



Evaluating case definitions of respiratory disease in dairy calves: A scoping review

K. J. Churchill,¹ C. B. Winder,¹ H. M. Goetz,¹ D. Wilson,¹ T. Uyama,¹
B. Pardon,² and D. L. Renaud^{1*}

¹Department of Population Medicine, University of Guelph, Guelph, ON, Canada N1G 2W1

²Department of Internal Medicine, Reproduction and Population Medicine, Faculty of Veterinary Medicine, Ghent University, 9820 Merelbeke, Belgium

ABSTRACT

Bovine respiratory disease (BRD) is one of the most common diseases observed in dairy calves in both the pre- and postweaning periods. Despite its common occurrence, no formal synthesis (e.g., scoping review) of the scientific literature has been conducted surrounding diagnosing BRD in dairy calves. As such, the objective of this scoping review was to describe what case definitions for diagnosing BRD in dairy calves have been used in research. An initial search of peer-reviewed journals and the gray literature yielded 12,001 unique articles, of which 307 records containing 358 case definitions were included. Seven categories of diagnostic methods were used to categorize case definitions: clinical scoring systems alone were the most used diagnostic method ($n = 148$, 41%); followed by physical exam parameters alone (i.e., the study did not provide a detailed description of how the calf was assessed for each component of the physical exam [i.e., no description or scoring options reported]), $n = 91$, 25%); physical exam parameters in combination with other diagnostic methods ($n = 40$, 11%); thoracic ultrasound alone ($n = 31$, 9%); clinical scoring system in combination with other diagnostic methods ($n = 23$, 7%); clinical scoring system in combination with thoracic ultrasound ($n = 10$, 3%); and other diagnostic methods ($n = 15$, 4%). Only two-thirds of publications provided a citation for their case definition ($n = 230$, 64%), of which the Wisconsin clinical scoring system was most referenced ($n = 100$, 43%), followed by the California clinical scoring system ($n = 19$, 8%). An attempt was made to compare entire case definitions within each diagnostic method; however, they displayed substantial diversity, which limited the ability to compare them at this level. Instead, counts of individual factors within each case definition

(e.g., individual clinical signs) were reported for each diagnostic method. From all case definitions across the 7 diagnostic method categories ($n = 358$), common factors assessed included cough ($n = 224$, 63%), nasal discharge ($n = 213$, 60%), and body temperature ($n = 142$, 54%). Across all case definitions, additional parameters such as validity were frequently not reported; only 4 (1.1%) were referred to as validated, 9 (2.5%) were being validated in the study, and 28 (7.8%) modified existing case definitions. Additionally, assessment of severity of disease when present (e.g., mild, moderate or severe) was frequently not reported ($n = 336$, 93.9%). Further, 106 (29.0%) reported assessments of accuracy, 66 (18.4%) of which were reports of qualitative oversight (e.g., training of assessors), and 40 (11.1%) of which were related to a quantitative assessment of accuracy (e.g., sensitivity and specificity). Parameters surrounding conducting the assessment were also extracted, for example how often it was conducted (most common: at least daily, $n = 120$, 33.5%) and who conducted it (most common: information not reported, $n = 144$, 40.2%). The findings of this scoping review highlighted the need for a validated, standardized set of criteria for BRD diagnosis in dairy calves, achievable through following the methodology of developing a core outcome set.

Key words: BRD, diagnosis, clinical scoring system, thoracic ultrasound, synthesis

INTRODUCTION

Bovine respiratory disease (BRD), also known as shipping fever or enzootic calf pneumonia, is an umbrella or container term for any infection of the upper and lower respiratory tract in cattle of all production systems (Woolums, 2015; Buczinski and Pardon, 2020; Gaudino et al., 2022). In dairy calves, BRD is one of the most important and common diseases observed in both the pre- and postweaning periods (Windeyer et al., 2014), with a recent systematic review estimating

Received October 5, 2024.

Accepted December 18, 2024.

*Corresponding author: renaudd@uoguelph.ca

The list of standard abbreviations for JDS is available at adsa.org/jds-abbreviations-25. Nonstandard abbreviations are available in the Notes.

Table 1. Search terms used in identifying literature on the definition of BRD in calves using CABI (August 25, 2022) from the scoping review of case definitions for respiratory disease in preweaning dairy calves (search term wildcards indicated by asterisks)

Search line	Search terms	Number of citations
1	Respirator* OR BRD OR BRDC OR “bovine respiratory disease” OR pneumoni* OR “shipping fever” OR “bronchopneumonia” OR “undifferentiated fever”	243,450
2	Calf OR calves OR heifer OR bull OR steer	204,738
3	1 and 2	8,033

that up to 61% of dairy heifer calves experience at least one case of BRD in their first year of life (Buczinski et al., 2021). Respiratory disease is a major reason for antimicrobial use (AMU) in the calf industry, especially in nonreplacement dairy calves due to the required transport and high degree of comingling as they move to calf raising facilities (Bokma et al., 2019). Although AMU has historically been essential to control diseases in calves under such conditions (Pardon et al., 2012), their use should be prudent to mitigate the development of resistant bacteria in both humans and animals (Uyama et al., 2022). Beyond AMU, BRD also has long-term performance consequences including higher odds of mortality and removal before first calving, reduced average daily gain, and reduced milk production during the first lactation (Buczinski et al., 2021).

A major challenge of managing BRD in calves is detection. Respiratory disease does not currently have an accepted gold standard (i.e., best accuracy) antemortem diagnostic method and case definition (Buczinski and Pardon, 2020; Kamel et al., 2024). A case definition is defined as set of criteria used to make a decision as to whether an individual has a specific disease (Basil, 2016). Furthermore, considerable variation exists between and within methods of diagnosis (e.g., clinical scoring systems and thoracic ultrasound [TUS]; Ferraro et al., 2021; Boccardo et al., 2023), with each diagnostic method having its own array of benefits and challenges. For example, clinical scoring systems are disease-specific assessment tools that are intended to be a systematic, quick, and inexpensive approach for disease diagnosis (Berman et al., 2022, Probo and Veronesi, 2022). Thoracic ultrasound is a noninvasive imaging technique that can be used to diagnose BRD through identification of lung consolidation (Ollivett and Buczinski, 2016). Further, it allows for the differentiation between lower (e.g., active bronchopneumonia) and upper respiratory tract infections (Ollivett and Buczinski, 2016), contributing to more prudent use of antimicrobials and ultimately reducing economic losses (Berman et al., 2021). It is considered the most accurate tool available to diagnose BRD in calves (Berman et al., 2022; Probo and Veronesi, 2022); however, TUS is inherently more labor-intensive as opposed to diagnosis with clinical scoring systems. Further, it requires equip-

ment and training to conduct and interpret, which may not be feasible or preferable for widespread on-farm use by producers (Love et al., 2014; Maier et al., 2019). Regardless of these potential hurdles, TUS has become a popular tool in research due to its ability to diagnose subclinical BRD, an important component of the disease that makes up a significant proportion of overall BRD cases (Jourquin et al., 2023).

As described above, a gold standard case definition is needed for antemortem BRD diagnosis in preweaning dairy calves. This can be established by following the rigorous, validated process of developing a core outcome set (COS; Williamson et al., 2017). The protocol for developing a COS on diagnosing BRD in preweaning dairy calves led by the first author on this scoping review is available on the University of Guelph Atrium (see Notes). A priori decisions were made that this COS would be specific to diagnosis with a clinical scoring system, TUS, and a combination of the 2 methods, due to the benefits and challenges outlined earlier. A combination category was included to balance the benefits and challenges of each individual diagnostic method and to facilitate early detection and accurate differentiation of clinical and subclinical BRD (Cuevas-Gómez et al., 2021).

As outlined in the protocol, an evidence synthesis project is generally the first step in developing a COS (Williamson et al., 2017). A scoping review was chosen to address this topic of BRD case definitions, as scoping reviews aim to identify knowledge gaps, scope a body of literature, and clarify concepts (Munn et al., 2018). Scoping reviews are appropriate for use when it is unclear what other, more specific questions can be posed and addressed by a systematic review (Armstrong et al., 2011). Scoping reviews can be conducted to examine and clarify definitions used in the literature, describing similarities and differences between case definitions (Munn et al., 2018). It is appropriate to scope the components of case definitions and identify if they're similar enough to endeavor comparing outcomes of those case definitions (i.e., in a systematic review and meta-analysis). The aim of a scoping review aligns with the aim of establishing a COS, as it provides a wholistic overview of all BRD case definitions used in the literature and provides a comprehensive

starting point. As such, the objective of this scoping review was to describe the literature examining BRD in young calves. Specifically, how it has been assessed in the literature (i.e., research), including describing case definitions used and diagnostic methods evaluated. This review aims to provide a first step in developing a consensus definition of BRD (i.e., COS).

METHODS

This manuscript was reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (**PRISMA-ScR**) guidance document (Tricco et al., 2018). A completed PRISMA-ScR checklist is available in Supplemental File S1 (see Notes).

Protocol and Registration

A formal guidance document for scoping review protocols does not currently exist; therefore, the protocol was developed following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Protocols, when appropriate (Moher et al., 2015). It can be found on the University of Guelph Atrium platform (see Notes). The protocol was published on September 1, 2022.

Eligibility Criteria

Studies were included in the scoping review if they were primary research articles published in English and available online with at least 500 words. Further, they had to provide a case definition for BRD in male or female dairy, dual-purpose, or dairy-beef calves at least once before 6 mo of age. Articles were excluded if they did not describe a BRD definition at the calf level, meaning there was no clear assessment of individual calves with a clear cut-off for diagnosing a case. For this scoping review, we included all definitions for BRD, inclusive of broad definitions as well as those specified as clinical or sub-clinical in nature, and of upper or lower respiratory tract infections. For example, studies that assessed groups of calves but did not provide individual calf information were excluded, as were studies that only diagnosed BRD based on postmortem evaluation.

Information Sources

Seven electronic databases were searched for relevant published and gray literature: Medline (via Ovid), Agricola (via ProQuest), CAB Direct (via CABI), Science Citation Index (via Web of Science), Conference Proceedings Citation Index (via Web of Science), Emerg-

ing Sources Citation Index (via Web of Science), and ProQuest Dissertation and Theses. No publication date restriction was applied, aside from that of the databases themselves. Conference proceedings were not searched, as previous reviews have shown that the limited word count of conference proceedings would make it highly unlikely that enough usable data would be provided to meaningfully contribute to the review (Sargeant et al., 2019a,b; Winder et al., 2019a,b,c).

Literature Search

The original search was run on September 1, 2022. After this, the strategy was evaluated for sensitivity by ensuring the identification of 20 relevant articles, which were identified by the second author and 2 research assistants (refer to the protocol [see Notes] for more information). As per the Cochrane Handbook, the search was updated on August 31, 2023, to ensure the search was run within 12 mo of submission for publication (Cumpston, 2023). However, upon final review of the manuscript (summer 2024) it was identified that a key relevant paper published in early 2023 was not picked up by the search update. The search update was re-run in CABI Digital Library on August 13, 2024 (CAB Direct migrated to CABI Digital Library in winter 2024), and in total, 3 additional publications reporting 6 case definitions were identified. The same search string was used in all 3 instances, which involved terms related to dairy calves and respiratory disease and is shown in Table 1 (as applied in CABI and then CABI Digital Library in 2024 due to CABI migrating to CABI Digital Library).

Data Management

Search results were uploaded into Endnote X9 (Clarivate Analytics, Philadelphia, PA) and subsequently de-duplicated using internal algorithms. Citations were then uploaded into DistillerSR (Evidence Partners Inc., Ottawa, Canada), a review management software with artificial intelligence capabilities. A second round of automated de-duplication was completed immediately following upload, and manual de-duplication was conducted throughout both rounds of citation screening.

