STANDARD ARTICLE

Journal of Veterinary Internal Medicine

ACVIM
American College of
Veterinary Internal Medicine

Comparison of thoracic ultrasonography and thoracic radiography to detect active infectious bronchopneumonia in hospitalized dairy calves

Julie Berman ![ORCID] | Isabelle Masseau | Gilles Fecteau | Sébastien Buczinski ![ORCID] |
David Francoz

Département des sciences cliniques, Faculté de médecine vétérinaire, Université de Montréal, Saint-Hyacinthe, Quebec, Canada

Correspondence
Julie Berman, Faculté de médecine vétérinaire, Université de Montréal, 3200 rue Sicotte, CP 5000, Saint-Hyacinthe, Quebec J2S 7C6, Canada.
Email: julie.berman@umontreal.ca

Abstract

Background: The best test between thoracic ultrasonography (TUS) and thoracic radiography (TR) or the best combination of tests (series or parallel) to detect active infectious bronchopneumonia (BP) in hospitalized dairy calves remains unknown.

Hypothesis/Objectives: To estimate performances of TUS and TR to detect active BP in hospitalized dairy calves and to determine the best strategy for using these tests based on a panel diagnosis method (PDM). Performances of TUS and TR were hypothesized to be equivalent.

Animals: Fifty hospitalized dairy calves (≥7 days old; ≤100 kg; standing; pCO_2 ≥ 53 mm Hg; any reason of presentation).

Methods: Each calf prospectively and sequentially underwent physical examination, thoracic auscultation, blood analyses, and TUS and TR. Three blinded experts determined whether active BP was present/absent based on PDM. Krippendorff's alpha measured interexpert agreement. The sensitivities (Se) and specificities (Sp) of TUS and TR alone and in series or parallel were compared (McNemar's test; $P < .05$).

Results: Interexpert agreement was moderate at 0.58 (95%CI: 0.42; 0.73). The Se and Sp of TUS were 0.84 (95%CI: 0.60; 0.97) and 0.74 (95%CI: 0.57; 0.86), respectively. The Se and Sp of TR were 0.89 (95%CI: 0.67; 0.99) and 0.58 (95%CI: 0.39; 0.75), respectively. No significant difference was found in the Se and Sp of TUS and TR when analyzed alone, in series or in parallel.

Conclusion: Thoracic ultrasonography or TR alone equally detected active BP in hospitalized dairy calves. Series or parallel analysis provided no additional benefit. Its

---

---

ease of use and widespread accessibility support using TUS as a first-line test to detect active BP in hospitalized dairy calves.

1 | INTRODUCTION

Infectious bronchopneumonia (BP) in cattle is an infection of the lower respiratory tract with a multifactorial etiology.[1] In the absence of established consensus,[2] several studies have relied on 1 or more of the following findings to declare an individual as having active BP: when the individual manifests clinical signs such as fever, cough, tachypnea, respiratory distress, develops an inflammatory leukogram (eg, abnormal neutrophil count, or presence of toxic neutrophils, bands, or both modalities) combined with an increase of acute phase protein concentration (eg, fibrinogen, haptoglobin), and shows metabolic abnormalities reflecting impaired gas exchange.[1,3,4] Active BP is commonly associated with lung lesions characterized on thoracic radiography (TR) by focal or multifocal areas of alveolar and interstitial patterns,[5] and on thoracic ultrasonography (TUS) by lung consolidation.[4] In contrast, nonactive BP implies that clinical signs and hematologic abnormalities have resolved with or without remaining TR or TUS lung lesions.[6] Resolution of lung opacities might lag behind clinical and hematological resolution of infection or scars might persist long after the initial infection is cured.[7] In dairy calves, failure to accurately diagnose active BP and prolonged use of antimicrobials in nonactive BP are associated with major economic losses and overuse of antimicrobials.[8-10] Our ability to accurately detect active BP will promote a prudent use of antimicrobials and ultimately reduce economic losses.

Currently, the field diagnosis of BP by veterinarians relies on abnormal physical examination findings and abnormalities detected during thoracic auscultation.[2] Unfortunately, both clinical tools provide suboptimal performances in detecting active BP with sensitivities (Se) (ie, the ability of a test to detect a calf with active BP) and specificities (Sp) (ie, the ability of a test to detect a healthy calf) of 62 to 100% and 62 to 87% for physical examination,[11-13] and 73% and 53% for thoracic auscultation, respectively.[14] Development of superior diagnostic strategies will improve the diagnosis of active BP.

