Characterization of controlled trials on probiotic supplementation to dairy calves: A scoping review

R. Branco-Lopes,1 C. Bernal-Córdoba,1 A. Valdecabres,2 C. Winder,3 M. E. Canozzi,4 and N. Silva-del-Río1,5*

1Veterinary Medicine Teaching and Research Center, Tulare, CA 93274
2Teagasc, Animal and Grassland Research and Innovation Center, Moorepark, Fermoy, Co. Cork, Ireland P61 C996
3Department of Population Medicine, University of Guelph, Guelph, ON, Canada, N1G 2W1
4Instituto Nacional de Investigación Agropecuaria (INIA), Programa Producción de Carne y Lana, Estación Experimental INIA La Estanzuela, Colonia, Uruguay, 70000
5Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis 95616

ABSTRACT

The objective of this scoping review was to identify, describe, and characterize the literature on probiotic supplementation in dairy calves. Eligible studies were nonrandomized, quasi-randomized and randomized controlled trials in English, Spanish, or Portuguese that evaluated the effect of probiotic supplementation on growth and health of dairy calves. The search strategies were based on a modification of the PICO (Population, Intervention, Comparator, Outcome) framework and used synonyms and words related to “dairy calves” (population), “probiotics” (intervention), and “growth and health measurements” (outcomes). No restrictions for publication year or language were applied. Searches were conducted in Biosis, CAB Abstracts, Medline, Scopus, and the Dissertations and Theses Database. In total, the search identified 4,467 records, of which 103 studies (110 controlled trials) met the inclusion criteria. The studies were published between 1980 and 2021 and originated from 28 countries. Trials were randomized (80.0%), nonrandomized (16.4%), and quasi-randomized (3.6%), ranging in sample size from 5 to 1,801 dairy calves (mode = 24; average = 64). Enrolled calves were frequently Holstein (74.5%), males (43.6%), and younger than 15 d at the beginning of probiotic supplementation (71.8%). Often, trials were conducted in research facilities (47.3%). Trials evaluated probiotics with single or multiple species of the same genus: Lactobacillus (26.4%), Saccharomyces (15.4%), Bacillus (10.0%), Enterococcus (3.6%), or multiple species of various genera (31.8%). Eight trials did not report the probiotic species used. Lactobacillus acidophilus and Enterococcus faecium were the species most supplemented to calves. The duration of probiotic supplementation ranged from 1 to 462 d (mode = 56; average = 50). In trials with a constant dose, it ranged from $4.0 \times 10^6$ to $3.7 \times 10^{11}$ cfu/calf per day. Most probiotics were administered mixed solely into feed (88.5%; whole milk, milk replacer, starter, or total mixed ration) and less frequently orally as a drench or oral paste (7.9%). Most trials evaluated weight gain (88.2%) as a growth indicator and fecal consistency score (64.5%) as a health indicator. Our scoping review summarizes the breadth of controlled trials evaluating probiotic supplementation in dairy calves. Differences in intervention design (mode of probiotic administration, dose, and duration of probiotic supplementation) and outcomes evaluation (type and methods) justify future efforts toward standardized guidelines in clinical trials.

Key words: calf, direct-fed microbial, feed additive, review

INTRODUCTION

Probiotics, also known as direct-fed microbials, have been proposed as a strategy to improve growth and enhance health of dairy calves (Cangiano et al., 2020). The International Scientific Association for Probiotics and Prebiotics defines probiotics as “live microorganisms that, when administered in adequate amounts, confer a health benefit on the host” (Hill et al., 2014). According to the US National Animal Health Monitoring System, 23% of dairy heifer operations supplement probiotics, and this management strategy is more common in large operations (38%; >500 cows; USDA, 2016).

Traditionally, lactic acid-producing bacteria (e.g., Lactobacillus spp.; Uyeno et al., 2015) and yeast-based products (e.g., Saccharomyces spp.; Alugongo et al., 2017) have been the most studied probiotics for dairy calves. Additionally, according to the 2019 Direct-fed-
Microbial, Enzyme and Forage Additive Compendium (Feedstuffs, 2019), most probiotics commercialized for cattle contain Lactobacillus acidophilus and Enterococcus faecium. In dairy calves, the proposed benefits associated with probiotic supplementation include improving ADG and feed efficiency (Sun et al., 2010; Le et al., 2017), reducing antibiotic use, maintaining growth rate in diarrheic calves (Villot et al., 2019), and decreasing the incidence of scours (Timmerman et al., 2005). Probiotics provide health or productivity benefits to the host by different mechanisms, such as competing for adhesion to the epithelium and for nutrients, producing antibacterial substances (e.g., bacteriocins and hydrogen peroxide), changing the gastrointestinal microenvironment (e.g., pH lowering), and modulating the immune system (Bermudez-Brito et al., 2012).