Selection of Sources of Evidence

DistillerSR was used to identify relevant articles using 2 stages of citation screening (Figure 1), with disagreements resolved by consensus. For the first round of screening, 2 independent reviewers evaluated the title and abstract of citations for relevancy using the following 3 questions:

1. Is the title or abstract written in English? (answer options: “yes,” “no,” “unclear”)
2. Does the title or abstract indicate as a primary research study? (answer options: “yes,” “no,” “unclear”)
3. Does the title or abstract describe or mention BRD in dairy or veal or dual-purpose or dairy-beef calves ≤ 6 mo of age? (answer options: “yes,” “no,” “unclear”)

These questions were piloted on 100 randomly selected citations by all 5 reviewers (KJC, HMG, DJW, TU, JM),

the principal investigator (DLR), and a co-author with experience in evidence synthesis (CBW). No changes were made to the questions following the pilot testing. Thereafter, citations were independently screened by 2 reviewers (HMG, KJC, DJW, TU, or JM). If the reviewers agreed that the answer to any question was “no,” the citation was not assessed at full-text screening. Consensus was at the form level: if both reviewers excluded a citation, the reason for exclusion did not need to match; only discrepancies where one reviewer included a citation and the other excluded it were flagged as conflicts that needed to be discussed by reviewers.

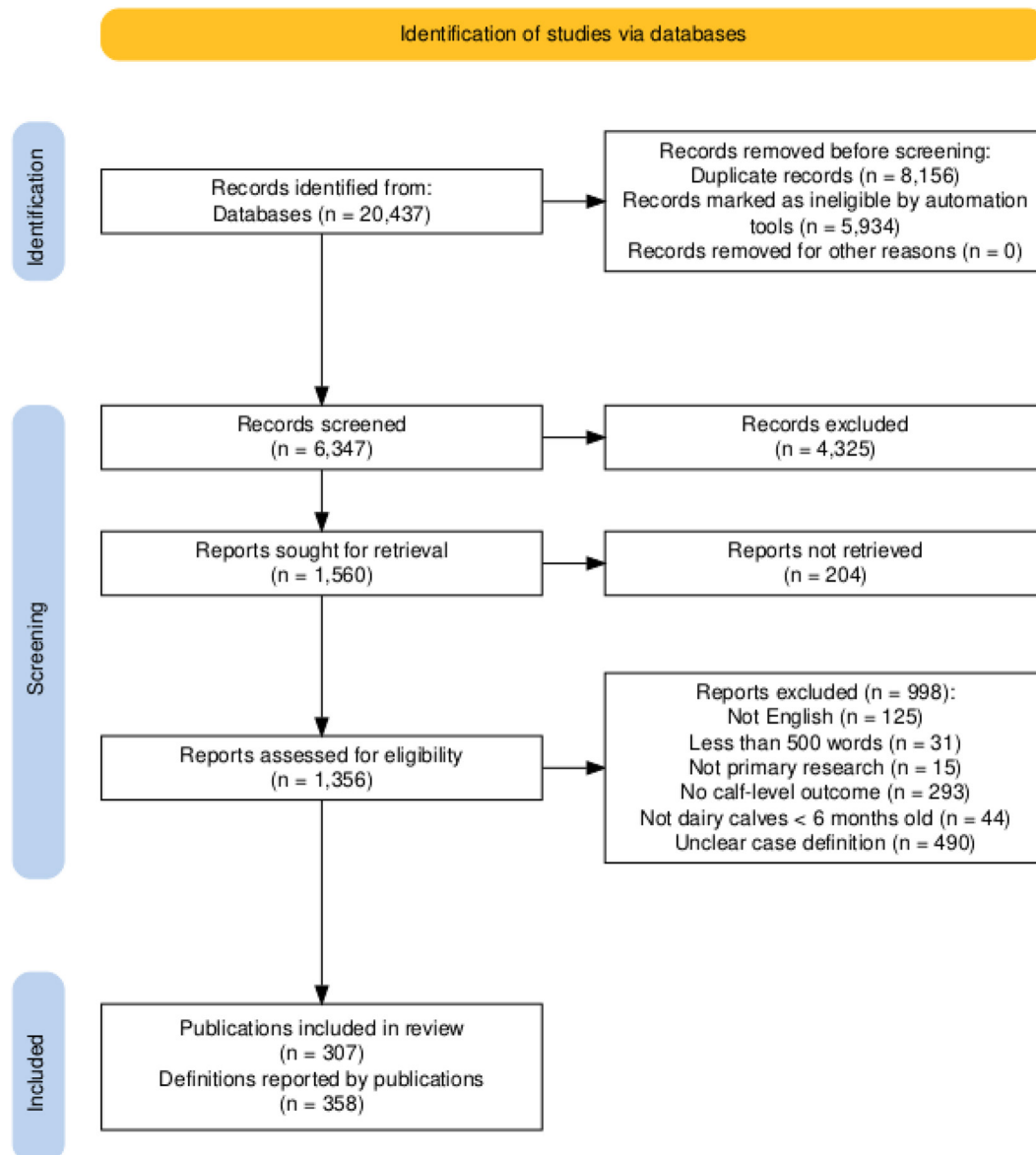


Figure 1. A Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) flowchart (Page et al., 2021), created using an online tool and R package (Haddaway et al., 2022) for the scoping review of case definitions for BRD dairy calves.

After an initial training set of 1,000 citations where 2 reviewers evaluated each citation and agreed on inclusion or exclusion, DistillerSR's internal machine learning ranking program (Artificial Intelligence Preview and Rank) was implemented to reduce the number of citations for initial screening, following methods similar to previous studies (Sargeant et al., 2021; Wilson et al., 2023). The ranking program automatically re-ranked citations for presentation to the reviewers based on the probability that the citation would be relevant. As outlined in the protocol, after screening 500 consecutively ranked citations with no additional eligible articles identified, further screening was not conducted, as it was assumed that all, or almost all, eligible references would have been identified.

For citations where the response to all screening questions was "yes" or "unclear," full-text articles were acquired through the University of Guelph's library resources. A screening form to evaluate eligibility based on information in the full-text articles was developed in DistillerSR and was pre-tested by all reviewers on 10 citations. Thereafter, full-text articles were independently screened by 2 reviewers (HMG, KJC, DJW, TU, or JM) using the following questions, with any disagreements resolved by consensus or consultation with a third reviewer (DLR or CBW):

- (1) Is the full text available? (answer options: "yes" or "no")
- (2) Is the full text available in English? (answer options: "yes" or "no")
- (3) Does the full text include at least 500 words? (answer options: "yes" or "no")
- (4) Does the study describe a primary study? (answer options: "yes" or "no")
- (5) Does the study investigate BRD at a calf level with at least one BRD-related outcome/exposure measured in dairy, veal, dual-purpose or dairy-beef calves, or calves where breed was not reported? (exclude beef studies, studies not in live animals, answer options: "yes" or "no")
- (6) Does the study investigate calves assessed for BRD at least once at ≤ 6 mo of age? (e.g., exclude studies without information on age or poorly reported such as "post-weaned" only, answer options: "yes" or "no")
- (7) Does the study include a definition or clinical signs specific to BRD? (e.g., scoring system, BRD clinical signs evaluated, answer options: "yes" or "no")

Questions were answered sequentially until a "no" was identified, or until all 7 questions were answered "yes," at which point the citation was included for data extraction. Consensus was at the question level, where reviewers had

to exclude citations on the same question (i.e., for the same reason).

Data Charting Process

Two reviewers (HMG, KJC, DJW, TU, or JM) charted data independently and in duplicate using DistillerSR. Reviewers were trained by CBW and DLR by pretesting 4 articles to ensure clarity of the form and consistency of application. In the event a citation reported more than one BRD case definition, the citation was duplicated and data pertaining to each case definition was individually extracted. Conflicts were resolved by consensus. Authors were not contacted for clarification or additional information.

Data Items

A full list of data items extracted can be found in the protocol (see Notes). Data extracted included the following:

- Study information: location of study, year of publication, and study design and objectives;
- Population information: sample size, type of farm, production system, breed, calf age at BRD assessment, and calf sex; and
- Information relevant to the BRD case definition: diagnostic tools and characteristics or components used (e.g., specific clinical signs); cut-off to define a case; any citation referenced as part the definition; the individual(s) who performed the diagnosis; insight to disease severity (e.g., mild, moderate, severe); frequency of assessment; any methods to determine validity of diagnosis; and method to determine the validity of diagnosis (i.e., qualitative [e.g., veterinary oversight] or quantitative [e.g., kappa]), if the case definition was referred to as "validated" or "modified" (post-hoc item; i.e., protocol deviation).

Critical Appraisal of Individual Sources of Evidence

As this is a scoping review, a critical appraisal of the included literature was not conducted.

Synthesis of Results

Data were exported from DistillerSR to Microsoft Excel (Microsoft 365, version 2312) and summarized descriptively. Relevant characteristics (e.g., number of studies by type of farm) were synthesized using pivot tables in Microsoft Excel. The results were presented

Table 2. Diagnostic method categories and definitions used to categorize the 358 case definitions from the scoping review of case definitions for respiratory disease in preweaning dairy calves

Diagnostic method category	Definition for this scoping review
(1) Clinical scoring systems alone	Disease-specific assessment tools that are intended to be a systematic, quick, and inexpensive approach for disease diagnosis (e.g., Wisconsin clinical scoring system; Berman et al., 2022; Probo and Veronesi, 2022)
(2) Clinical scoring systems in combination with other diagnostic methods	(1) in combination with at least one other diagnostic method other than thoracic ultrasound (e.g., Wisconsin clinical scoring system, plus review of farm treatment records)
(3) Clinical scoring systems in combination with thoracic ultrasound	(1) in combination with thoracic ultrasound (e.g., Wisconsin clinical scoring system, plus lung consolidation depth of ≥ 3 cm)
(4) Physical exam parameters alone	The study did not provide a detailed description of how the calf was assessed for each component of the physical exam (i.e., no description or scoring options reported [e.g., “the calf was assessed for cough, fever and nasal discharge”])
(5) Physical exam parameters in combination with other diagnostic methods	(4) in combination with at least one other diagnostic method (e.g., “the calf was assessed for cough, fever and nasal discharge,” plus thoracic ultrasound)
(6) Thoracic ultrasound alone	Noninvasive imaging technique that can be used to diagnose BRD through identification of lung consolidation (Ollivett and Buczinski, 2016)
(7) Other diagnostic methods	Any diagnostic method other than those described in categories (1) to (6), for example, farm treatment records alone, or lung imaging other than ultrasound

following recommendations by PRISMA-ScR (Tricco et al., 2018).

Due to the high number and diversity of diagnostic methods, 7 diagnostic method categories were created (definitions for each category are provided in Table 2). These categories were as follows:

- (1) Clinical scoring system (i.e., disease-specific assessment tools that are intended to be a systematic, quick, and inexpensive approach for disease diagnosis [e.g., Wisconsin clinical scoring system; Berman et al., 2022; Probo and Veronesi, 2022]) alone;
- (2) Clinical scoring system in combination with other diagnostic methods (e.g., farm treatment record review); or
- (3) Clinical scoring system in combination with TUS;
- (4) Physical exam parameters (i.e., the study did not provide a detailed description of how the calf were assessed for each component of the physical exam [i.e., no description or scoring options reported]) alone;
- (5) Physical exam parameters in combination with other diagnostic methods;
- (6) Thoracic ultrasound (noninvasive imaging technique that can be used to diagnose BRD through identification of lung consolidation; Ollivett and Buczinski, 2016) alone; and
- (7) Other diagnostic methods alone (any diagnostic method other than those described in categories (1) to (6), for example farm treatment records alone, or lung imaging other than ultrasound).

For this scoping review, physical exam parameters differed from clinical scoring systems, as physical exam parameters did not describe numeric scoring options

or criteria (e.g., 0 to 3, where 0 = absence of clinical signs), nor was information provided related to using the sum of assessment scores to determine the presence or absence of BRD. Additionally, clinical scoring systems may be reported under physical exam parameters if scoring details (i.e., scores and descriptions) of the scoring system were not provided by the publication. The “other” category was an umbrella category for all diagnostic methods other than the categories listed above, which included “based on another study” (scoring system not apparent), hematological parameters, farm treatment records, lung imaging other than ultrasound (e.g., thoracic radiography), or lung auscultation.