Thoracic ultrasonography and TR detect thoracic lesions in dairy calves.[15,16] TUS is a noninvasive, inexpensive, and relatively easy to interpret imaging test.[15] Previous on-farm studies have enabled TUS detection of active BP in preweaned dairy calves with Se and Sp of 89% and 95%, respectively.[4] However, TUS is limited to detecting lesions beneath the chest wall.[17] In comparison, TR detects thoracic lesions in calves with Se and Sp of 86% and 89%, respectively.[18] However, its ability to discriminate between active and nonactive BP in calves remains unknown and expertise required for interpretation limits its use in practice.[16] In contrast to TUS, TR can detect lesions deep into the lung parenchyma and provides an

overview of the entire lung, displayed on 1 or 2 images according to the size of the calf. For these reasons, TR is often preferred to assess BP in hospital settings, combined or not with TUS. Yet, a comparison of diagnostic performance of TUS alone, TR alone, or their combination in detecting active BP in hospitalized dairy calves is currently unavailable.

Recently, TUS and TR showed similar performances in detecting thoracic lesions in hospitalized dairy calves.[18] These data are inconsistent with those reported for the detection of active BP in neonatal calves for which TUS was less sensitive than TR,[19] while TUS was more sensitive in detecting the same condition in adult cattle.[20,21] The current literature offers insufficient evidence to determine which test is more accurate or whether combinations of tests represent the most appropriate strategy to detect active BP in calves.

To date, no gold standard (ie, test with a Se and Sp of 100%) has been validated for the antemortem diagnosis of active BP in dairy calves.[2] To overcome this deficiency, results from a series of imperfect tests (ie, test with a Se and Sp < 100%) can be combined to construct a reference standard outcome based on a panel diagnosis method (PDM) (ie, a consensus procedure among experts concluding on whether each calf has active BP or not based upon a set of imperfect test results).[22] Briefly, a PDM considers all relevant information gathered from medical history, clinical examination, and other test results. Because experts make their diagnosis of active BP based on all available information, the PDM classification as active or nonactive BP could more closely reflect the diagnostic process taking place in a clinical setting.[22] Based on published guidelines and ability to determine the performance of TUS, TR, or both tests for the detection of active BP in a clinical setting, a PDM was selected.[23]

The main objective of this study was to use PDM to estimate the Se and Sp of TUS and TR in detecting active BP in hospitalized dairy calves. The secondary objective was to determine which test (TUS or TR) alone or which combination of tests (ie, TUS and TR performed in series or in parallel) best detects active BP in hospitalized dairy calves. We hypothesized that performances of TUS and TR would be similar in detecting active BP in hospitalized calves.

2 | MATERIALS AND METHODS

The design was a prospective cohort study. The design, conduct, and results are reported according to STARD (Standards for Reporting of Diagnostic Accuracy Studies) 2015 guidelines (Additional File 1). The study protocol was approved by the institutional ethical committee (#20160111).

2060 | Journal of Veterinary Internal Medicine

ACVIM
Open Access
American College of
Veterinary Internal Medicine

BERMAN ET AL.

2.1 | Study sample

The study sample was the subject of a previous study aimed at estimating the performances of TUS and TR in detecting thoracic lesions assessed by CT. Further demographic details on the study sample are described elsewhere.[18] Briefly, all calves admitted to the Veterinary Medical Teaching Hospital (VMTH) of University of Montreal from March 22nd to May 6th, 2016, June 13th to November 4th, 2016, and January 16th to October 4th, 2017, were prospectively enrolled.

No sample size calculation was estimated prior to the study since little information was available on performances of TR and TUS in detecting active BP in hospitalized dairy calves. The number of calves was determined based on previous reports in newborn calves (n = 56)[19] and 1 human study revealing differences between diagnostic procedures in 52 pediatric participants.[24] Fifty calves of various dairy breeds (ie, Holstein, Ayrshire, Jersey, Brown Swiss, Red Holstein), older than 7 days of age and weighing <100 kg (constraint related to CT gantry size), were included.