Evidence synthesis methods have been used to evaluate the efficacy of probiotics on calves (Frizzo et al., 2011; Signorini et al., 2012; Alawneh et al., 2020). However, none of these previous reviews has mapped current literature trends or addressed differences in study characteristics (e.g., funding source and facility where experiment was carried out), study design (e.g., sample size justification and randomization process), and outcomes characteristics (e.g., description of evaluators and their training) across controlled studies. This information is key to interpreting results from existing studies and informing future research during the process of study design. A scoping review (ScR) is the appropriate evidence synthesis method to examine how research on a certain topic is conducted and to identify and analyze knowledge gaps (Munn et al., 2018). A scoping review addresses broader review questions than systematic reviews and meta-analyses, which are frequently focused on effectiveness (Peters et al., 2020). The results of ScR are not used for recommendations; these reviews describe, rather than analyze, the available information (Lockwood et al., 2019). Therefore, the objective of this study was to identify, describe, and characterize the literature on probiotic supplementation in dairy calves.

METHODS

The protocol for this ScR was adapted from a protocol developed for a systematic review on probiotic use for dairy calves, which is deposited on University of California (UC) Davis eScholarship (https://escholarship .org/uc/item/2r93v26f) and is available via Systematic Reviews for Animals and Food (https://www.syreaf .org/protocol/). This article is reported following the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) extension for scoping reviews (PRISMA-ScR; Tricco et al., 2018). Because no human or animal subjects were used, this analysis did not require approval by an Institutional Animal Care and Use Committee or Institutional Review Board.

Eligibility Criteria

Eligible studies were primary research studies, including nonrandomized, quasi-randomized, and randomized controlled trials, written in English, Spanish, or Portuguese (at least one reviewer was proficient in these languages). Published studies as peer-reviewed and non-peer-reviewed articles (thesis and dissertations) were included in the ScR. Eligibility criteria were defined based on the PICO (Population, Intervention, Comparator, Outcome) framework (Cooper et al., 2018). The target population was dairy calves (up to 7 mo of age at enrollment), with no restrictions on calves’ breed or sex. The intervention of interest was probiotic supplementation, regardless of probiotic type, dose, or supplementation duration. Studies that named probiotics as direct-fed microbials were also included. Studies using probiotics as treatment therapy were not considered. The eligible comparator was control group consisting of calves that were untreated or received placebo. Eligible studies included publications that measured a growth performance outcome (e.g., ADG, body traits) or a health outcome (e.g., serum metabolites, immunological parameters).

Information Sources

Identification of the electronic databases was completed with the assistance of a UC Davis research librarian specialized in veterinary science. The following databases were searched to identify relevant literature: Biosis (Web of Science, 1926–2021), CAB Abstracts (CAB Direct, 1973–2021), Medline (PubMed, 1966–2021), and Scopus (Scopus, 1996–2021). The Dissertations and Theses Database (ProQuest, 1861–2021) was searched to retrieve gray literature. In addition, the bibliographies of relevant studies were searched by hand.

Search Strategy

The search strategy was developed based on the PICO framework and in collaboration with a UC Davis librarian. Key words were selected for each PICO concept from the relevant literature. Posteriorly, the librarian performed a keyword mining in PubMed and CAB Direct. The Yale MeSH analyzer was also used to compare common Medical Subject Headings across studies for PubMed. No restrictions for publication year, study design, or language were applied during the
searches. The literature search was conducted on February 27 and March 3, 2020, and updated on August 19, 2021. The search strategy used in Medline (PubMed) is described in Table 1; search strings were adjusted accordingly to each database to fit its specific formats (Supplemental File S1; https://osf.io/bwfjd). Search results were uploaded to the reference manager F1000 (Faculty of 1000 Ltd.), and duplicates were removed. The de-duplicated results were then exported to the Covidence systematic review management software (Veritas Health Innovation) for screening.

### Selection of Sources of Evidence

All records were screened for eligibility twice by 2 reviewers independently. Both reviewers were trained on the methodology, and neither was blinded to journal or author names. First, the titles were screened, using the following questions to identify eligible studies:

1. Does the title describe a study involving dairy calves?
2. Does the title describe a study with probiotic(s) supplementation?

Second, the abstracts were screened, using the following questions:

3. Does the abstract describe a primary research study?
4. Does the abstract describe a study involving dairy calves supplemented with probiotic(s)?
5. Does the abstract describe one or more measurements of performance (e.g., ADG, feed efficiency) or health (e.g., fecal consistency score, diarrhea incidence)?

For title and abstract screening questions, the possible answers were “no,” “maybe,” and “yes.” Studies were excluded if both reviewers answered “no” to any question. In both screenings (title and abstract), conflicts between the 2 reviewers were discussed until a consensus was reached.

Full-text screening was conducted by one reviewer. This screening included questions (3), (4), and (5), which were adapted for full-text screening (the word “study” was used instead of “abstract”), plus

6. Is the study a trial with a control group?
7. Is the study written in English, Spanish, or Portuguese?
8. Is the probiotic a supplementation strategy (not treatment for sick animals)?
9. Is the study population (dairy calves) ≤7 mo old at enrollment?