The primary objective of this review was to summarize entire case definitions (i.e., comparison of all components [e.g., clinical signs], scales, and cut-points) of BRD by diagnostic method. However, definitions within each diagnostic method category displayed significant diversity. Therefore, comparison of entire case definitions (e.g., these 5 clinical signs vs. these 5 clinical signs, with these scales [i.e., scoring options], descriptions and cut-offs for diagnosis) was not possible. Instead, individual components of case definitions were compared (i.e., how many times were individual clinical signs [e.g., cough] were reported across case definitions).

RESULTS

Study Characteristics

Selection of Sources of Evidence. Of the initial retrieval of 20,437 articles across all databases, de-duplication yielded 12,001 unique publications (records) which were then screened for relevance using the title and abstract. A total of 1,353 publications were then included for full-text screening, following which 310 relevant publications

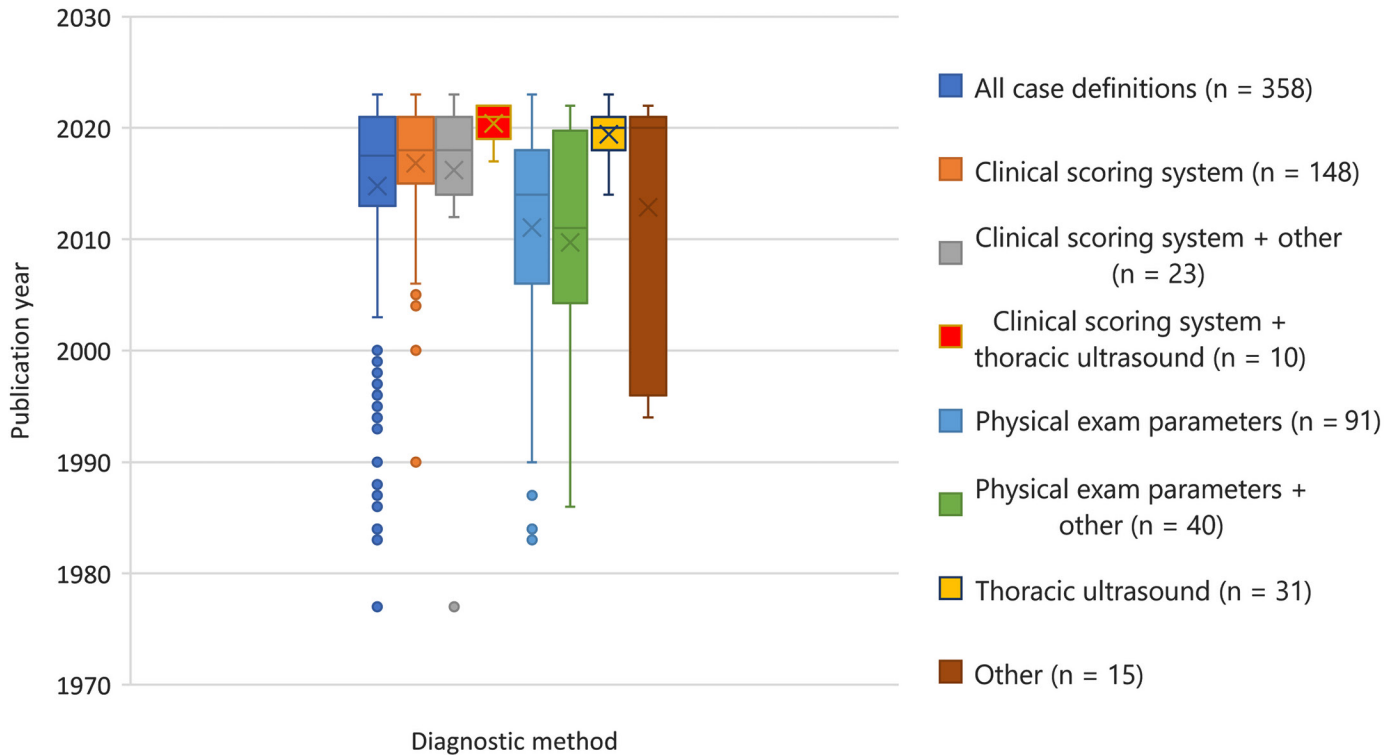


Figure 2. Publication year (minimum, first quartile, median, third quartile, maximum, and outliers) by diagnostic method for 358 publications included in the scoping review of case definitions for BRD dairy calves. Minimum = lower reach of error bar, first quartile = bottom of box to median line (the area marked with X), median = horizontal line in box, third quartile = top of box above median line, maximum = top reach of error bar, and outliers = individual circles outside the error bar reach.

were included in the scoping review, composed of 294 peer-reviewed publications (94.8%) and 16 academic theses (gray literature, 5.2%). These 310 publications reported 358 BRD case definitions, as 5 academic theses and 32 peer-reviewed publications reported multiple relevant case definitions. As the aim of this scoping review was to synthesize and assess BRD case definitions, 358 is used as the denominator throughout the synthesis. Extracted data were exported from DistillerSR into a Microsoft Excel spreadsheet for the descriptive results to be characterized. The flow of studies through the screening process including reasons for the full-text exclusions, are shown in Figure 1, generated using a PRISMA flow generator (Haddaway et al., 2022). Refer to Supplemental File S2 (see Notes) for the list of publications excluded at full-text screening and Supplemental File S3 (see Notes) for the list of included publications.

Characteristics of Sources of Evidence. Included publications were published between 1977 and 2023, with a median publication year of 2017 (interquartile range [IQR]: 2013 to 2020) and mode (i.e., most common) publication year of 2021 (Figure 2). Included publications were conducted in developed (n = 289, 77.7%), developing (n = 43, 11.6%), and in-transition economies (n = 4,

1.1%; UN DESA, 2024), although some studies did not report country of conduct (n = 36, 9.7%; Supplemental File S4, see Notes). Across these economies, 39 countries were reported, with almost 90.0% of studies conducted in the developed economies of the United States (n = 129, 34.7%), the European Union (n = 83, 22.3%), and Canada (n = 44, 11.8%; Supplemental File S4). Seven publications were conducted in multiple countries: 4 publications were conducted in 2 countries, 2 publications were conducted in 3 countries, and 1 publication was conducted in 7 countries (Supplemental File S4). Most studies were observational (n = 207, 57.8%) or controlled trials (n = 102, 28.5%), with some challenge (n = 30, 8.4%) and diagnostic accuracy studies (n = 19, 5.3%; Supplemental File S4). Publications reported between 1 and 3 objectives (mean = 2), resulting in 407 total objectives. The most common objectives were to evaluate the effects of an intervention (n = 113, 27.2%) and evaluate disease or mortality risk factors (n = 102, 24.6%), alone or in combination with at least one other listed objective (Figure 3, Supplemental File S5, see Notes).

Characteristics of Included Populations. As per the inclusion criteria for this scoping review, calves in included publications were ≤ 6 mo of age, and of dairy

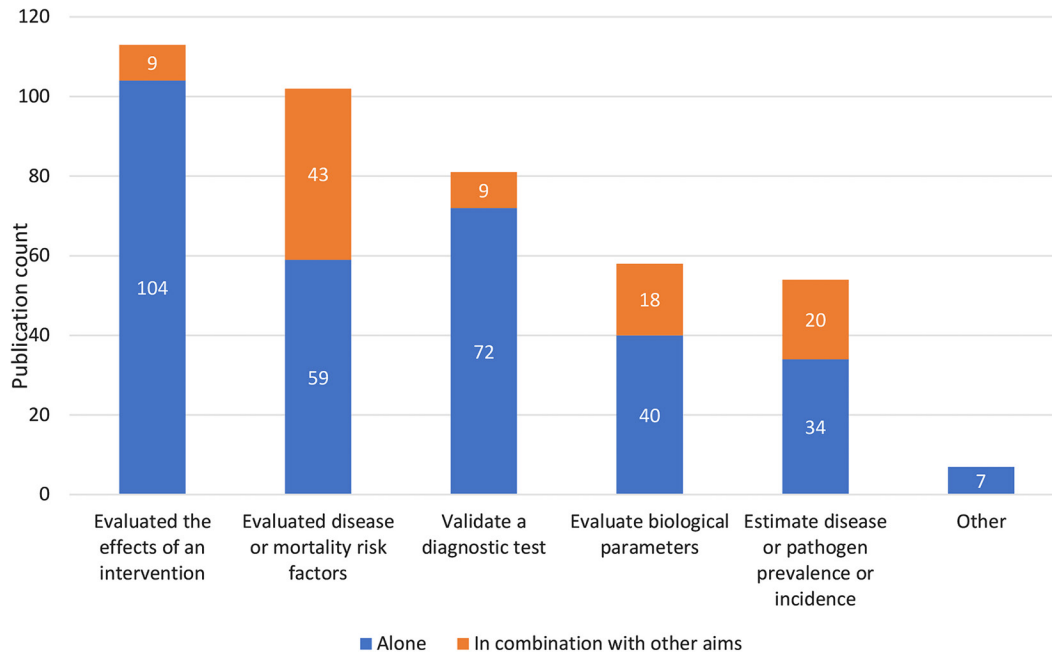


Figure 3. Distribution of objectives of included studies from the scoping review of case definitions for BRD in dairy calves. The 358 publications included reported between 1 and 3 aims (mean: 2), resulting in a total of 415 aims.

($n = 342$, 95.5%), breed not reported ($n = 7$, 1.9%), cross-bred ($n = 5$, 1.4%), other breeds ($n = 2$, 0.6%), or mixed-purpose breeds ($n = 2$, 0.6%). Across all breeds, almost a third of calves each were mixed sex ($n = 120$, 33.5%) or female only ($n = 111$, 31.0%). Some studies did not report sex ($n = 71$, 19.8%) or were male-only ($n = 56$, 15.6%). For studies using dairy breed calves ($n = 342$, 95.5%), sex of calves included mixed sex ($n = 115$, 33.6%), female only ($n = 110$, 32.2%), sex not reported ($n = 64$, 18.7%), or male only ($n = 53$, 15.5%). For studies where breed of calves was not reported ($n = 7$, 1.9%), sex of calves was often not reported ($n = 6$, 1.8%), and one study used female calves (0.1%). For studies with mixed-purpose or crossbreed calves ($n = 5$, 1.4%), sex of calves included male ($n = 2$, 0.6%), mixed sex ($n = 2$, 0.6%) or sex not reported ($n = 1$, 0.3%). For the 2 studies using other breeds (0.6%), sex was not reported in either. For the 2 studies using mixed-purpose breeds (0.6%), calves were of mixed sex ($n = 1$, 0.3%) or sex was not reported ($n = 1$, 0.3%; Supplemental File S6, see Notes).

Almost half of publications were conducted on commercial dairy farms ($n = 172$, 48.0%), with some research ($n = 62$, 17.3%), commercial calf-rearing ($n = 52$, 14.5%), or mixed facilities ($n = 38$, 10.6%) being used. A few publications did not report this information ($n = 24$, 6.7%), were conducted in veterinary hospitals ($n = 8$, 2.2%), or took place in another facility type ($n = 2$, 0.6%; Supplemental File S6).

Synthesis of Results

The 358 included BRD case definitions were sorted into 7 diagnostic methods categories for synthesis (Table 2): (1) clinical scoring system alone ($n = 148$, 41.3%), (2) in combination with other diagnostic methods (e.g., farm treatment record review, $n = 23$, 6.4%), or (3) in combination with TUS ($n = 10$, 2.8%); (4) physical exam parameters alone ($n = 91$, 25.4%) or (5) in combination with other diagnostic methods ($n = 40$, 11.1%); (6) TUS alone ($n = 31$, 8.6%); and (7) other diagnostic methods alone ($n = 14$, 3.9%; Figure 4, Supplemental File S7, see Notes).