Calves with arterial $pCO_2 \geq 53$ mm Hg or presented in decubitus were excluded to allow immediate treatment intervention.[25] Calves for which the consent form was not signed in time for enrolment or considered unsuitable for transport to the diagnostic imaging suite (eg, biosecurity) were also excluded (Figure 1).

2.2 | Clinical definition of active BP

The presence or absence of active BP in calves was determined by PDM based on all relevant available information collected by 1 single American College of Veterinary Internal Medicine (Large Animal Internal Medicine) food animal emphasis (ACVIM-FA) board certified veterinarian (first author) on each calf including: history, physical examination and thoracic auscultation findings, and results from arterial blood gas analysis, CBC and serum biochemistry profile. Collection of data is detailed in Table 1.

Component test results were reported on a single spreadsheet (MS Excel, Microsoft Corporation, Redmond, Washington) facilitating blind review (Additional File 2). Three experts consisting of board-certified specialists from the ACVIM-FA (Expert 1, Expert 2, Expert 3), unaware of the TUS and TR results, independently reviewed the spreadsheet 5 months after the end of the study. Aiming for each calf to be evaluated by at least 2 experts, each expert analyzed two-thirds of the calves. Overall, each third of the study sample was evaluated by a different pair of experts (17 calves by Expert 1 and Expert 2, 17 calves by Expert 1 and Expert 3, and 16 calves by Expert 2 and Expert 3). Where there was disagreement between the 2 experts, the third expert (Expert 1, 2, or 3 according to the initial pair) was asked to rule on the presence/absence of active BP.

2.3 | Index tests

Each calf underwent TUS first and then TR (paired design). Diagnostics were completed within 24 hours of admission to the VMTH to avoid or minimize changes after treatment initiation or disease progression.

2.3.1 | Thoracic ultrasonography

Each calf was clipped from the 3rd to the 10th intercostal space on the left side, and from the 1st to the 10th intercostal space on the right side. Acoustic coupling gel was then applied to the entire area.[15] A portable ultrasound unit equipped with a 7.5 MHz linear transducer (Sonoscape S6V, Shenzhen, China) was used.
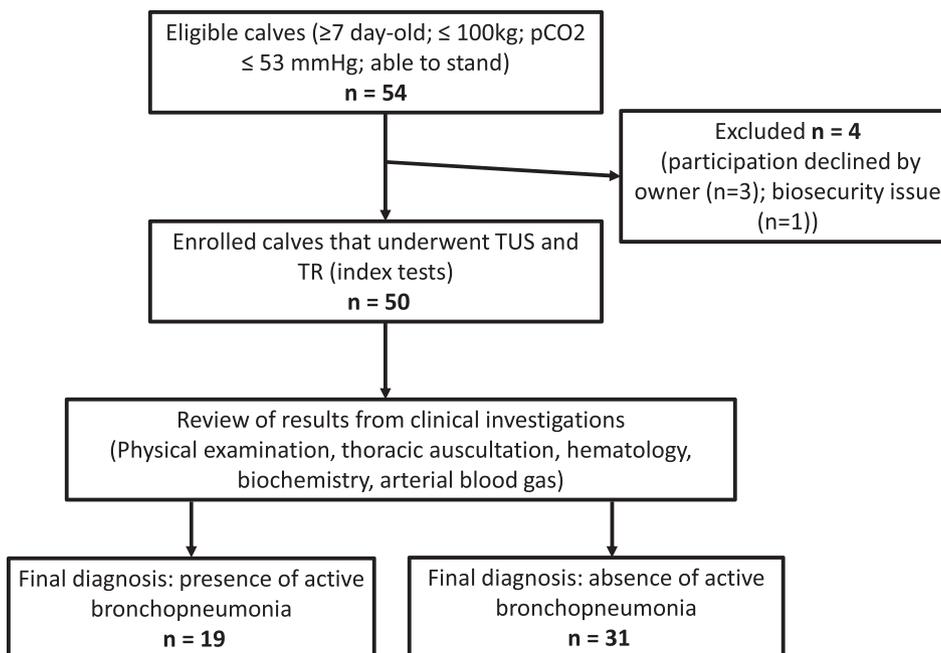
```
┌─────────────────────────────┐
│ Eligible calves (≥7 day-old; │
│ ≤ 100kg; pCO2 ≤ 53 mmHg;    │
│ able to stand)               │
│          n = 54              │
└─────────────────────────────┘
              │         ┌──────────────────────────────┐
              │─────────▶ Excluded n = 4               │
              │         │ (participation declined by    │
              │         │ owner (n=3); biosecurity issue│
              │         │          (n=1))               │
              ▼         └──────────────────────────────┘
┌─────────────────────────────┐
│ Enrolled calves that         │
│ underwent TUS and            │
│ TR (index tests)             │
│          n = 50              │
└─────────────────────────────┘
              ▼
┌─────────────────────────────────────────────┐
│ Review of results from clinical investigations│
│ (Physical examination, thoracic auscultation, │
│ hematology, biochemistry, arterial blood gas) │
└─────────────────────────────────────────────┘
      │                            │
      ▼                            ▼
┌──────────────────────┐  ┌──────────────────────┐
│ Final diagnosis:      │  │ Final diagnosis:      │
│ presence of active    │  │ absence of active     │
│ bronchopneumonia      │  │ bronchopneumonia      │
│       n = 19          │  │       n = 31          │
└──────────────────────┘  └──────────────────────┘
```

