chosen to treat calf diarrhea in California commercial operations (sulfonamides as the first choice; ceftiofur products as the second choice; Okello et al., 2021). Although the knowledge that SR provides about AM efficacy and effectiveness is important, it is clearly not the only metric of importance in AM selection. The work of veterinarians and practitioners is key to improving AM use in livestock; other relevant factors that guide veterinarians in AM treatment selection are treatment algorithms and protocols, AM stewardship guidelines, local AM prescribing policies, label recommendations, sensitivity testing results for target animals, and cost–benefit analysis (O’Connor et al., 2019).

Consistent with previous SR in livestock, issues with the risk of bias assessment were observed related to incomplete reporting of the randomization process and the blindness of personnel who delivered treatments and outcome assessors (Francoz et al., 2017; Sargeant et al., 2019b). Randomization was classified as a high-risk grade based on unclear allocation to the treatment as the randomization process was not described, or randomization was outside label recommendations.

### Table 3. Scoring systems for clinical outcomes evaluated by studies included in the SR of the efficacy of AMs in the treatment of calf diarrhea

<table>
<thead>
<tr>
<th>Study</th>
<th>Evaluation period (days)</th>
<th>Fecal consistency</th>
<th>Attitude</th>
<th>DH (criteria)</th>
<th>Appetite</th>
<th>Fever</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Challenge</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bywater (1977)</td>
<td>0 to 10</td>
<td>0–3&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0–2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0–2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Fecteau et al. (2003)</td>
<td>−4 to 13</td>
<td>0–1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0–4</td>
<td>–</td>
<td>0–3</td>
<td>&gt;39.2°C</td>
</tr>
<tr>
<td>Lofstedt et al. (1996)</td>
<td>0 to 7</td>
<td>0–4</td>
<td>0–4</td>
<td>0–3 (skin)</td>
<td>–</td>
<td>&gt;40°C</td>
</tr>
<tr>
<td>Ollivett et al. (2009)</td>
<td>−5 to 10</td>
<td>1–3&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1–4&lt;sup&gt;c&lt;/sup&gt;</td>
<td>–</td>
<td>–</td>
<td>NR</td>
</tr>
<tr>
<td>Schnyder et al. (2009)</td>
<td>−1 to 28</td>
<td>Solid-liquid&lt;sup&gt;d&lt;/sup&gt;</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Silva et al. (2010)</td>
<td>−2 to 5</td>
<td>0–2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>NR</td>
</tr>
<tr>
<td>White et al. (1998)</td>
<td>0 to 6</td>
<td>0–3&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0–3</td>
<td>Absent-severe&lt;sup&gt;b&lt;/sup&gt;</td>
<td>NR</td>
<td>–</td>
</tr>
<tr>
<td><strong>Natural infection</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grimshaw et al. (1987)</td>
<td>0 to 7</td>
<td>0–3</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Grandemange et al. (2002)</td>
<td>0 to 3, 7</td>
<td>1–5</td>
<td>1–4</td>
<td>1–4 (skin)</td>
<td>1–2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>&gt;39.5°C</td>
</tr>
<tr>
<td>Sheldon (1997)</td>
<td>0, 2, 4, 10</td>
<td>0–3&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0–4&lt;sup&gt;b&lt;/sup&gt;</td>
<td>NR</td>
<td>–</td>
<td>≥39.5°C</td>
</tr>
<tr>
<td>Sunderland et al. (2003)</td>
<td>0 to 11</td>
<td>Absent–severe diarrhea</td>
<td>Absent–severe depression</td>
<td>Absent–severe (skin)</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