Across all case definitions ($n = 358$), an assessment of severity of disease when present was frequently not reported ($n = 336$, 93.9%), but some 2-level ($n = 13$, 3.6%), 3-level ($n = 7$, 2.0%), and 4-level ($n = 1$, 0.3%) severity classifications were reported. Only 4 case definitions (1.1%) were referred to as validated (all of which were referencing the California clinical scoring system), 9 (2.5%) were being validated by the publication they were extracted from, 28 (7.8%) were modifying an existing case definition, and the remainder did not mention validity of the case definition ($n = 317$, 88.5%). Further, 106 (29.0%) reported assessments of accuracy, 66 (18.4%) of which were reports of qualitative oversight (e.g., training of BRD assessors), and 40 (11.1%) of which were related to a quantitative assessment of accuracy (e.g., sensitivity [Se] and specificity [Sp]). It

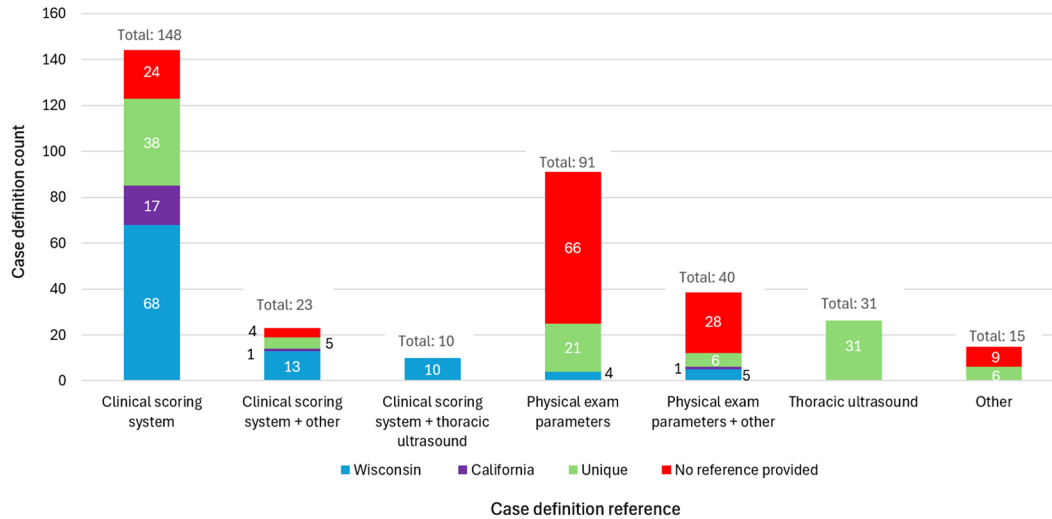


Figure 4. Distribution of diagnostic methods and composition of case definition references from the scoping review of 358 case definitions for BRD in preweaning dairy calves. Where Wisconsin = McGuirk and Peek (2014) or McGuirk (2008), California = Love et al. (2014), and unique = a reference used by ≤ 3 case definitions.

was frequently not reported who conducted the calf assessment ($n = 144$, 40.2%); however, a veterinarian ($n = 64$, 17.9%), researcher ($n = 61$, 17.0%), farmer ($n = 56$, 15.6%), multiple assessors from the previously mentioned stakeholder categories ($n = 29$, 8.1%), and other assessors ($n = 4$, 1.1%) were also reported. Assessments for BRD were conducted at least daily ($n = 120$, 33.5%), only once ($n = 76$, 21.2%), frequency not reported ($n = 67$, 18.7%), 1 to 6 times weekly ($n = 54$, 15.1%), or another, longer frequency ($n = 41$, 11.5%).

Clinical Scoring System in Combination with Other Diagnostic Methods. Twenty-three case definitions used a clinical scoring system in combination with other diagnostic methods ($n = 12$, 52.2%). Most case definitions provided a reference (Figure 4) for at least one of the diagnostic methods ($n = 19$, 82.6%). The the Wisconsin clinical scoring system was most frequently referenced for the clinical scoring system, but no reference was provided for the “other” diagnostic method ($n = 13$, 56.5%) portion of the case definition. Individual or combinations of citations not referenced by other publications in this scoping review ($n = 3$, 13.0% and $n = 2$, 8.7%, respectively) and the California clinical scoring system (Love et al., 2014; $n = 1$, 4.3%) were referenced for the clinical scoring system, without a reference for the “other” diagnostic method. The median publication year of studies using clinical scoring systems in combination with other diagnostic methods was 2018 (IQR: 2014 to 2021; Figure 2). Comparisons of individual components of case definitions were made: the number of clinical signs ranged from 1 to 7, with a median of 4 clinical signs (IQR: 2 to 5). Further, the number of levels a clinical sign could be assigned ranged from 1 to 4, with a median of 4 levels

(IQR: 2 to 4). As evident through the variability in number of clinical signs and number of levels each clinical sign could be assigned, a synthesis of total scoring cut-offs to define a case would not be meaningful. Eleven clinical sign categories were reported by these case definitions (Figure 6), with the most reported clinical signs being cough ($n = 21$, 91.3%), nasal discharge ($n = 18$, 78.3%), and body temperature ($n = 18$, 78.3%). Most of the publications did not report if the combination of clinical scoring systems and other diagnostic methods were referred to as validated ($n = 16$, 69.9%), provide insight into disease severity when present ($n = 21$, 91.3%), or report the accuracy (e.g., Se and Sp) of the case definition ($n = 16$, 69.9%). For almost half of case definitions, the person who conducted the assessment was not reported ($n = 9$, 39.1%); however, veterinarians ($n = 5$, 21.7%) or a combination of researchers and farmers ($n = 4$, 17.4%) were also reported. Most often, calves were assessed for BRD at least once daily ($n = 9$, 39.1%) or only a single time ($n = 5$, 21.7%). A comprehensive overview of the details of the case definitions using a clinical scoring system in combination with other diagnostic methods can be found in Table 7C of Supplemental File S7.

Clinical Scoring System in Combination with TUS. Of the 10 case definitions using a clinical scoring system in combination with TUS, all referenced the Wisconsin clinical scoring system, and most provided a citation not used by other included publications for the TUS methodology ($n = 7$, 70.0%; Figure 4, Supplemental File S7). The 5 clinical signs of the Wisconsin clinical scoring system were nasal discharge, cough, ocular discharge, body temperature, and ear position (Table 2). For the TUS component, lung consolidation surface area was

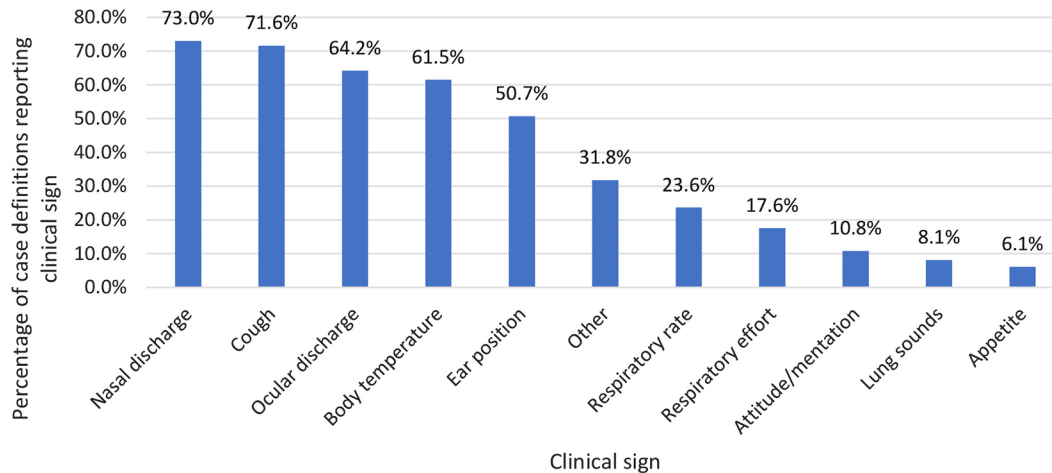


Figure 5. Components of the 148 case definitions using clinical scoring systems from the scoping review of case definitions for BRD in preweaning dairy calves. Other clinical signs were diverse but included general state, lacrimal consistency, hematological parameters, and tachycardia.

most often used (≥ 1 cm²: $n = 4$, 40.0%; ≥ 2 cm²: $n = 1$, 10.0%; ≥ 3 cm²: $n = 3$, 30.0%; Supplemental File S7). The median publication year of studies using clinical scoring systems in combination with TUS was 2021 (IQR: 2019 to 2022; Figure 2). None of the publications reported if the referenced diagnostic methodology was validated or provided insight to disease severity when present. In half of the publications, the person who conducted the examination was not reported, and in the other half, a researcher conducted the examinations. Most often, assessments were conducted at least once daily ($n = 4$, 40.0%) or 1 to 6 times weekly ($n = 3$, 30.0%). The accuracy of

the case definition was not reported by most studies ($n = 7$, 70.0%). A comprehensive overview of the details of the case definitions using a clinical scoring system in combination with TUS can be found in Table 7D of Supplemental File S7 (Figure 7).

Physical Exam Parameters Alone. Of the 91 case definitions utilizing physical exam parameters, the majority did not provide a reference ($n = 66$, 72.5%; Figure 4). For the 25 case definitions that provided a reference (27.5%), most used individual citations not used in other included publications ($n = 18$, 19.8%), a few referenced the Wisconsin clinical scoring system but did not pro-

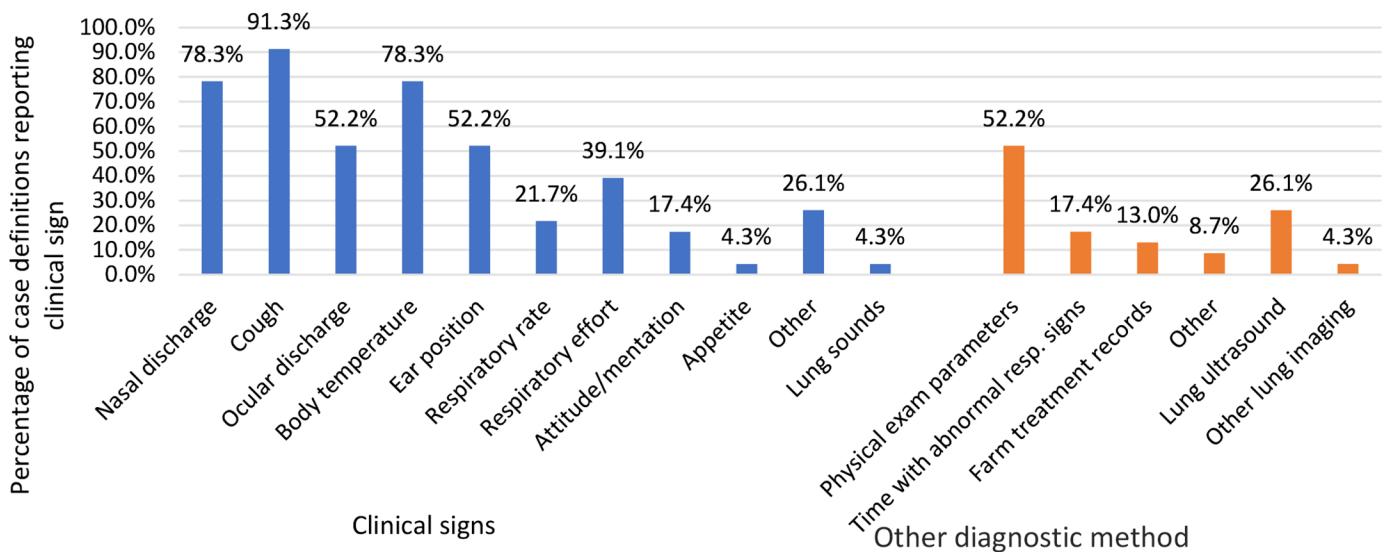


Figure 6. Components of the 23 case definitions using clinical scoring systems in combination with other diagnostic methods (i.e., physical exam parameters, imaging, or both) from the scoping review of case definitions for BRD in preweaning dairy calves. Other clinical signs included hydration and tachypnea, other diagnostic methods included referencing a scoring system without describing it, and other lung imaging included radiography.