FIGURE 1 Flowchart illustrating the overall flow of calves in the study

TABLE 1 Relevant information collected by 1 single ACVIM-FA board certified veterinarian on the 50 calves included in the study and used for the panel diagnosis method

Component tests	Details of procedure	Available to experts
History and reason for presentation	Breed, age, sex, medical history, colostrum intake, weaning, reason for presentation, previous treatments	No
Physical examination	Rectal temperature, heart rate, respiratory rate, respiratory pattern, nasal or ocular discharge, ears/head carriage, inductive or spontaneous cough, joint distension, diarrhea, umbilical enlargement	Rectal temperature, heart rate, respiratory rate and any additional abnormal clinical signs
Thoracic auscultation	Abnormal sounds: wheezes, crackles, and increased or decreased bronchovesicular sounds[a] An evaluation sheet was used to record the location of abnormal sounds, as previously described in Buczinski et al[26]	Presence of abnormal sounds, their grade, and their location
Arterial blood gas	Blood was collected from the medial intermediate branch of the caudal auricular artery with a 1 mL syringe and a heparinized 25G needle, as previously described in Bleul et al[27] The analysis took place <5 minutes after collection PH, pO2, pCO2, SatO2 and HCO3- were obtained using an automated analyzer (ABL 80 FLEX, Radiometer America, Inc, Brea, California)	pH, pO2, pCO2, SatO2 and HCO3-values
Complete blood count	Twenty milliliters of jugular venous blood was introduced into an EDTA vacutainer tube Analyses were performed in the hour following collection Hematocrit, platelets count, neutrophils count, lymphocytes count, eosinophils count, and fibrinogen concentration measured by heat precipitation were obtained with an automated hematology analyzer (Advia 120, Siemens, Tarrytown, New York)	Hematocrit, regenerative anemia comments, fibrinogen concentration, and neutrophil morphology and count Platelets count, lymphocytes count, and eosinophils count were available only if values were abnormal
A serum biochemistry profile	Twenty milliliters of jugular venous blood was deposited into a redtop vacutainer tube Analyses were performed in the hour following collection Glucose, BUN, creatinine, GGT, GLDH, AST, CK, albumin, globulin, sodium, chloride, magnesium, calcium, phosphorus, bicarbonate, and anion gap, were obtained with an automated biochemistry analyzer (UniCel DxC 600, Beckman Coulter)	Globulin concentration Glucose, BUN, creatinine, GGT, GLDH, AST, CK, albumin, sodium, chloride, magnesium, calcium, and phosphorus were available only if values were abnormal