During the full-text article screening, the available answers were “no” and “yes.” Studies were excluded if the answer was “no” for at least one of the questions. The exclusion reason was recorded at this screening level.

### Table 1. Results of the search strategy used to identify records; search strategy used in Medline (PubMed)

<table>
<thead>
<tr>
<th>Search ID</th>
<th>Terms</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>#4</td>
<td>#1 AND #2 AND #3</td>
<td>661</td>
</tr>
<tr>
<td>#5</td>
<td>#4 AND (2020:2021[pdat])</td>
<td>121</td>
</tr>
</tbody>
</table>

1 [tiab = title/abstract (words included in the title or abstract of a citation); MeSH = Medical Subject Headings (a controlled vocabulary of terms from the US National Library of Medicine); Pdat = publication date.]
All screening questions were pilot tested. For this, the questions were developed according to the eligibility criteria and tested on the first 30 studies listed in the systematic review management software (Covidence). Primary screening questions (1 to 5) were tested by 2 reviewers and secondary questions (6 to 9) by one reviewer; subsequently, the clarity and objectivity of questions were discussed with a third reviewer.

Data Charting Process and Data Items

Data from studies that met eligibility criteria were extracted into a predesigned form built in Excel (Excel 2022 v. 2206, Microsoft Corp.). The data extraction form was pilot tested on 5 studies randomly selected, and the relevance of the data extracted was discussed with a second reviewer. Study-level data consisted of authors’ names, journal’s name, language, trial’s country (when not available, the country was based on first author’s affiliation), year of publication, funding source, randomization process, sample size, sample size justification, and facility type (commercial vs. research). Population characteristics consisted of calf breed, sex, and age at the start of probiotic supplementation. Intervention and comparator data consisted of description of comparator, commercial name of probiotic, scientific name, dose, mode of probiotic administration (e.g., whole milk, milk replacer), and duration of probiotic supplementation. Outcomes data consisted of all outcomes evaluated, method of measurement, evaluation frequency, evaluators, and their training and blinding. To guarantee the accuracy of data extracted from selected studies, the extraction process was performed twice, and discrepancies were corrected.

Synthesis of Results

The screening process was described using the PRISMA flowchart (Page et al., 2021). Excel (Microsoft Corp.) was used to summarize eligible studies in tables and graphs and to generate numerical outcomes’ metrics such as average, mode, and range. The choropleth map was developed using MapChart (https://www.mapchart.net/). Based on the reported probiotic genus, trials were classified into 6 categories: *Bacillus* (BC, including monospecies and multispecies), *Enterococcus* (ENT, including monospecies), *Lactobacillus* (LB, including monospecies and multispecies), *Saccharomyces* (SC, including monospecies), multiple genera (Multi, including monospecies and multispecies), or undefined species and less frequent genera (Other, including monospecies). Throughout the article, data are presented at 3 different levels: study, trial, and arm (or treatment group).

RESULTS

Selection of Sources of Evidence

Results of the database searches, screening process, and exclusion reasons are summarized in Figure 1. In total, the search identified 4,467 records, of which 3,426 remained after de-duplication. After all screening levels, 3,323 records were considered irrelevant and excluded. During the full-text article screening, 46 articles were excluded, mainly due to lack of clarity on breed purpose of enrolled animals (41.3%; 19/46 studies with inconclusive dairy or beef breed). Overall, 103 studies met the eligibility criteria and were included in the ScR, reflecting 110 controlled trials and 322 arms (165 arms evaluated probiotics). The complete list of included papers is presented in Supplemental File S2 (https://osf.io/rj8d2).

General Study Characteristics

Study-Level Characteristics. The included studies were published between 1980 and 2021 as peer-reviewed articles (n = 100; 97.1%) or as dissertations or theses (n = 3; 2.9%). Most of the studies were published within the last decade (n = 61; 59.2%), indicating increased research interest in probiotic supplementation to dairy calves over the years. The number of studies published per year stratified by probiotic type is shown in Figure 2. The geographical distribution of studies is shown in Supplemental File S2. Studies were conducted in 28 countries, mainly in the United States (n = 22; 21.3%), Brazil (n = 13; 12.6%), and China (n = 9; 8.7%). Most studies were written in English (n = 88; 85.4%), followed by Portuguese (n = 11; 10.7%), Spanish (n = 2; 1.9%), or Spanish and English (n = 2; 1.9%).