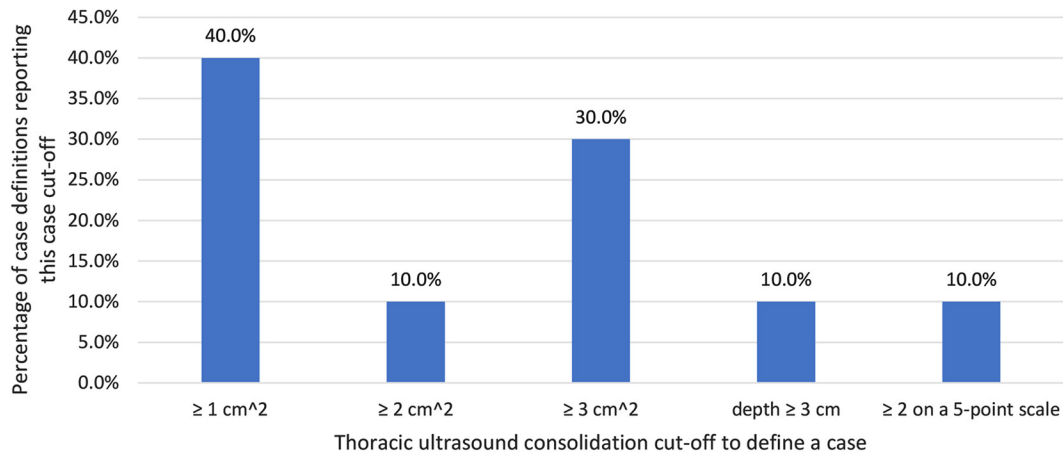


Figure 7. Thoracic ultrasound components of the 10 case definitions using clinical scoring systems in combination with thoracic ultrasound from the scoping review of case definitions for respiratory disease in preweaning dairy calves. In all 10 case definitions, the Wisconsin clinical scoring system was used.

vide scoring details ($n = 4$, 4.4%), and a few referenced combinations of citations not used in other included publications ($n = 3$, 3.3%). The median publication year of studies using physical exam parameters alone was 2014 (IQR: 2006 to 2018; Figure 2). Comparisons of individual components of case definitions were made: the number of clinical signs ranged from 1 to 10, with a median of 5 clinical signs (IQR: 4 to 6). Thirteen clinical signs or other assessment categories were reported by these case definitions (Figure 8), with body temperature ($n = 68$, 74.7%), cough ($n = 66$, 72.7%), nasal discharge

($n = 56$, 63.7%), and respiratory rate ($n = 54$, 59.3%) most commonly reported. Almost all publications did not report if the set of physical exam parameters being referenced were validated ($n = 88$, 96.7%) nor provide insight to disease severity when present ($n = 87$, 95.6%). The person who conducted the physical examination was not reported by many studies ($n = 34$, 37.4%), but farmers ($n = 22$, 24.2%) and veterinarians ($n = 21$, 23.1%) were common. Calves were most frequently assessed for BRD at least once daily ($n = 35$, 38.5%), at a single time point ($n = 22$, 24.2%), or this information was not reported (n

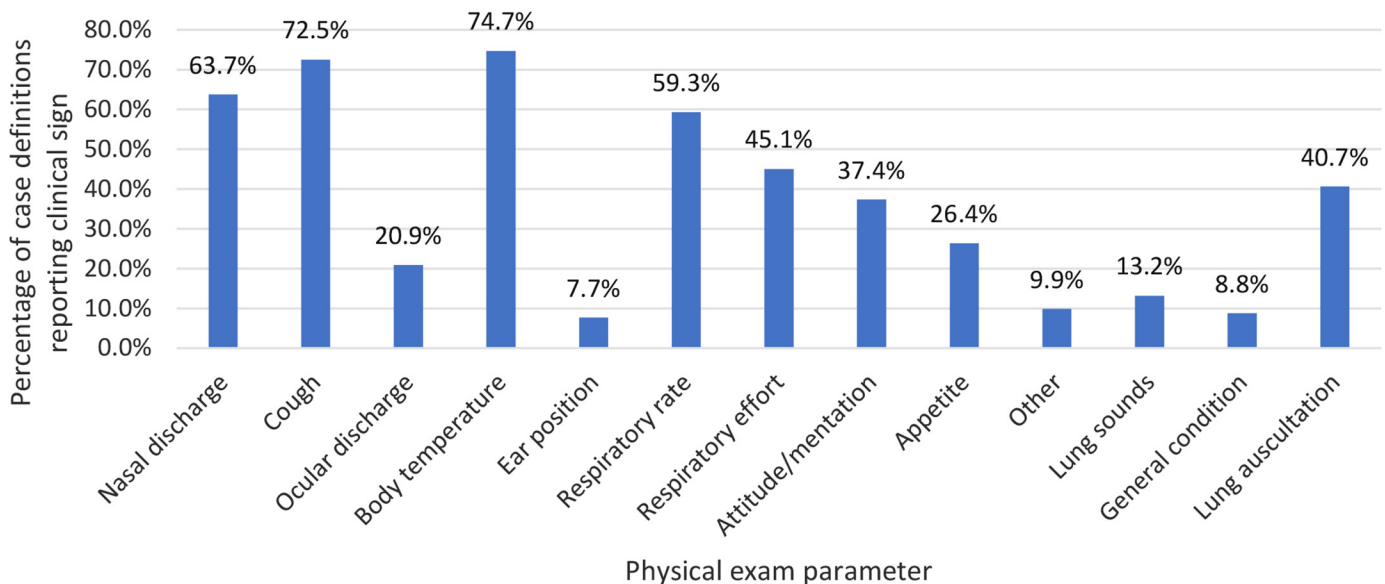


Figure 8. Components of the 91 case definitions using physical exam parameters from the scoping review of case definitions for respiratory disease in preweaning dairy calves. Examples of “other” physical exam parameters included diarrhea and lack of liquid or soft feces in combination with fever and aggravated breathing.

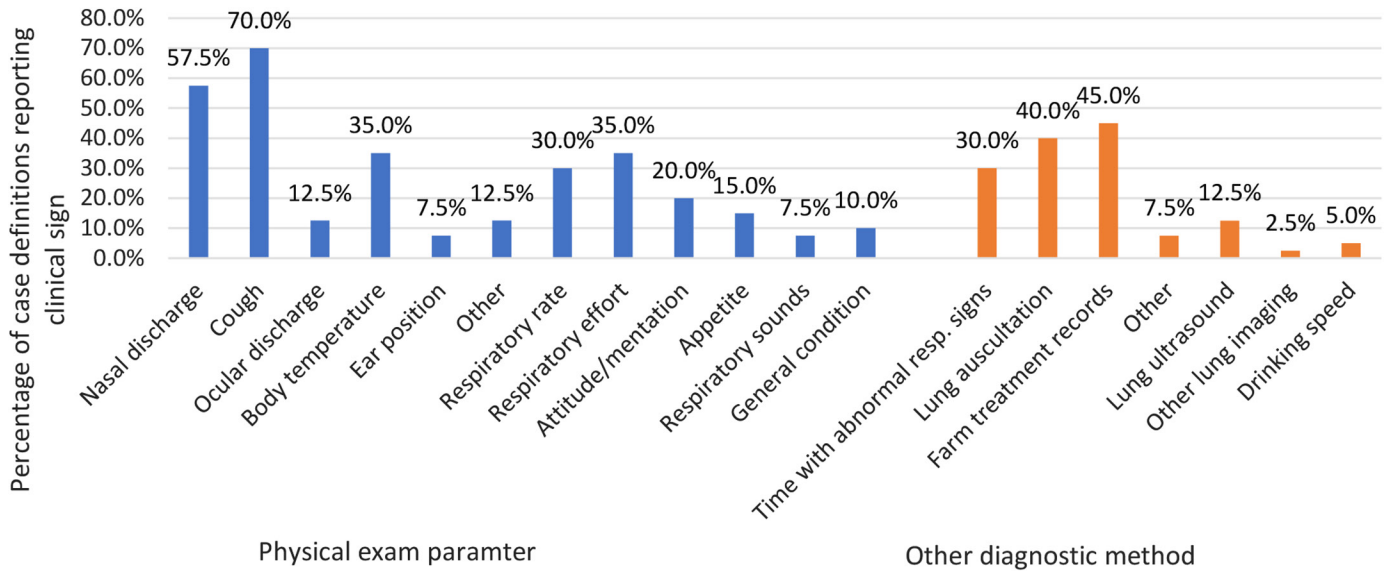


Figure 9. Components of the 40 case definitions using physical exam parameters in combination with other diagnostic methods (e.g., lung auscultation or farm treatment record review) from the scoping review of case definitions for respiratory disease in preweaning dairy calves.

= 17, 18.7%). The accuracy (e.g., Se and Sp) of the case definition used was not reported by most studies ($n = 67$, 73.6%). A comprehensive overview of the details of the case definitions using physical exam parameters can be found in Table 7E of Supplemental File S7.

Physical Exam Parameters in Combination with Other Diagnostic Methods. Of the 40 case definitions using physical exam parameters in combination with other diagnostic methods, the majority did not provide a reference for the case definition ($n = 28$, 70.0%; Figure 4). For the 12 case definitions that provided a reference, some referenced the Wisconsin clinical scoring system but did not provide scoring details ($n = 5$, 41.6%), some used combinations of citations not used in other included publications ($n = 4$, 33.3%), a few referenced an individual citation not used in other included publications ($n = 2$, 16.6%), and 1 referenced the California clinical scoring system but did not provide scoring details (8.3%). The median publication year of studies using physical exam parameters in combination with other diagnostic methods was 2011 (IQR: 2005 to 2019; Figure 2). Comparisons of individual components of case definitions were made: the number of clinical signs ranged from 1 to 10, with a median of 3 clinical signs (IQR: 3 to 5). Twelve clinical sign categories were reported by these case definitions (Figure 8). Seven other diagnostic methods were used in combination with physical exam parameters to diagnose BRD (Figure 9), most commonly farm treatment records ($n = 18$, 45.0%) or lung auscultation ($n = 16$, 40.0%). Almost all of the publications did not report if the combination of physical exam parameters and other diagnostic

methods were considered validated ($n = 38$, 95.0%). Insight to disease severity when present was not reported by most case definitions ($n = 37$, 92.5%). A farmer ($n = 16$, 40.0%) or multiple assessors (e.g., farmer and veterinarian, $n = 11$, 27.5%) most often conducted the examinations. Frequency of BRD assessment was not reported by almost half of studies ($n = 16$, 40.0%), but for those that did report this information, calves were assessed at least once daily ($n = 10$, 25.0%) or used another frequency ($n = 9$, 22.5%). The accuracy (e.g., Se and Sp) of the case definition used was not reported by almost half of studies ($n = 19$, 47.5%). A comprehensive overview of the case definitions using physical exam parameters in combination with other diagnostic methods can be found in Table 7F of Supplemental File S7.