[a]A scoring system was developed to quantify increasing auscultated lung sounds based on comparison between bronchovesicular and heart sounds was used irrespectively of the location of the auscultation (i.e., independent of how distant from the heart); Grade 0: Bronchovesicular sounds barely perceptible and cardiac sounds predominated; Grade 1: Bronchovesicular sounds easily auscultated but remaining less audible than cardiac sounds; Grade 2: Bronchovesicular sounds of equal intensity compared to cardiac sounds perceived; Grade 3: Bronchovesicular sounds of superior intensity compared to cardiac sounds perceived. In one calf, the presence of a heart murmur grade 4/6 hampered determination of a grade.
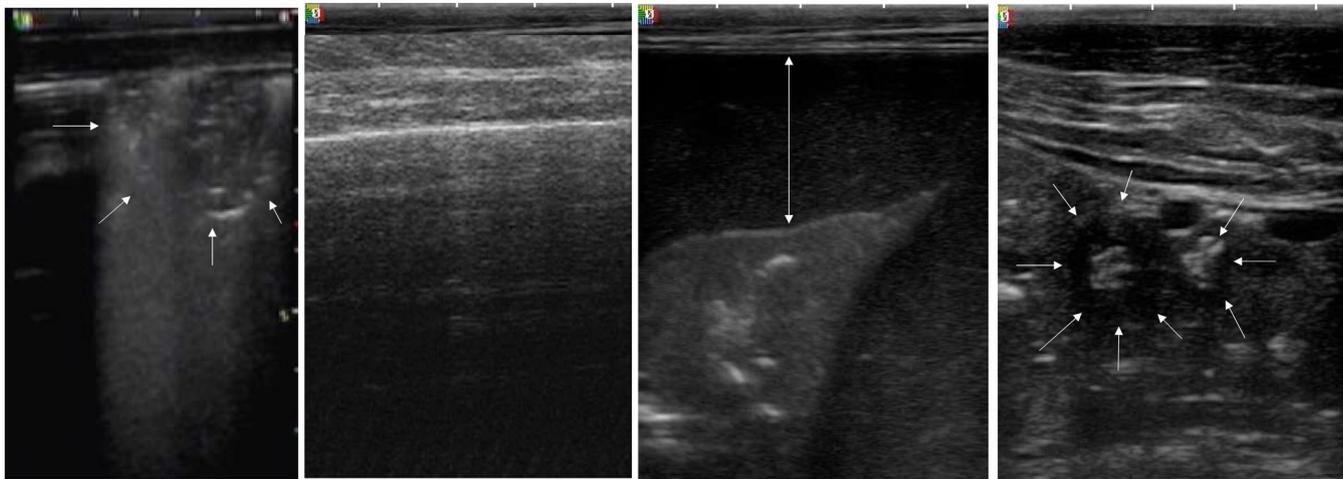
The real-time sonographic evaluation of the pleurae and lungs is described in Figure 2. Maximal depth of lung consolidations (cm), cavitary lesions (cm), and effusions (cm) were measured using the ultrasound software analyzing tools. Only lesions deeper than or equal to 1 cm were considered clinically relevant.[28]

Thoracic ultrasonography was considered positive (TUS+) if any of the predefined lesions were present. In their absence, TUS was considered negative (TUS−). The depth and intercostal space of each lesion was recorded on a scoring sheet, as previously described.[26] Standardized TUS were performed by a single experienced operator (first author), unaware of TR results, after physical examination, arterial blood gas analysis and thoracic auscultation.

2.3.2 | Thoracic radiography

Each calf underwent standing left to right latero-lateral and recumbent ventro-dorsal radiographs. A 500 mA and 150 kVp Siemens Vertix machine (Siemens, Mississauga, ON, Canada) and an AGFA DX-S CR system, using a focused grid (grid ratio of 6:1; Reina Imaging, Crystal Lake, IL, USA), were utilized.[18]

Radiographs were transferred to a Picture Archiving and Communication System (PACS) and viewed on a single workstation using AGFA Impax 6 (Agfa, Toronto, ON, Canada). A single American College of Veterinary Radiology (ACVR) certified diagnostic imaging specialist, unaware of the component tests and TUS results, reviewed the radiographs 10 months after enrolment of the last calf in the study.

2062 | Journal of Veterinary Internal Medicine

ACVIM
Open Access
American College of
Veterinary Internal Medicine

BERMAN ET AL.

Lung consolidation defined as hypoechoic lung tissue resembling the sonographic appearance of hepatic parenchyma (arrows)

Pneumothorax diagnosed when the sliding movement of the pleurae was no longer visible

Pleural effusion diagnosed when the pleurae were separated from one another by anechoic or echogenic fluid (double-arrow)

Nodule described as a well-defined discrete hypoechoic circular region inside the lung parenchyma (arrows)

FIGURE 2   Ultrasound images corresponding to the 4 lesions assessed in dairy calves based on Babkine et al.[17] Pleural irregularities, B-lines, and comet-tails were overlooked since their clinical relevance in calves is controversial[6]