Study Design and Randomization. The 103 studies included 110 trials (control vs. treatment: n = 52; control vs. multiple treatments: n = 58), of which 88 were randomized, 18 nonrandomized, and 4 quasi-randomized trials. Randomized trials were designed as randomized controlled (n = 50; 56.8%), randomized block (n = 21; 23.9%), factorial (n = 16; 18.2%), or incomplete block designs (n = 1.1%). Among randomized and quasi-randomized trials (n = 92), only 9 trials described the randomization process (i.e., preassigned ballots “in a hat” strategy, random list, ear tag, location, alternation), and the remaining trials solely noted that calves were randomly allocated to treatments (n = 83). Randomized block trials (n = 21) used the following variables alone or in combination as blocking factors: body weight or birth body weight (n = 10), sex (n = 7), birth date (n = 4), age (n = 2),
serum total protein (n = 2), breed (n = 1), and farm of origin (n = 1).

**Sample Size.** Trial sample size ranged from 5 to 1,801 dairy calves (mode = 24; average = 64). Half of trials (n = 56; 50.9%) enrolled 30 or fewer calves (Table 2). The sample size was justified in 9 trials based on power analysis (n = 8) and facilities capacity (n = 1). Also, one trial reported a sample size justification, but the outcome used for it was unclear. The outcomes used for power analysis calculations were growth performance (n = 3), gut total bacteria concentration (n = 1), antimicrobial resistance of fecal coliform counts (n = 1), diarrhea incidence (n = 1), *Mannheimia haemolytica* incidence (n = 1), and both rectal temperature and total peripheral blood neutrophil counts (n = 1).

**Funding.** Studies were funded by public institutions (n = 39; 37.9%), private companies (n = 13; 12.6%), both public and private institutions (n = 10; 9.7%), did not receive financial support (n = 2; 1.9%), or did not disclose funding (n = 39; 37.9%). Some studies that did not provide a funding source had probiotic-, feed-, seed-, or pharmaceutical-industry authorship (n = 10; 25.6%).

---

**Figure 1.** Flowchart describing the number of records identified, included, and excluded, and the reasons for exclusions through the screening process of the scoping review. The left-hand Identification box represents records identified by search from February 27 to March 3, 2020; the right-hand box shows records identified by an updated search on August 19, 2021. The chart format was adapted from Page et al. (2021).
Trials were conducted in research facilities (n = 52; 47.3%), commercial dairy facilities (n = 28; 25.4%), or did not provide the facility type (n = 30; 27.3%; Table 2). Among trials performed in commercial settings, 78.6% were conducted in farms (n = 17; dairy farm and n = 5; unspecified type of farm) and, less frequently, in calf-raising operations (n = 3; 10.7% calf ranch and n = 3; 10.7% veal facility).

**Facility Type.** Trials were conducted in research facilities (n = 52; 47.3%), commercial dairy facilities (n = 28; 25.4%), or did not provide the facility type (n = 30; 27.3%; Table 2). Among trials performed in commercial settings, 78.6% were conducted in farms (n = 17; dairy farm and n = 5; unspecified type of farm) and, less frequently, in calf-raising operations (n = 3; 10.7% calf ranch and n = 3; 10.7% veal facility).
Population

Breed. Enrolled calves were mainly Holstein (n = 82; 74.5%); others were Holstein crossbreed (n = 7; 6.4%), Jersey (n = 3; 2.7%), at least 2 different breeds (n = 7; 6.4%), and other breeds (n = 9; 8.2%). Calf breed was not reported in 2 trials, although the authors stated that calves were from dairy herds (Table 2).

Sex. Often trials included only males (castrated or not; n = 48; 43.6%; Table 2), followed by trials that enrolled both males and females (n = 31; 28.2%) and those including only females (n = 21; 19.1%); sex of calves was not provided in 10 trials (9.1%).

Age. The calves’ age at start of supplementation ranged from birth to 192 d (mode = 1; average = 21; Table 2). Most trials (n = 79; 71.8%) started probiotic supplementation when calves were younger than 15 d. Calf age was unclear or not reported in 9 trials.

Intervention and Comparator

Probiotics. A full description of the species investigated by the studies is provided in Supplemental File S2. The individual probiotic genus most investigated was Lactobacillus (n = 29 trials; 26.4%), followed by Saccharomyces (n = 17 trials; 15.4%) and Bacillus (n = 11 trials; 10.0%); however, most of the studies investigated multi-species probiotics that contained diverse genera (n = 35 trials; 31.8%). Among trials, Lactobacillus acidophilus was the species most studied (n = 58 arms); 11 different strains of L. acidophilus were evaluated. Enterococcus faecium was the second most evaluated strain (n = 33 arms), and 6 different E. faecium strains were assessed. Eight trials did not report the probiotic species used. Forty-eight different commercial products were assessed; but only a few of those were evaluated 3 or more times (Levucell SC (n = 6 trials; Lallemand Biochem International), Levucell SB20 (n = 6 trials; Lallemand Biochem International), Protexin (n = 4 trials; Probiotics International Ltd.), Biomate FG (n = 3 trials; Chr Hansen), and BioPLus 2B (n = 3 trials; Chr Hansen)]).