Thoracic Ultrasound Alone. Of the 31 case definitions using TUS alone, all provided a reference (Figure 3). Most commonly, unique citations (combinations [$n = 12$, 38.7%] or individual citations [$n = 6$, 19.4%]) not used in other included publications were reported; however, Ollivett and Buczinski (2016; $n = 10$, 32.7%), and Ollivett et al. (2015; $n = 3$, 9.6%) were also referenced (Figure 4). The median publication year of studies using TUS alone was 2020 (IQR: 2018 to 2021; Figure 2). Most definitions used depth of lung consolidation to define a case, specifically presence of ≥ 1 cm ($n = 13$, 41.9%) or ≥ 3 cm ($n = 6$, 19.4%). The remaining studies used consolidation surface area, such as ≥ 1 cm² ($n = 3$, 19.7%) or ≥ 3 cm² ($n = 1$, 3.2%), or another cut-off parameter ($n = 8$, 25.8%: e.g., a score ≥ 2 [3 to 5 cm of consolidation] on a 4-point scale, or “numerous small comet

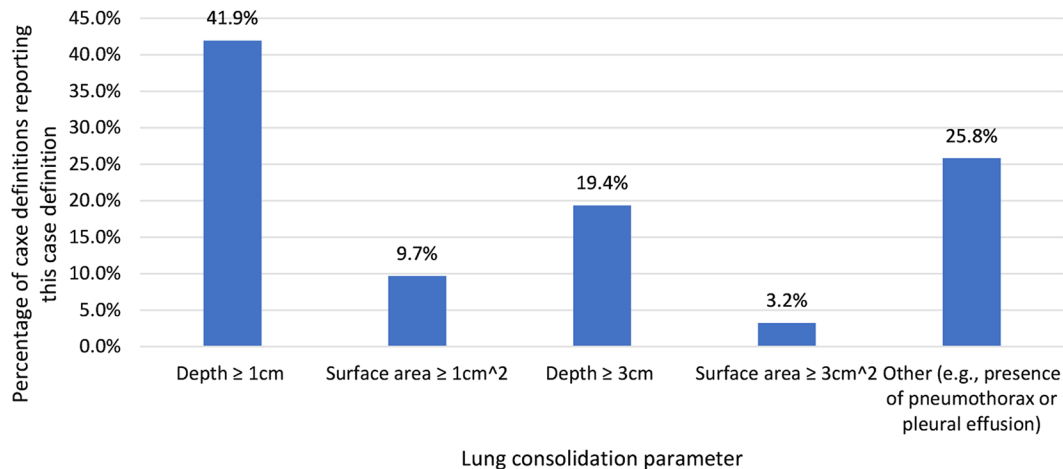


Figure 10. Components of the 31 case definitions using thoracic ultrasound from the scoping review of case definitions for respiratory disease in preweaning dairy calves.

tails in multiple fields, multiple large comet tails, focal or extensive consolidation, or abscesses”; Deepak et al., 2021, p. 3). See Figure 10 and Table 7G of Supplemental File S7 for more details. Almost all of the publications did not report if the referenced TUS methodology was validated ($n = 28$, 90.3%), or provide insight to disease severity when present ($n = 28$, 90.5%). A researcher most often conducted the examinations ($n = 13$, 41.9%), but almost as often this information was not reported ($n = 9$, 29.0%). Most often, calves were assessed for BRD at a single time point ($n = 15$, 48.4%) or another frequency ($n = 7$, 22.6%). The accuracy (e.g., Se and Sp) of the case definition used was not reported by most studies ($n = 19$, 61.3%). A comprehensive overview of the details of the case definitions using TUS alone can be found in Table 7G of Supplemental File S7.

Clinical Scoring Systems Alone. Of the 148 case definitions utilizing a clinical scoring system alone, most provided a reference ($n = 124$, 83.7%; Figure 4). The Wisconsin clinical scoring system (including McGuirk, 2008; McGuirk and Peek, 2014; and respiratory components of the Calf Health Scoring Chart [https://fyi.extension.wisc.edu/heifermgmt/files/2015/02/calf_health_scoring_chart.pdf], which all reference the same clinical signs and scoring structure) was cited in more than half of these case definitions ($n = 68$, 53.1%), followed in frequency by the California clinical scoring system ($n = 17$, 13.2%). Due to the commonality of referencing to these scoring systems, a breakdown of their components is provided in Table 3. For the remaining case definitions for which the publication they were extracted from provided a citation for the case definition, some cited a unique citation (i.e., not otherwise used by other included publications) case definitions ($n = 14$, 10.9%), others reported combinations of unique citations ($n = 10$, 8.6%), and 4 citations were

cited 3 or 4 times each (Heinrichs et al., 2003; McGuirk, 2005; Lago et al., 2006; Berge et al., 2009). Details of these citations can be found in Supplemental File S7. Twenty-four (18.8%) case definitions did not provide a reference. The median publication year of studies using clinical scoring systems alone was 2018 (IQR: 2015 to 2021; Figure 2). The case definitions displayed significant diversity which limited the ability to compare them at the full case definition level. Instead, comparisons of individual components of case definitions were made. The number of clinical signs included in these case definitions ranged from 1 to 14, with a median of 5 clinical signs (IQR: 1 to 6). Further, the number of levels a clinical sign could be assigned ranged from 1 to 5 (e.g., 3 levels: 0 = not present, 1 = mild discharge, 2 = heavy discharge), with a median of 3 levels (IQR: 2 to 4). As evident through the variable number of clinical signs and number of levels each clinical sign could be assigned, a synthesis of total scoring cut-offs to define a case would not be meaningful. Eleven clinical sign categories were reported by these case definitions (Figure 5), with the most commonly reported clinical signs being nasal discharge ($n = 108$, 72.9%), cough (induced or spontaneous, $n = 106$, 71.6%), ocular discharge ($n = 95$, 64.2%), body temperature ($n = 91$, 61.5%), and ear position ($n = 75$, 50.7%). Calves were most commonly assessed at least once daily ($n = 59$, 39.9%) or 1 to 6 times weekly ($n = 34$, 23.0%). The majority of publications did not report the following: if the clinical scoring system was validated ($n = 122$, 82.4%); provide insight to disease severity when present (e.g., mild, moderate, severe; $n = 138$, 93.2%); indicate who conducted the clinical scoring assessment ($n = 76$, 51.3%); or include the diagnostic accuracy of the case definition ($n = 112$, 79%). Among the 14 publications that reported a quantitative assessment of the diag-

Table 3. Wisconsin (McGuirk, 2008; McGuirk and Peek, 2014; Calf Health Scoring Chart) and California (Love et al., 2014) clinical scoring systems comparison, from the scoping review of case definitions for respiratory disease in preweaning dairy calves¹

Clinical sign	Scales and descriptions	
	Wisconsin clinical scoring system	California clinical scoring system ²
Rectal temperature, °C (°F)	0: 37.8–38.2 (100–100.9) 1: 38.3–38.8 (101–101.9) 2: 38.9–39.4 (102–102.9) 3: ≥39.5 (103)	0: <39.2 (102.5) 2: ≥39.2 (102.5)
Cough score	0: No cough 1: Induce single 2: Induce repeated or occasional spontaneous 3: Repeated spontaneous coughing	0: No spontaneous cough 2: Any spontaneous cough
Nasal discharge score	0: Normal, serous discharge 1: Small amount of unilateral, cloudy discharge 2: Bilateral, cloudy, or excessive mucus 3: Heavy ocular discharge, severe head tilt, or bilateral ear droop	0: Normal 4: Any abnormal discharge
Eye score	0: Normal eye 1: Mild ocular discharge 2: Moderate bilateral ocular discharge 3: Heavy ocular discharge	0: Normal 2: Any abnormal discharge
Ear score	0: Normal ear 1: Mild ear flicking 2: Slight unilateral ear drop 3: Severe head tilt or bilateral ear droop	0: No ear droop, head tilt, or ear flick 5: Any ear droop or head tilt
Respiratory quality	NA	0: Normal respiration 2: Abnormal respiration
Defining a BRD case	Total score ≥5 (provided that ≥2 abnormal parameters are observed)	Total score ≥5

¹BRD = bovine respiratory disease, NA = not applicable.

²This categorization comparison of the 2 scoring systems is narratively outlined in the California clinical scoring system paper (Love et al., 2014).

nostic accuracy of the case definition, 5 reported inter-rater reliability (i.e., kappa, agreement beyond chance) and 8 reported Se and Sp (Supplemental File S7). Similar inter-rater reliability was reported for the Wisconsin (3 publications: 0.31 to 0.97) and California (2 publications: 0.49 to 0.87) BRD scoring systems. Similar accuracy was reported for the Wisconsin (Se: 0.26 to 0.94; Sp: 0.26 to 1.0) and California (Se: 0.67 to 0.82; Sp: 0.26 to 0.79) scoring systems. Calves were often assessed for BRD at least once a day (n = 59, 39.9%) or 1 to 6 times weekly (n = 34, 23.0%). A comprehensive overview of the details of the case definitions using a clinical scoring system can be found in Table 7B of Supplemental File S7.

Other Diagnostic Methods. Of the 15 case definitions using other diagnostic methods, two-thirds did not provide a reference (n = 9, 60.0%). Of the references provided, all were individual citations not used in other included publications (n = 6, 40.0%; Figure 4). The median publication year of studies using other diagnostic methods was 2020 (IQR: 2001 to 2021; Figure 2). Almost half of these case definitions used retrospective review of farm treatment records (n = 6, 40.0%), and a few each used lung imaging other than ultrasound (n = 3, 20.0%), hematological parameters (n = 3, 20.0%, i.e., haptoglobin concentration [n = 2, 13.3%] or neutrophil levels [n = 1, 6.7%]), referred to another study, but the scoring method was not apparent (n = 2, 13.3%) or used

farm treatment records in addition to another diagnostic method (n = 1, 6.7%; Figure 11). None (100%) of the publications reported if the diagnostic methodology was validated or provide insight to disease severity when present. A farmer most often conducted the examinations (n = 6, 40.0%), but this information was also often not reported (n = 4, 26.7%). Most often, frequency of assessment was not reported (n = 8, 53.3%), but a single time point (n = 4, 26.7%) was also commonly used. The accuracy (e.g., Se and Sp) of the case definition was not reported by most studies (n = 11, 73.3%).

A comprehensive overview of the details of the case definitions using other diagnostic methods can be found in Table 7H of Supplemental File S7.

DISCUSSION

Summary of Evidence

Reflection on Study Characteristics. A total of 358 case definitions for evaluating BRD in dairy calves were identified in the literature from peer-reviewed journals and the gray literature. Publications included in this scoping review demonstrate a growing body of research, with the earliest included publication being from 1977, but a median publication year of 2017 and mode (i.e., most common) publication year of 2021. This indicates a

growing interest in BRD in dairy calves. Among the 256 publications published in the last decade (2014 to 2023), 79.7% (n = 204) provided a reference for the method used to classify BRD. Among these, nearly half (n = 100, 49%) referenced either the Wisconsin clinical scoring system (n = 79, 39%) or the California clinical scoring system (n = 21, 10%).

Reflection on Diagnostic Methods and Case Definitions. Clinical scoring systems alone (1) were the most commonly used diagnostic method (n = 148, 41.3%), followed by (4) physical exam parameters alone (n = 91, 25.4%), a term used to describe case definitions where publications did not provide assessment guidance for assessment of clinical signs. Case definitions were also reported for (5) physical exam parameters in combination with other diagnostic methods (n = 40, 11.1%), (6) TUS alone (n = 31, 8.6%), (2) clinical scoring system in combination with other diagnostic methods (n = 23, 6.2%) or (3) in combination with TUS (n = 10, 2.8%), and (7) other diagnostic methods (n = 15, 4.2%). Three categories involved clinical scoring systems, 2 of which extrapolate the diagnosis of BRD beyond the intended assessment parameters, to include other diagnostic methods. As these extrapolations were not intended by the scoring system authors, and have not been validated, it is difficult to assess their appropriateness or accuracy. Two categories involved physical exam parameters, which due to the category definition in this scoping review likely have subjective applications and are therefore difficult to replicate accurately or apply consistently. An attempt was made to compare case definitions at the entire case definition level (i.e., comparison of all components [e.g., these 5 clinical signs vs. these 4 clinical signs], scales, and case cut-points) within each diagnostic method; however, they displayed significant diversity which limited the ability to compare them. Instead, the commonality of individual factors used for each diagnostic method was evaluated. The common factors included in case definitions across diagnostic method categories involving clinical scoring systems and physical exam parameters included cough, nasal discharge, and body temperature. Diagnostic methods categories involving TUS commonly used lung consolidation depth ≥ 1 or ≥ 3 cm as a cut-off to define or contribute to defining a case.

Other diagnostic methods (e.g., farm treatment records or hematological parameters; in categories 2, 5, and 7) had extremely diverse compositions with no common trends identified or reports of validity. Further, many of these other diagnostic methods (farm treatment records or hematological parameters) have not been well defined in the literature, are difficult to implement in real-time for BRD assessment, and may have other barriers such as cost and equipment availability.