Qualitative variables	Numbers	Proportion (%)
Breed		
Holstein	39	78
Jersey	6	12
Red Holstein	3	6
Ayrshire	2	4
Sex		
Female	46	92
Male	4	8
Reason for presentation		
Digestive[a]	16	32
Locomotor[b]	8	16
Umbilical[c]	7	14
Pneumonia	6	12
Otitis	4	8
Teaching	4	8
Ocular	2	4
Other[d]	3	6
Treatment prior to admission[e]		
Yes	40	80
No	10	20
Continuous variables	**Mean**	**SD**
Age (days)	30	24.5
Weight (kg)	53	16

TABLE 2   Descriptive data of the 50 calves recruited in the study (≥7 days-old; $pCO_2 ≤ 53$ mm Hg; ≤100 kg; able to stand)

[a]Digestive: Diarrhea, colic, regurgitation, anorexia, intestinal obstruction, bloat.
[b]Locomotor: Septic arthritis, contracture, laceration.
[c]Umbilical: Hernia, urachus.
[d]Other: Fever of unknown origin.
[e]Treatment prior to admission: antimicrobials and/or anti-inflammatory.

TABLE 3 Descriptive data summarizing thoracic lesions detected on thoracic ultrasonography and thoracic radiography in 50 calves with (active BP +) or without (active BP−) active bronchopneumonia (BP) based on a panel diagnosis method

		Groups	
	Thoracic lesions	Active BP+ (n = 19)	Active BP− (n = 31)
Thoracic ultrasonography[a]	Lung consolidation	16	8
	Pneumothorax	1	0
	Pleural effusion	0	0
	Nodule	1	0
	Absence of lesion	3	23
Thoracic radiography[b]	Bronchial pattern	3	3
	Interstitial pattern	7	8
	Alveolar pattern	17	13
	Nodule	1	0
	Pleural effusion	0	0
	Pneumothorax	1	0
	Absence of lesion	2	18

[a]Thoracic ultrasonography: (a) lung consolidation defined as hypoechoic lung tissue resembling the sonographic appearance of hepatic parenchyma; (b) pneumothorax diagnosed when the sliding movement of the pleurae was no longer visible; (c) pleural effusion diagnosed when the pleurae were separated from one another by anechoic or echogenic fluid; and (d) nodule described as a well-defined discrete hypoechoic circular region inside the lung parenchyma.[17]

[b]Thoracic radiography: (a) bronchial pattern defined as the presence of airways with prominent and/or thickened walls; (b) interstitial pattern defined as the presence of an area of the lung field with increased opacity combined with blurred vascular borders; (c) nodule defined as the presence of a focal area of increased soft tissue opacity (<3 cm) with relatively well-defined borders; (d) alveolar pattern defined as the presence of an area of increased soft tissue opacity silhouetting with pulmonary vessels located within the opacified area; (e) pleural effusion diagnosed based on the retraction of lung lobe margins, pleural fissure lines, and fluid opacity between the thoracic wall and the lung margins; and (f) pneumothorax diagnosed based on the presence of air between the lung lobes and the thoracic wall causing retraction of the lung margins.[18]

Radiographic variables used to determine if TR was considered positive (TR+) are described in Table 3. Thoracic radiographic studies exempt of any of the listed thoracic lesions were considered negative (TR−).

2.4 | Statistical analysis

All statistical analyses were performed using commercial software (SAS v 9.4, SAS, Cary, North Carolina) and R statistical software (R Core Team 2020).[29] Additional details of statistical analyses are present in Additional File 3.

2.4.1 | Assessment of the interexpert agreement

To determine whether expert experience influenced the classification of calves with or without active BP, the interexpert agreement was calculated using a Krippendorff's alpha (Kalpha; kripp. boot package https://github.com/MikeGruz/kripp.boot).[30,31] Briefly, the Kalpha represents the pairwise agreement between 2 experts, averaged over all expert pairs and number of calves, handling for missing values.[32] The confidence intervals (CIs) were obtained using 5000 bootstrap replicates in the absence of validated parametric method to calculate them.[31] In the absence

of specific benchmarks of agreement, Kalpha was interpreted according to LandisKoch.[33]