Duration of Supplementation. The duration of probiotic supplementation is shown in Figure 3. Overall, the duration of probiotic supplementation ranged from 1 to 462 d (mode = 56; average = 50). Probiotics were often supplemented for 60 d or less (n = 109 arms; 66.1%), and, less frequently, for more than 76 d (n = 17 arms; 10.3%). The duration of supplementation was unclear or not reported in 16.4% of arms (n = 27).

Mode of Administration. The mode of probiotic administration is represented in Figure 4. Probiotics were administered mixed solely into feed (n = 146 arms; 88.5%; Colostrum, milk, milk replacer, starter, concentrate, or TMR) and less frequently orally as drench, gel, or paste (n = 13 arms; 7.9%), ruminally (n = 1 arm; 0.6%), or nasally (n = 1 arm; 0.6%). For all 6 probiotic categories, milk (colostrum, skim, whole or waste milk, and milk replacer) was the most adopted mode of probiotic administration (BC: 53.3%; ENT: 50.0%; LB: 76.6%; Multi: 67.3%; SC: 43.3%; Other: 64.7%).

Dose. The probiotic dose administered to dairy calves in each trial is presented in Supplemental File S2. In trials with a constant dose throughout the experiment, the probiotic dose ranged from \(4.0 \times 10^6\) to \(3.7 \times 10^{11}\) cfu/calf per day. The constant dose was \(\geq 10^9\) cfu/calf per day in 100% of ENT and SC, 85.7% of BC and Other, 63.1% of LB, and 57.8% of Multi trials. The dose was not reported in 7 trials (6.4%).

Comparator. Few trials used a placebo (n = 19; 17.3%) as the comparator group; most trials adopted a nonsupplemented group as control (n = 91; 82.7%).

Outcomes

Growth. The most common growth performance outcomes were weight gain (n = 97; 88.2%), feed intake (n = 75; 68.2%), and feed efficiency (n = 53; 48.2%; Figure 5). The method used to measure or estimate weight was specified in 16.5% of the trials (scale: n = 15; tape: n = 1). Calves’ weight was determined weekly (n = 43; 44.3%), every other week (n = 10; 10.3%), or in a different pattern (e.g., at enrollment and at the end of the experimental period, monthly, 3 times during experiment; n = 35; 36.1%). The frequency of weight measurement was not reported in 9 trials (9.3%).

Health. Common health outcomes evaluated included fecal consistency score (n = 71; 64.5%), fecal, intestinal, or ruminal microflora (n = 52; 47.3%), and blood parameters (n = 51; 46.4%; Figure 5). Most trials (n = 42; 59.1%) provided a reference for the fecal consistency scoring system used; 19 different references were used across trials. Larson et al. (1977) was the most commonly adopted scoring system (n = 13; 30.9%). The frequency of fecal consistency score assessments varied across trials from daily (n = 51; 71.8%) to 2 or 3 times/d (n = 5; 7.0%) to less frequently (e.g., every other day, twice a week, or weekly; n = 7; 9.8%). Eleven percent of the trials did not report the frequency of assessment (n = 8).

Evaluator. The evaluator of the outcomes was reported in 12 trials (10.9%), and in 7 trials (6.4%) evaluators were blinded while assessing at least one outcome. Measurement of at least one outcome was performed by veterinarians (n = 5), farm workers (n = 3), researchers
DISCUSSION

This scoping review aimed to identify, describe, and characterize controlled trials evaluating probiotic supplementation in dairy calves. The large number of studies included in this review demonstrates extensive research on probiotics supplementation to dairy calves, which has spanned 4 decades and has been performed worldwide. This substantial literature supports future systematic reviews and meta-analyses, which may be more comprehensive than previous reviews (Frizzo et al., 2011; Signorini et al., 2012; Alawneh et al., 2020). However, the variability of experimental designs and the incompleteness of reporting among studies may make meta-analyses challenging. Given the observed trend for increased research interest in probiotic supplementation to dairy calves, it is of utmost importance that researchers exercise best practices when designing future experiments so the aforementioned powerful synthesis methods can help the research community and additional stakeholders reach conclusions on a demonstrated area of research interest.

General Characteristics of Selected Studies

**Randomization.** In our ScR, description of the randomization process was uncommon among the randomized trials (9.7%). Adequate reporting of the allocation method is essential in systematic reviews, which evaluate the risk of bias (Page et al., 2021). The assessment of risk of bias considers biases arising from different domains (different stages of a trial). There are different tools to evaluate the risk of bias; for example, the Cochrane tool includes 5 bias domains, 1 of which is related to randomization (Sterne et al., 2019). Inadequate reporting of a randomization process does not mean that the methods were inappropriate or that the study had flaws (details could be omitted due to...
Randomization may not always result in comparable groups, and pre-treatment assessment may be important to inspect baseline differences. Jarett et al. (2021)