Several knowledge gaps and areas for improvement were identified in this scoping review. Case definitions displayed high diversity which inhibited comparison of entire definitions. Comparison of individual components of case definitions resulted in the identification of high variability of components between and within the 7 diagnostic method categories. However, as seen in Figure 2, it appears that a move away from unclear case definitions using predominantly physical exam parameters (a term used to describe case definitions where publications did not provide assessment guidance for assessment of clinical signs) and toward diagnosis with clinical scoring systems (providing assessment guidance for clinical signs) and TUS in recent years is occurring. Although this recent shift in popularity toward standardized case definitions of clinical scoring systems is an improvement (n = 148), as previously mentioned, this diagnostic method exhibits some key challenges, including diagnosis restricted to clinical cases; only moderate Se, Sp, and inter-rater reliability; and a still-present diversity of case definitions within this diagnostic method category. Diagnosis with TUS exhibits many previously mentioned benefits, including the ability to identify subclinical BRD cases; higher Se, Sp, and inter-rater reliability; and the similarity of case definitions identified by this scoping review (generally 1 or 3 cm depth or surface area [cm²] cut-off). Publications reporting case definitions using TUS are even more recent (Figure 2) and may contribute to fewer case definitions being identified (n = 31), which may be due to a few reasons. Thoracic ultrasound does require equipment and training and therefore may have more challenges for widespread adoption than clinical scoring systems. Additionally, it may be attributable to what is referred to in human medical research as the implementation gap. This is the observation that it takes an average of 17 yr for evidence to change practice (Morris et al., 2011), and with the first case definition identified by this scoping review that leveraged TUS alone for diagnosis being published in 2014, in 2024 we are only 10 yr into this 17-yr gap before widespread adoption. Overall, these findings appear to indicate that clinical scoring systems remain popular for use, but TUS provides more accurate diagnosis of BRD and may be in the preliminary phase of its use.

Frequently, no reference was provided for the case definition, calling into question the rationale of how it was formulated, and its validity and accuracy. The most commonly referenced methodology was the Wisconsin clinical scoring system (McGuirk, 2008; McGuirk and Peek, 2014; Calf Health Scoring Chart [https://fyi.extension.wisc.edu/heifermgmt/files/2015/02/calf_health_scoring_chart.pdf]). Of the 358 included case definitions, only 4 (1.1%) were referred to by the publication they were extracted from as validated (all of

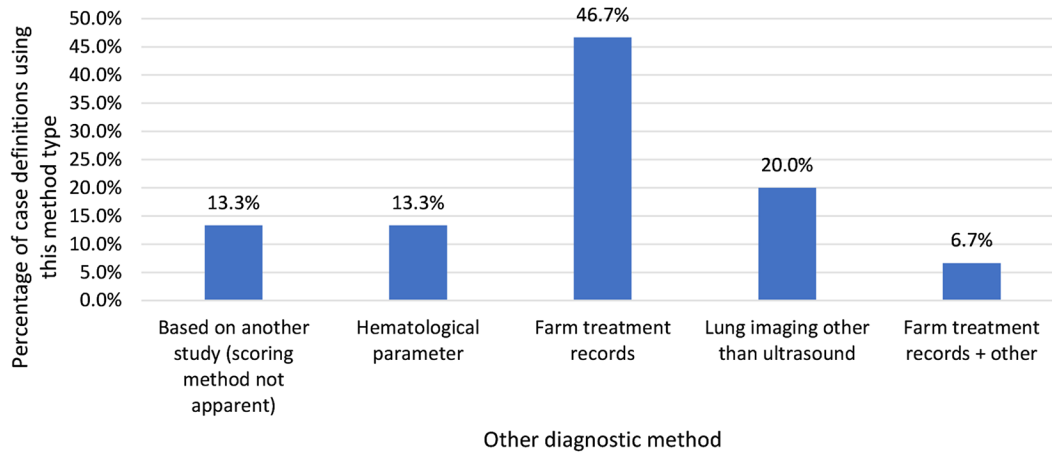


Figure 11. Components of the 15 case definitions other diagnostic methods from the scoping review of case definitions for respiratory disease in preweaning dairy calves. Other = drinking speed.

which referenced Love et al., 2014), and 9 (2.5%) were being validated by the publication they were extracted from. This may raise questions about how researchers choose their case definition, and the ability for them to identify and reproduce validated methodologies as they become available.

Severity of disease was not reported by the majority of case definitions, but a proportional comparison between diagnostic method categories shows that severity was most commonly reported by (6) TUS alone (9.7%, $n = 3$); (2) clinical scoring systems in combination with other diagnostic methods (8.7%, $n = 2$); (5) physical exam parameters in combination with other diagnostic methods (7.5%, $n = 3$); (1) clinical scoring systems (6.8%, $n = 10$); and (4) physical exam parameters alone (4.4%, $n = 4$). In 2 of the categories, (3) clinical scoring systems in combination with TUS and (7) other diagnostic methods, no case definitions reported severity of disease. Severity of disease is important to report because it gives insight to the burden of disease on the calf and likely will guide treatment decisions.

Parameters around implementing the case definition (e.g., who conducted the assessment or how often) were not well reported across all diagnostic methods, limiting reproducibility. Reporting these parameters is important for replication and aiding in the identification of reasons for inconsistent estimates of disease prevalence and should therefore be reported when possible.

With the growing interest in BRD diagnosis, a need and opportunity exists to establish a validated and standardized list of criteria to diagnose BRD, through the development of a COS. The findings of this scoping review (i.e., that clinical scoring systems are most commonly used but Se, Sp, and inter-rater reliability are only moderate, and TUS is the most accurate for diagnosis) align

with the a priori decision outlined in the COS protocol (see Notes) to develop a COS for BRD diagnosis using TUS alone, a clinical scoring system alone, and a combination of both (to balance the pros and cons of each individual diagnostic method).

Limitations

More than 75.0% ($n = 289$) of included publications were conducted in developed economies. In 2022, 76.5% of the world's reported dairy cow population ($n = 617,171,000$) was located in India, a developing economy (Agriculture Canada, 2024; UN DESA, 2024). The developing economies of Brazil ($n = 16,111,000$), Pakistan ($n = 15,100,000$), and China ($n = 10,943,000$) also surpassed the United States, the developed economy with the highest reported population of dairy cows ($n = 9,448,000$; Agriculture Canada, 2024). The numbers of included publications for India ($n = 1$), Brazil ($n = 16$), Pakistan ($n = 1$), and China ($n = 2$) were exceedingly low, far surpassed by publications from the United States ($n = 129$), the European Union ($n = 83$), or Canada ($n = 40$). Given this information, it is possible that parameters surrounding BRD in dairy calves worldwide vary by the country's economic status (due to factors such as breeds of cattle used, management practices and housing), and therefore that the findings of this scoping review may not have global representation or application.

Only English-language articles were eligible for inclusion due to language capacity of the review team. A total of 125 articles were excluded at full-text screening due to not being available in English, and many others were excluded at title and abstract screening for this reason. More than 75% ($n = 283$) of included publications were

conducted in developed economies, the majority of which were conducted in the English-language dominant countries of the United States, Canada, and the United Kingdom, which may reflect a bias due to the language restriction imposed, due in the search terms used to identify studies and article screening criteria.

CONCLUSIONS

This scoping review characterized the literature investigating BRD in dairy calves, including the study and population characteristics, the case definitions used, and the diagnostic method parameters used to implement the case definitions. Substantial diversity was present throughout the identified body of evidence, indicating probable difficulty in further synthesizing the data in a systematic review and meta-analysis. The development of a COS for BRD diagnosis in calves would support more consistent assessment of the disease and opportunities for evidence synthesis.

NOTES

The first author was supported by Dairy Farmers of Canada (Ottawa, Ontario, Canada). Although funders provided the graduate student stipend, they were not otherwise involved in study design or conduct. The authors thank Janelle Morrison (Department of Population Medicine, University of Guelph, Guelph, ON, Canada) for their contribution to citation screening and data extraction. Bart Pardon is the co-founder of a company named qTUS, which aims to make TUS more accessible for clinical practice. Supplemental material for this article is available at <https://doi.org/10.5683/SP3/ZIQBCD>. The protocol for developing a COS on diagnosing BRD in preweaning dairy calves led by the first author on this scoping review is available at <https://hdl.handle.net/10214/28405>. The scoping review protocol following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Protocols is available at <https://hdl.handle.net/10214/27276>. No human or animal subjects were used, so this analysis did not require approval by an Institutional Animal Care and Use Committee or Institutional Review Board. The authors have not stated any conflicts of interest.

Nonstandard abbreviations used: AMU = antimicrobial use; BRD = bovine respiratory disease; COS = core outcome set; IQR = interquartile range; N/A = not applicable; PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-Analysis; PRISMA-ScR = Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews; Se = sensitivity; Sp = specificity; TUS = thoracic ultrasound.

REFERENCES

- Agriculture Canada. 2024. World dairy situation. Number of dairy cows for selected countries (DP002A). Accessed May 1, 2024. <https://agriculture.canada.ca/en/sector/animal-industry/canadian-dairy-information-centre/statistics-market-information/world-situation>.
- Armstrong, R., B. J. Hall, J. Doyle, and E. Waters. 2011. 'Scoping the scope' of a Cochrane review. *J. Public Health (Oxf.)* 33:147–150. <https://doi.org/10.1093/pubmed/fdr015>.
- Bassil, K. 2016. Case definition. *Encyclopedia Britannica*. Accessed May 1, 2024. <https://www.britannica.com/science/case-definition>.
- Berge, A. C., D. A. Moore, T. E. Besser, and W. M. Sischo. 2009. Targeting therapy to minimize antimicrobial use in preweaned calves: Effects on health, growth, and treatment costs. *J. Dairy Sci.* 92:4707–4714. <https://doi.org/10.3168/jds.2009-2199>.
- Berman, J., D. Francoz, A. Abdallah, S. Dufour, and S. Buczinski. 2022. Development and validation of a clinical respiratory disease scoring system for guiding treatment decisions in veal calves using a Bayesian framework. *J. Dairy Sci.* 105:9917–9933. <https://doi.org/10.3168/jds.2021-21695>.
- Berman, J., I. Masseau, G. Fecteau, S. Buczinski, and D. Francoz. 2021. Comparison of thoracic ultrasonography and thoracic radiography to detect active infectious bronchopneumonia in hospitalized dairy calves. *J. Vet. Intern. Med.* 35:2058–2068. <https://doi.org/10.1111/jvim.16157>.
- Boccardo, A., S. Ferraro, G. Sala, V. Ferrulli, D. Pravettoni, and S. Buczinski. 2023. Bayesian evaluation of the accuracy of a thoracic auscultation scoring system in dairy calves with bronchopneumonia using a standard lung sound nomenclature. *J. Vet. Intern. Med.* 37:1603–1613. <https://doi.org/10.1111/jvim.16798>.
- Bokma, J., R. Boone, P. Deprez, and B. Pardon. 2019. Risk factors for antimicrobial use in veal calves and the association with mortality. *J. Dairy Sci.* 102:607–618. <https://doi.org/10.3168/jds.2018-15211>.
- Buczinski, S., D. Achard, and E. Timsit. 2021. Effects of calfhood respiratory disease on health and performance of dairy cattle: A systematic review and meta-analysis. *J. Dairy Sci.* 104:8214–8227. <https://doi.org/10.3168/jds.2020-19941>.
- Buczinski, S., and B. Pardon. 2020. Bovine respiratory disease diagnosis: What progress has been made in clinical diagnosis? *Vet. Clin. North Am. Food Anim. Pract.* 36:399–423. <https://doi.org/10.1016/j.cvfa.2020.03.004>.
- Cuevas-Gómez, I., M. McGee, J. M. Sanchez, E. O'Riordan, N. Byrne, T. McDaniel, and B. Earley. 2021. Association between clinical respiratory signs, lung lesions detected by thoracic ultrasonography and growth performance in pre-weaned dairy calves. *Ir. Vet. J.* 74:7. <https://doi.org/10.1186/s13620-021-00187-1>.
- Cumpston, M. F. E. 2023. Chapter IV: Updating a review. In *Cochrane handbook for systematic reviews of interventions version 6.4*. J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, and V. A. Welch, ed. Cochrane.
- Deepak, S. S. Aly, W. J. Love, P. C. Blanchard, B. Crossley, A. L. Van Eenennaam, and T. W. Lehenbauer. 2021. Etiology and risk factors for bovine respiratory disease in pre-weaned calves on California dairies and calf ranches. *Prev. Vet. Med.* 197:105506. <https://doi.org/10.1016/j.prevetmed.2021.105506>.
- Ferraro, S., G. Fecteau, J. Dubuc, D. Francoz, M. Rousseau, J.-P. Roy, and S. Buczinski. 2021. Scoping review on clinical definition of bovine respiratory disease complex and related clinical signs in dairy cows. *J. Dairy Sci.* 104:7095–7108. <https://doi.org/10.3168/jds.2020-19471>.
- Gaudino, M., B. Nagamine, M. F. Ducatez, and G. Meyer. 2022. Understanding the mechanisms of viral and bacterial coinfections in bovine respiratory disease: A comprehensive literature review of experimental evidence. *Vet. Res.* 53:70. <https://doi.org/10.1186/s13567-022-01086-1>.
- Haddaway, N. R., M. J. Page, C. C. Pritchard, and L. A. McGuinness. 2022. Prisma2020: An R package and shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and open synthesis. *Campbell Syst. Rev.* 18:e1230. <https://doi.org/10.1002/cl2.1230>.