NR, outcome with assessment methods not reported.
<sup>a</sup>Relative to treatment onset; evaluation days are indicative of the period reported by the original manuscript, so these could differ from the original value.
<sup>b</sup>Not clearly described.
<sup>c</sup>Reference provided for the scoring system method.
The most frequent reason for a high-risk classification for blinding of both personnel who delivered treatments and outcome assessors was the absence of blinding reporting. The lack of blinding of personnel delivering treatments could influence the care for the calves during the study. Similarly, outcome assessment could be influenced by knowledge of the intervention delivered, especially for subjective outcomes, such as FCS and DH (Francoz et al., 2017; Sterne et al., 2019). All relevant manuscripts except one were linked to a pharmaceutical company, and that could potentially introduce a source of bias. Furthermore, none of the studies included in our SR sample size calculation, which is consistent with previous reviews (Haimerl et al., 2012; Winder et al., 2019). This might have introduced a publication bias, as underpowered studies with non-significant results are less likely to reach peer-review journals (Sargeant et al., 2009, 2019a). The risk of bias tool was modified to introduce a new domain related to the definition of disease. Most studies were classified with a high risk of bias based on this domain, as the diarrhea definition was missing in about half of the studies, and when reported, the definition was only based on a single outcome. This could lead to biased results due to unnecessary AM administration to diarrheic calves without signs of systemic illness (Constable et al., 2008). Our results are consistent with other SRs which highlighted the lack of disease definition in clinical trials in cattle (Naqvi et al., 2018). Similarly, treatment success and failure definitions were rarely reported; thus, it was difficult to accurately evaluate study results, assessment methods of treatment efficacy, and likely variation sources related to health definitions (Kelly and Janzen, 1986; Wellman and O’Connor, 2007).

In the relevant manuscripts, the evaluation of clinical signs of health disorders was subjective and very diverse across studies. Although FCS was evaluated in all studies, the scoring systems varied highly, even when FCS had the same numerical scale. Additionally, many studies provided a vague description of the FCS categories with only two studies stating a reference; however, those references for FCS methods reported non-validated, unreferenced, and incomplete FCS evaluation methods. No other fecal features beyond consistency were evaluated, and diarrhea

![Fig. 2. Forest plot to illustrate the results about mortality from the studies included in the systematic review of the efficacy of antimicrobials in the treatment of calf diarrhea. P = Placebo; BAQ:SUL = Baquiloprim:Sulfadimidine; CI = Confidence intervals.](https://doi.org/10.1017/S1466252322000032)
severity classification was provided in a single study. Some of the secondary clinical signs evaluated included DH, fever, anorexia, and depression, but their evaluation methods varied across studies, lacked references, and were subjective. DH was evaluated based on skin elasticity without considering body fat, skin location, animal position, and age (Constable et al., 1998). Four studies evaluated fever, but the difference between the maximum and minimum threshold for fever definition reached nearly 1°C across studies, and none of the studies accounted for possible inaccuracies in body temperature assessment related to physiological, environmental, and procedure methods (Hill et al., 2016).

Similarly, scoring systems for attitude and appetite were based on empirical and subjective measurements, and highly differed across studies. Overall, the lack of standardized evaluation methods across the 11 relevant studies was concerning, as low reliability in both the measurement of outcomes and health definitions could contribute to decreased statistical power and thereby an under- or over-estimation of treatment effects (Sargeant et al., 2009). Over four decades ago, calf health evaluation guidelines were proposed to make reporting more uniform across research studies (Larson et al., 1977); however, these guidelines have not been adopted, most likely because of their level of complexity (Kertz and Chester-Jones, 2004). Current industry guidelines for calf diarrhea suggest limiting AM treatments to calves with loose stools that also show systemic signs of illness (e.g. inappetence, DH, lethargy, pyrexia), blood or mucosal shreds in their stool, or concurrent infections (Constable et al., 2008; McGuirk, 2008). None of the relevant studies attained this definition; challenge studies treated all exposed animals, and in natural infection studies, treatment was merely based on FCS. Future studies should address this lack of standardized, validated calf health definitions, which results in heterogeneous treatment decisions and cure definitions. The incorporation and combination of validated health assessment methods are key to accurately identifying sick calves, increasing treatment success, and improving animal welfare both inside and outside of research (McGuirk, 2008; Cramer et al., 2016). Furthermore, standardized assessment methods would lead to greater uniformity in study designs (Larson et al., 1977), making the interpretation and comparison of livestock experiments easier. Moreover, objective outcomes, such as ADG, mortality, and laboratory outcomes, could increase the reliability of studies and the ability to summarize the effect size of interventions. However, in our SR, only one study assessed ADG, six studies reported mortality, and the reported laboratory outcomes were highly diverse and limited to a single evaluation.