2.4.2 | Estimation of performances of TUS and TR

The calf was considered as the unit of interest. The Se, Sp, positive likelihood ratio (LR+) (ie, percentage of calves with active BP having a positive test result (Se), divided by the percentage of calves without active BP but deemed positive based on the test result (ie, false positive fraction = 1 - Sp)), and negative likelihood ratio (LR-) (ie, percentage of calves with active BP classified as negative based on the test result (ie, false negative fraction = 1- Se), divided by the percentage of calves without active BP with a negative test result (Sp)) were estimated for each test based on the 2 by 2 table.[34,35] Confidence intervals were calculated using exact binomial methods for Se and Sp, and using a natural logarithmic transform and the delta method for LR.[36]

2.4.3 | Comparison of TUS and TR

Thoracic ultrasonography and TR were assumed dependent since both tests aimed to detect the same pathological process: lung lesions. Covariance between both tests was estimated as in Dohoo et al.[35] Considering

an assumption of covariance between both tests and the paired design of the study, a 2-sided McNemar's test was performed (Additional File 3).

Additionally, as recommended by Pepe,[36] a comparison between both tests was performed by estimating the relative probabilities of TUS and TR. This metric was performed in addition to the McNemar's test since it offers superior flexibility and robustness across a broader range of study designs, promoting the external validity of the results.[36] Relative probabilities were estimated as follows:

$$rSe\,(TUS, TR) = Se\,of\,TUS/Se\,of\,TR.$$

$$rSp\,(TUS, TR) = Sp\,of\,TR/Sp\,of\,TR.$$

The difference between both tests was assessed by calculating the joint 95%CI of both rSe (TUS, TR) and rSp (TUS, TR) as described by Pepe,[36] (Additional File 3).

A rSe (TUS, TR) (rSp (TUS, TR)) equal to a number x > 1 denoted that the Se of TUS (Sp of TUS) was x times superior to the Se of TR (Sp of TR). Inversely, an rSe (TUS, TR) (rSp (TUS, TR)) equal to a number x < 1 denoted that the Se of TUS (Sp of TUS) was x times inferior to the Se of TR (Sp of TR). The relative probabilities were considered significant if 1 was not included in the 95%CIs for the ratios of Se and Sp.[36]

2.4.4 | Sequential use of TUS and TR

Sensitivity and Sp of each test were calculated to assess the sequential parallel or serial use of TR and TUS. Briefly, calves positive on 1 or both tests were considered positive for the parallel assessment. For the assessment of TR and TUS performed in series, only calves that tested positive on both tests were considered positive. Considering the covariance between both tests, TR and TUS in series or in parallel were assessed based on Dohoo et al.[35] Confidence intervals were estimated using an exact binomial method,[36] (Additional File 3).

The difference between the Se (Sp) of TUS alone and TR alone and the Se (Sp) of both tests used in series or in parallel was estimated by assessing the difference between marginal probabilities in calves with (without) active BP using the 2-sided McNemar's test. A *P*-value <.05 was considered significant (Additional File 3).

3 | RESULTS

3.1 | Study sample

Fifty-four calves met the inclusion criteria. Figure 1 illustrates the step-by-step approach to the study, the number of calves that underwent each step, and their final classification. Three calves were excluded since the owner refused to sign the statement of informed consent. An additional calf was excluded for biosecurity reasons (BVDV positive). Consequently, 50 calves were enrolled in the study; their demographic information is detailed in Table 2.

3.2 | Clinical diagnosis of active BP

For each calf, results of component tests and classification by experts are shown in Additional File 2. There was no missing data. Diagnosis of active BP was consensual between 2 experts for 80% (n = 40) of the study sample. Ten remaining calves required assessment by a third expert (highlighted in dark gray). The final classification identified 19 calves (38%) with active BP, this number being higher than the proportion of calves referred for active BP (6/50; 12%). The inter-expert agreement was moderate (Kalpha = 0.58 [95%CI: 0.42; 0.73]).

3.3 | Performances of TUS and TR

Thoracic lesions identified on TUS and TR are presented in Table 3. There was no missing data or any adverse events. Lung consolidation on TUS and presence of an alveolar pattern on TR were the most prevalent thoracic lesions and were both detected in calves with and without active BP. As expected, more calves in the nonactive BP group were exempt of lesions on both modalities compared to calves with active BP. Conversely, TUS and TR failed to identify thoracic lesions in 3 and 2 calves with active BP, respectively. Table 4 shows descriptive data summarizing the comparison of TUS and TR in calves with and without active BP. Out of 19 calves with active BP, TUS and TR were positive for 16 and 17 calves, respectively, resulting in respective sensitivities of 84% and 89% (Table 5). Thoracic lesions were identified in more calves without active BP by TR than by TUS, leaving only slightly more than half of the calves recognized as free of thoracic lesions. Details on tests performances are reported in Table 5.