- Heinrichs, A. J., C. M. Jones, L. R. Vanroekel, and M. A. Fowler. 2003. Calf track a system of dairy calf workforce management, training, and evaluation and health evaluation. *J. Dairy Sci.* 86 (Suppl. 1):115.
- Jourquin, S., T. Lowie, F. Debruyne, L. Chantillon, N. Vereecke, F. Boyen, R. Boone, J. Bokma, and B. Pardon. 2023. Dynamics of sub-clinical pneumonia in male dairy calves in relation to antimicrobial therapy and production outcomes. *J. Dairy Sci.* 106:676–689. <https://doi.org/10.3168/jds.2022-22212>.
- Kamel, M. S., J. L. Davidson, and M. O. Verma. 2024. Strategies for bovine respiratory disease (BRD) diagnosis and prognosis: A comprehensive overview. *Animals (Basel)* 14:627. <https://doi.org/10.3390/ani14040627>.
- Lago, A., S. M. McGuirk, T. B. Bennett, N. B. Cook, and K. V. Nordlund. 2006. Calf respiratory disease and pen microenvironments in naturally ventilated calf barns in winter. *J. Dairy Sci.* 89:4014–4025. [https://doi.org/10.3168/jds.S0022-0302\(06\)72445-6](https://doi.org/10.3168/jds.S0022-0302(06)72445-6).
- Love, W. J., T. W. Lehenbauer, P. H. Kass, A. L. Van Eenennaam, and S. S. Aly. 2014. Development of a novel clinical scoring system for on-farm diagnosis of bovine respiratory disease in pre-weaned dairy calves. *PeerJ* 2:e238. <https://doi.org/10.7717/peerj.238>.
- Maier, G. U., J. D. Rowe, T. W. Lehenbauer, B. M. Karle, D. R. Williams, J. D. Champagne, and S. S. Aly. 2019. Development of a clinical scoring system for bovine respiratory disease in weaned dairy calves. *J. Dairy Sci.* 102:7329–7344. <https://doi.org/10.3168/jds.2018-15474>.
- McGuirk, S. M. 2005. Otitis media in calves. Pages 228–230 in *Proc. 23. Am. Coll. Vet. Intern. Med.*
- McGuirk, S. M. 2008. Disease management of dairy calves and heifers. *Vet. Clin. North Am. Food Anim. Pract.* 24:139–153. <https://doi.org/10.1016/j.cvfa.2007.10.003>.
- McGuirk, S. M., and S. F. Peek. 2014. Timely diagnosis of dairy calf respiratory disease using a standardized scoring system. *Anim. Health Res. Rev.* 15:145–147. <https://doi.org/10.1017/S1466252314000267>.
- Moher, D., L. Shamseer, M. Clarke, D. Ghersi, A. Liberati, M. Petticrew, P. Shekelle, and L. A. Stewart. 2015. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst. Rev.* 4:1. <https://doi.org/10.1186/2046-4053-4-1>.
- Morris, Z. S., S. Wooding, and J. Grant. 2011. The answer is 17 years, what is the question: Understanding time lags in translational research. *J. R. Soc. Med.* 104:510–520. <https://doi.org/10.1258/jrsm.2011.110180>.
- Munn, Z., M. D. J. Peters, C. Stern, C. Tufanaru, A. McArthur, and E. Aromataris. 2018. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Med. Res. Methodol.* 18:143. <https://doi.org/10.1186/s12874-018-0611-x>.
- Ollivett, T. L., and S. Buczinski. 2016. On-farm use of ultrasonography for bovine respiratory disease. *Vet. Clin. North Am. Food Anim. Pract.* 32:19–35. <https://doi.org/10.1016/j.cvfa.2015.09.001>.
- Ollivett, T. L., J. L. Caswell, D. V. Nydam, T. Duffield, K. E. Leslie, J. Hewson, and D. Kelton. 2015. Thoracic ultrasonography and bronchoalveolar lavage fluid analysis in Holstein calves with subclinical lung lesions. *J. Vet. Intern. Med.* 29:1728–1734. <https://doi.org/10.1111/jvim.13605>.
- Page, M. J., J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, and D. Moher. 2021. Updating guidance for reporting systematic reviews: Development of the PRISMA 2020 statement. *J. Clin. Epidemiol.* 134:103–112. <https://doi.org/10.1016/j.jclinepi.2021.02.003>.
- Pardon, B., B. Catry, J. Dewulf, D. Persoons, M. Hostens, K. De Bleeker, and P. Deprez. 2012. Prospective study on quantitative and qualitative antimicrobial and anti-inflammatory drug use in white veal calves. *J. Antimicrob. Chemother.* 67:1027–1038. <https://doi.org/10.1093/jac/dkr570>.
- Probo, M., and M. C. Veronesi. 2022. Clinical scoring systems in the newborn calf: An overview. *Animals (Basel)* 12:3013. <https://doi.org/10.3390/ani12213013>.
- Sargeant, J. M., M. D. Bergevin, K. Churchill, K. Dawkins, B. Deb, J. Dunn, D. Hu, C. M. Logue, S. Meadows, C. Moody, A. Novy, A. M. O'Connor, M. Reist, Y. Sato, C. Wang, and C. B. Winder. 2019a. The efficacy of litter management strategies to prevent morbidity and mortality in broiler chickens: A systematic review and network meta-analysis. *Anim. Health Res. Rev.* 20:247–262. <https://doi.org/10.1017/S1466252319000227>.
- Sargeant, J. M., B. Deb, M. D. Bergevin, K. Churchill, K. Dawkins, J. Dunn, D. Hu, C. Moody, A. M. O'Connor, T. L. O'Sullivan, M. Reist, C. Wang, B. Wilhelm, and C. B. Winder. 2019b. Efficacy of bacterial vaccines to prevent respiratory disease in swine: A systematic review and network meta-analysis. *Anim. Health Res. Rev.* 20:274–290. <https://doi.org/10.1017/S1466252319000173>.
- Sargeant, J. M., S. C. Totton, M. Plishka, and E. R. Vriezen. 2021. *Salmonella* in animal feeds: A scoping review. *Front. Vet. Sci.* 8:727495. <https://doi.org/10.3389/fvets.2021.727495>.
- Tricco, A. C., E. Lillie, W. Zarin, K. K. O'Brien, H. Colquhoun, D. Levac, D. Moher, M. D. J. Peters, T. Horsley, L. Weeks, S. Hempel, E. A. Akl, C. Chang, J. McGowan, L. Stewart, L. Hartling, A. Aldcroft, M. G. Wilson, C. Garrity, S. Lewin, C. M. Godfrey, M. T. Macdonald, E. V. Langlois, K. Soares-Weiser, J. Moriarty, T. Clifford, Ö. Tunçalp, and S. E. Straus. 2018. Prisma extension for scoping reviews (Prisma-ScR): Checklist and explanation. *Ann. Intern. Med.* 169:467–473. <https://doi.org/10.7326/M18-0850>.
- UN DESA (United Nations Department of Economic and Social Affairs). 2024. World Economic Situation and Prospects 2024. United Nations Department of Economic and Social Affairs, New York, NY.
- Uyama, T., D. F. Kelton, E. I. Morrison, E. De Jong, K. D. McCubbin, H. W. Barkema, S. Dufour, J. Sanchez, L. C. Heider, S. J. Leblanc, C. B. Winder, J. T. McClure, and D. L. Renaud. 2022. Cross-sectional study of antimicrobial use and treatment decision for preweaning Canadian dairy calves. *JDS Commun.* 3:72–77. <https://doi.org/10.3168/jdsc.2021-0161>.
- Williamson, P. R., D. G. Altman, H. Bagley, K. L. Barnes, J. M. Blazeby, S. T. Brookes, M. Clarke, E. Gargon, S. Gorst, N. Harman, J. J. Kirkham, A. McNair, C. C. Prinsen, J. Schmitt, C. B. Terwee, and B. Young. 2017. The COMET Handbook: Version 1.0. *Trials* 18(Suppl. 3):280. <https://doi.org/10.1186/s13063-017-1978-4>.
- Wilson, D. J., G. Habing, C. B. Winder, and D. L. Renaud. 2023. A scoping review of neonatal calf diarrhea case definitions. *Prev. Vet. Med.* 211:105818. <https://doi.org/10.1016/j.prevetmed.2022.105818>.
- Winder, C. B., J. M. Sargeant, D. Hu, C. Wang, D. F. Kelton, M. A. Godkin, K. J. Churchill, and A. M. O'Connor. 2019a. Comparative efficacy of antimicrobials for treatment of clinical mastitis in lactating dairy cattle: A systematic review and network meta-analysis. *Anim. Health Res. Rev.* 20:229–246. <https://doi.org/10.1017/S1466252319000318>.
- Winder, C. B., J. M. Sargeant, D. Hu, C. Wang, D. F. Kelton, S. J. Leblanc, T. F. Duffield, J. Glanville, H. Wood, K. J. Churchill, J. Dunn, M. D. Bergevin, K. Dawkins, S. Meadows, B. Deb, M. Reist, C. Moody, and A. M. O'Connor. 2019b. Comparative efficacy of antimicrobial treatments in dairy cows at dry-off to prevent new intramammary infections during the dry period or clinical mastitis during early lactation: A systematic review and network meta-analysis. *Anim. Health Res. Rev.* 20:199–216. <https://doi.org/10.1017/S1466252319000239>.
- Winder, C. B., J. M. Sargeant, D. Hu, C. Wang, D. F. Kelton, S. J. Leblanc, T. F. Duffield, J. Glanville, H. Wood, K. J. Churchill, J. Dunn, M. D. Bergevin, K. Dawkins, S. Meadows, B. Deb, M. Reist, C. Moody, and A. M. O'Connor. 2019c. Comparative efficacy of teat sealants given prepartum for prevention of intramammary infections and clinical mastitis: A systematic review and network meta-analysis. *Anim. Health Res. Rev.* 20:182–198. <https://doi.org/10.1017/S1466252319000276>.
- Windeyer, M. C., K. E. Leslie, S. M. Godden, D. C. Hodgins, K. D. Lissimore, and S. J. Leblanc. 2014. Factors associated with morbidity, mortality, and growth of dairy heifer calves up to 3 months of age. *Prev. Vet. Med.* 113:231–240. <https://doi.org/10.1016/j.prevetmed.2013.10.019>.
- Woolums, A. R. 2015. Lower respiratory tract disease. Pages 583–617 in *Large animal medicine*. 5 ed. B. P. Smith, ed. Elsevier Mosby.

ORCIDS

K. J. Churchill,  <https://orcid.org/0000-0002-8767-3181>
C. B. Winder,  <https://orcid.org/0000-0002-7314-3657>

H. M. Goetz,  <https://orcid.org/0000-0001-9838-2921>
D. Wilson,  <https://orcid.org/0000-0001-8027-7998>
T. Uyama,  <https://orcid.org/0000-0002-2252-3043>
B. Pardon,  <https://orcid.org/0000-0003-1026-8433>
D. L. Renaud  <https://orcid.org/0000-0002-3439-3987>