Figure 4. Mode of probiotic administration across probiotic arms included in the scoping review (n = 165 arms). BC = Bacillus (n = 15), ENT = Enterococcus (n = 4), LB = Lactobacillus (n = 47), SC = Saccharomyces (n = 30), Multi = multiple genera (n = 52), Other (n = 17); less frequent genera (n = 6) and not reported species (n = 11). The milk category included trials that mixed probiotics into colostrum, whole, skimmed or waste milk, a mixture of whole milk and milk replacer, and milk replacer solely. The starter category included trials that administered the probiotic into starter or concentrate. The milk and starter category included trials that mixed the probiotics into milk (replacer or whole) and starter during the whole experiment or separately. The other mode category included trials that administered the probiotic via oral (syringe, drench, paste, gel), ruminal (fistula), feed (milk replacer combined with oral application, milk replacer combined with TMR, TMR), nasal. The not reported category included trials that did not reported mode of probiotic administration.

Figure 5. Heat map representing the frequencies of growth and health outcomes evaluated in the included trials (n = 110). Frequencies of 0 to 25% are white, 26 to 50% light gray, 51 to 75% dark gray, and 76 to 100% black. BC = Bacillus, ENT = Enterococcus, LB = Lactobacillus, SC = Saccharomyces, Multi = multiple genera, Other [less frequent genera (n = 6) and not reported species (n = 8)]. Outcomes: Weight gain = ADG and body gain; Rumen parameters = VFA, ruminal histomorphology, ruminal pH, enzymatic activity, ruminal ammonia; Fecal consistency score = fecal consistency score and diarrhea; Microbiota = fecal, intestinal, or ruminal microbiota assessments; Blood parameters = glucose, blood urea nitrogen, cholesterol, triglycerides, cortisol, plasma total protein, albumin, globulin, BHB, nonesterified fatty acids, insulin, ghrelin, glucagon, creatinine kinase, aspartate aminotransferase, alkaline phosphatase, glutamate dehydrogenase, γ-glutamyl transferase, lactate dehydrogenase, catalase, inorganic phosphorus, inorganic calcium, inorganic iron, hematocrit, red and white blood cells, IgA, IgE, IgM, IgG, IL-1β, IL-4, IL-6, IL-10, IFN-γ, tumor necrosis factor-α; Clinical examination = rectal temperature, heart rate, respiration rate, ear score, eye score, navel score, attitude score, nasal discharge, dehydration score, joint score, bloating, umbilical palpation, clinical examination to identify bovine respiratory disease; Other = duration of diarrhea treatment, days with diarrhea, antibiotic use, colon histomorphology, intestine pH, fecal pH, fecal bacteria antibiotic susceptibility test, Johne’s disease, meat and carcass traits, nasal microbiota, fecal DM, gene expression.
proposed best practices for microbiome study design in companion animals and recommended pre-screening of the gut microbiome of all animals before enrollment. Previous research in humans suggests that probiotic efficacy might depend on the initial characteristics of the microbial ecosystem (e.g., butyrate concentration; Ferrario et al., 2014). There are different study designs to tackle interindividual microbiome variability; for example, one study included in our ScR adopted a crossover design with a 20-d washout period (Watanabe et al., 2019). However, determining the length of the washout period may be challenging, and the probiotic may persist after the washout (Gibson et al., 2011). Although the consideration of initial microbiome composition in the study design (e.g., treatment allocation) may not be always feasible, future studies should assess and report microbiome composition before probiotic treatment as it may be a key determinant of probiotic effectiveness.

**Sample Size.** Fifty percent of trials enrolled 30 or fewer calves, and only 10% justified the sample size. Consequently, some of the trials included in the present ScR might have underpowered statistical tests, which are prone to incur a type II error (Christley, 2010), and fail to reject a null hypothesis that is actually false. A power analysis estimates the minimum sample size needed to statistically detect the difference between treatment groups (Cohen, 1992). It should be calculated during the experimental design, as post hoc analyses are considered conceptually flawed and analytically misleading (Zhang et al., 2019). If power analysis is used to justify the sample size for a primary outcome as shown in this ScR (e.g., based on diarrhea incidence or Mannheimia haemolytica incidence), the remaining outcomes of that study should be interpreted with caution.

**Funding.** In our ScR, 37.9% of the reviewed studies failed to disclose sponsorship; this is a larger proportion than previously reported in probiotics human studies (23.9 to 32.0%; Mugambi et al., 2013; Saa et al., 2019). The review by Mugambi et al. (2013) reported that industry-funded studies were more likely to report at least one favorable outcome than non-industry-funded trials, but Saa et al. (2019) did not find that association. Disclosing the funding source would help readers evaluate publication bias, which occurs when studies with positive results are more likely to be published (Song et al., 2010).