Finally, this SR had several strengths; it followed a protocol that was reported in accordance with PRISMA-P (Moher et al., 2015); it adhered to the guidelines for SR in animal agriculture and veterinary medicine (O’Connor et al., 2014a, 2014b; Sargeant and O’Connor, 2014); the search strategy, which used multiple electronic databases, was designed with support from a librarian in order to identify the highest number of available studies; and to increase the reliability of the process, the screening, data extraction, and the risk of bias assessment were independently performed by two or more reviewers with a background in veterinary and animal science as well as in research synthesis methods (Sargeant and O’Connor, 2014).

On the other hand, our SR could have some limitations. We did not consider gray literature as a relevant source. On average, only 50% of abstracts reporting the results of RCTs reach full publication, and the calculated abstract-to-publication ratio for some bovine conferences is <10% (Dickersin et al., 1994; Brace et al., 2016).
Thus, excluding these studies could result in lower precision in the estimate of intervention effect and may result in biased results by introducing publication bias (Dickersin et al., 1994; Sargeant and O’Connor, 2014). However, excluding gray literature may have had a limited impact, as it usually involves short abstracts with not enough data to conduct research synthesis methods (Burns and O’Connor, 2008; Brace et al., 2010; Sargeant and O’Connor, 2014). Another source of bias may be the exclusion of potentially relevant articles published in a language other than English. In veterinary medicine, the impact of language restrictions remains unknown (Burns and O’Connor, 2008), while in human medicine, limiting the language of publication of trial reports to English in SR of conventional interventions (e.g. AMs) does not change the estimates of the effectiveness of an intervention (Moher et al., 2003; Pham et al., 2005). Therefore, the impact of exclusion of manuscripts in languages other than English was likely minimal in the present SR.

Conclusions

At present, the efficacy of AMs in the treatment of calf diarrhea cannot be evaluated using MA methods, as the SR identified few relevant studies testing heterogeneous interventions. Our SR revealed important limitations in study design and reporting, which future studies should overcome in order to perform a valuable MA evaluation of the efficacy of AMs in the treatment of calf diarrhea. The interventions tested should reflect common on-farm treatment approaches, the research community needs to reach an agreement on the definition and outcome evaluation systems of diarrheal disease, and studies should adhere to reporting guidelines.

Supplementary material. The supplementary material for this article can be found at https://escholarship.org/uc/item/0nw528h4supplemental.

Acknowledgments. The authors gratefully acknowledge the support from all members of the Dr Silva-del-Río Lab for their contributions, and special thanks to Dr Ainhoa Valdecabres for her support in developing the forest plots.

Financial support. Financial support for this research was partially provided by the California Department of Food and Agriculture (CDFA) Antimicrobial Use Stewardship Program (the sponsor had no role in either protocol or review).

Conflict of interest. No author has conflicts to declare.

References


European Food Safety Authority – EFSA (2010) Application of systematic review methodology to food and feed safety assessments to support decision making. EFSA Journal 8, 1637–1727.


Nagvi SA, Nobrega DB, Ronski PD and Barkema HW (2018) Invited review: Effectiveness of precalving treatment on postcalving udder health: effect size analysis to update a Cochrane review. Evidence-Based Veterinary and Comparative Medicine, 4728.


Pham B, Klassen TP, Lawson ML and Moher D (2005) Language of publication restrictions in systematic reviews gave different results depending on whether the intervention was conventional or complementary. Journal of Clinical Epidemiology 58, 769–776.

Sargeant JM and O’Connor AM (2014) Conducting systematic reviews of intervention questions II: relevance screening, data extraction, assessing risk of bias, presenting the results and interpreting the findings. Zoonoses and Public Health 61, 39–51.