**Facility Type.** Our ScR revealed that trials were often conducted at research facilities (47.3%). It is generally assumed that research facilities are managed following industry best practices, which may decrease the challenges that animals are exposed to. Moreover, personnel of research facilities might be better trained to implement treatments and handle samples. In contrast, using commercial facilities in research studies may require undesired adjustments to the study protocol based on their husbandry practices, and farm personnel may have little experience or time to comply with the study protocol. However, Johnston et al. (2003) pointed out that experiments conducted on commercial farms may provide valuable information on the efficacy of new technologies. Additionally, large commercial operations may offer the opportunity to enroll larger numbers of animals (Engstrom et al., 2010). Thus, study methods should clearly describe the trial facilities, as well as implemented husbandry practices (Nevalainen, 2014) and the role of farm personnel in study implementation, regardless of whether studies are conducted in commercial or research facilities.

**Population.** Based on the PICO framework, our eligibility criteria for population (dairy calves up to 7 mo at enrollment) were established to include studies supplementing probiotics in preweaning and weaned dairy calves. Traditionally, probiotics have been studied as a strategy to improve gut health and decrease diarrhea (Cangiano et al., 2020), which is more likely in preweaning calves, especially during the first 3 wk of life (Klein-Jöbstl et al., 2014; Cruvinel et al., 2020). The high interest in probiotic supplementation in early life is seen in this ScR, with most of the trials starting probiotic supplementation at <15 d of life. At early ages, the gastrointestinal microbiota is highly variable and prone to change (Jami et al., 2013; Yáñez-Ruiz et al., 2015), potentially making microbial modulation by probiotics more effective. Probiotics have also been explored as growth promoters in calves and other livestock animals (Barba-Vidal et al., 2019; Jha et al., 2020). Thus, a different reasoning may explain the use of probiotics at an older age, as shown in the range of age at start of supplementation summarized in this ScR.

Our ScR revealed that probiotic research in dairy calves was frequently undertaken with Holstein and male calves. We hypothesized that the lower purchase price of male compared with female dairy calves (Marquou et al., 2019) may have influenced the greater use of males, especially in experiments that euthanized animals. It has been suggested that the composition of gastrointestinal microbiota may differ by breed (Gonzalez-Recio et al., 2018) and sex (Li et al., 2019). Thus, because probiotic effectiveness may be influenced by the initial microbiome, it is important to conduct probiotic research with different sexes and breeds, and these need to be taken into consideration when evaluating probiotic research.

**Intervention.** Our ScR found that most studied genus was *Lactobacillus*. Additionally, it is notewor-
hly that 8 studies did not state the species evaluated. *Lactobacillus acidophilus* and *E. faecium* were the most evaluated strains among trials. Both species are highly used in commercial probiotics designed for cattle use (Feedstuffs, 2019) and are constituents of calves’ gut microbiota (Rodriguez-Palacios et al., 2009; Kumar et al., 2022). Autochthonous strains seem to establish more efficiently in the gastrointestinal tract compared with allochthonous strains (Frese et al., 2012). In our ScR, recent studies reported 2 species (*Bacillus megaterium* and *Candida tropicalis*) that had not been studied in previous years. *Bacillus megaterium* was isolated from chicken manure (Deng et al., 2021), whereas the source of *C. tropicalis* was not specified (Bi et al., 2017; Kong et al., 2019). However, it has been reported that *Candida* spp. can be found in the gastrointestinal tracts of different animals (Sidrîm et al., 2010; Marrero et al., 2011; Mandal and Ghosh, 2013). The screening of autochthonous strains is becoming more feasible with the accessibility of genetic sequencing tools, which is contributing to the emergence of a new kind of probiotics termed “next-generation probiotics” (microorganisms identified based on comparative microbiota analyses; Martín and Langella, 2019). Future researchers have the opportunity to incorporate the new technologies into their screening for new probiotics strains.

Our ScR revealed that the probiotic dose supplemented to calves varied among studies. The probiotic dose is fundamental for its effectiveness, as shown by Renaud et al. (2019) in a study where only the highest evaluated dose of *Saccharomyces cerevisiae boulardii* CNCM 1-1079 had an effect on calf growth performance. The probiotic definition states that “probiotics are live microorganisms that, when administered in adequate amounts, confer a health benefit on the host” (Hill et al., 2014). However, “adequate amounts” is subjective, and Hill et al. (2014) did not provide a dose recommendation for humans or animals. The effective dose for each probiotic might depend on the strain, desired health outcome, vehicle, and mode of probiotic administration (Ouwehand, 2017). However, shipment and storage conditions (e.g., temperature, humidity, pressure) may decrease probiotic viability (Gurram et al., 2021). Thus, an overage is commonly included in commercial probiotics to ensure that colony-forming units are equal to or greater than the label dose, and to take into account potential losses in viability (Weitzel et al., 2021).

In line with young calves’ feeding management, most of the studies included in the ScR provided the probiotic mixed into liquid feed. Some included studies that evaluated probiotics with expected rumen effects mixed them into milk (Kong et al., 2019; Takemura et al., 2020). This likely implies probiotics bypassing the rumen, reticulum, and omasum through the esophageal groove (Ørskov, 1972), and directly reaching the abomasum. Calf probiotics aiming to promote rumen development may be more effective if incorporated into starter (Diao et al., 2019), although Stefafiska et al. (2021) hypothesized that probiotics that bypass the rumen may contribute to ruminal development by alteration of small intestine metabolism.

**Outcomes.** Although the definition of “probiotics” is linked to health benefits (Hill et al., 2014), weight gain was the most evaluated outcome across trials. Calf growth performance may be an indirect assessment of overall calf health status, supported by the reported negative relationship between health status and growth performance (Shivley et al., 2018). Feed efficiency (calculated based on feed intake and calf weight), which potentially holds more information than ADG and feed intake, was the third most reported outcome. Feed efficiency is an economically important outcome, because feed represents the main cost when raising a heifer from birth to calving (Heinrichs et al., 2013; Boulton et al., 2017). In dairy calves, the gut microbiota may play a role in feed efficiency by increasing the availability of energy substrates (VFA) and essential nutrients (AA and vitamins) from the diet (Elolimy et al., 2020).

The most assessed health outcome across studies was fecal consistency score, alone or as a tool to identify diarrhea. However, there was disagreement across studies on references used for fecal consistency scoring and a variation in assessment frequency. These unstandardized measurements make the comparison among trials challenging. Training for evaluation of fecal consistency was reported in only 2 trials included in our ScR. In a small study, Steen Pedersen et al. (2011) reported inter-evaluator discrepancies in fecal consistency score across swine veterinarians (kappa = 0.24), partially attributed to the previous experience and knowledge of evaluators. Variation between evaluators has been reported even for objective measurements, such as rectal temperature (Naylor et al., 2012). Therefore, training the evaluators, independently of their previous knowledge, will increase the repeatability of measurements. Fecal consistency score was validated to assess fecal DM in dairy calves by Renaud et al. (2020), and the authors recommended that research using this metric should assess internal and external reliability to guarantee repeatability. Furthermore, agreement on a fecal consistency scoring system would facilitate study comparisons and summarization using evidence synthesis methods.

Our ScR identified some limitations in the selected studies; these include inconsistency of outcomes measured (e.g., fecal consistency score with different scoring systems) and incomplete data reporting (e.g., missing breed or age). Efforts have been made to decrease the
inconsistency in methods and data reporting. In 1977, a standardized guideline was published to orient researchers designing calf experiments (Larson et al., 1977), and the Reporting Guidelines for Randomized Controlled Trials for Livestock and Food Safety (REFLECT) was launched in 2010 to provide a checklist of essential items that authors should report when publishing their study (O’Connor et al., 2010). Additionally, McNamara et al. (2016) reviewed key aspects of experimental design and data reporting, and encouraged data sharing in support of animal systems modeling research. More recently, in human health research, the Core Outcome Measures in Effectiveness Trials (COMET) initiative has been developed, aiming to diminish the heterogeneity in outcome reporting (Williamson et al., 2017). The COMET initiative proposes a Core Outcome Set (COS), a minimum set of outcomes that should be measured and reported in experiments evaluating a specific condition. This ScR has several strengths; to identify the largest number of available studies, the search strategy was designed with librarian support, and language was not restricted to English (12.6% of the studies were not written in English). Also, multiple databases were searched to maximize coverage, and the gray literature was included. To minimize bias, the title and abstract screenings were performed independently by 2 reviewers using pretested forms. To decrease errors, the data extraction was performed twice. This study had some limitations: it focused only on controlled trials; observational studies were not included. Nineteen references were included. To minimize reporting (Williamson et al., 2017). The COMET initiative proposes a Core Outcome Set (COS), a minimum set of outcomes that should be measured and reported in experiments evaluating a specific condition. This ScR has several strengths; to identify the largest number of available studies, the search strategy was designed with librarian support, and language was not restricted to English (12.6% of the studies were not written in English). Also, multiple databases were searched to maximize coverage, and the gray literature was included. To minimize bias, the title and abstract screenings were performed independently by 2 reviewers using pretested forms. To decrease errors, the data extraction was performed twice. This study had some limitations: it focused only on controlled trials; observational studies were not included. Nineteen reports were excluded due to the lack of clarity on cattle production system (dairy or beef). Moreover, the last literature search was performed on August 19, 2021. Thus, relevant recent literature might be missing. Additionally, only one person screened full-text articles and extracted the data, even though data were extracted twice to minimize error.

CONCLUSIONS

This scoping review reveals the breadth of controlled trials evaluating probiotic supplementation in dairy calves. Future research should describe the randomization process, provide sample size justification, identify the evaluators, and follow the available guidelines for evaluating and reporting data.

ACKNOWLEDGMENTS

This project was supported by the California Department of Food and Agriculture Antimicrobial Use Stewardship (Sacramento, CA). The authors have not stated any conflicts of interest.

REFERENCES