TABLE 4 Descriptive data summarizing the comparison of thoracic ultrasonography (TUS) and thoracic radiography (TR) in 50 calves with (active BP+) or without (active BP−) active bronchopneumonia (BP) based on a panel diagnosis method

Active BP−			
	TR+	TR−	Total
TUS+	7	1	8
TUS−	6	17	23
Total	13	18	31
Active BP+			
	TR+	TR−	Total
TUS+	16	0	16
TUS−	1	2	3
Total	17	2	19

Notes: TUS+: presence of lung consolidation (depth ≥ 1 cm), pneumothorax, pleural effusion, or nodule (depth ≥ 1 cm) on thoracic ultrasonography. TUS−: absence of lung consolidation (depth ≥ 1 cm), pneumothorax, pleural effusion, or nodule (depth ≥ 1 cm) on thoracic ultrasonography. TR+: presence of bronchial pattern, interstitial pattern, alveolar pattern, nodule, pleural effusion, or pneumothorax on thoracic radiography. TR−: absence of bronchial pattern, interstitial pattern, alveolar pattern, nodule, pleural effusion, or pneumothorax on thoracic radiography.

TABLE 5   Test sensitivity, test specificity, positive likelihood ratio (LR+) and negative likelihood ratio (LR−) along with their 95% confidence intervals for thoracic ultrasonography (TUS) and thoracic radiography (TR) using a panel diagnosis method as a gold standard

	Sensitivity	Specificity	LR+	LR−
TUS	0.84 (0.60; 0.97)	0.74 (0.57; 0.86)	3.2 (1.7; 5.9)	0.22 (0.08; 0.64)
TR	0.89 (0.67; 0.99)	0.58 (0.39; 0.75)	2.1 (1.4; 3.3)	0.19 (0.14; 0.71)

3.4 | Comparison between TUS and TR

The marginal probabilities of TUS and TR were not significantly different in calves with active BP since no significant difference in Se was found between both tests (McNemar's; $P > .05$). Additionally, there was no difference between marginal probabilities in calves without active BP since no significant difference in Sp was found between both tests (McNemar's; $P = .06$). The relative probabilities were rSe(TUS, TR) = 0.94 (95%CI: 0.84; 1.06) and rSp(TUS, TR) = 1.28 (95% CI: 0.77; 2.1). Both rSe and rSp were not significantly different from 1.

3.5 | Sequential use of TUS and TR

A parallel analysis of the combined Se and Sp of both tests achieved 0.89 (95%CI: 0.67; 0.99) and 0.52 (95%CI: 0.36; 0.73), respectively. The series analysis revealed an Se of 0.84 (95%CI: 0.60; 0.97) and an Sp of 0.77 (95%CI: 0.60; 0.88), respectively. The marginal probabilities of using both tests in series vs parallel were significantly different in calves without active BP since the Sp when using both tests in series was significantly different from the Sp of using both tests in parallel (McNemar's; $P < .05$). No other differences in marginal probabilities of TUS or TR in calves with or without active BP were found, as the performances of TUS or TR used in series or in parallel were not different from those of TUS (Se of 0.84 (95%CI: 0.60; 0.97) and Sp of 0.74 (95%CI: 0.57; 0.86)) or TR (Se of 0.89 (95%CI: 0.67; 0.99) and Sp of 0.58 (95%CI: 0.39; 0.75)) used alone (McNemar's; $P > .05$).

4 | DISCUSSION

In this study, TR and TUS performed comparably in detecting active BP in hospitalized dairy calves based on PDM. Despite wide CI of accuracy estimates, the Se of both TUS and TR in detecting active BP was >80%, with Sp of TR < 60% and Sp of TUS close to 75%. The serial interpretation of both tests was more specific in detecting active BP in hospitalized dairy calves than with parallel analysis. The accuracy of TUS or TR alone was not different from that obtained when using both tests in series or in parallel.

The study employed a prospective cohort design offering the advantage of enrolling calves with a large spectrum of clinical presentation, ranging from healthy calves to calves with mild and moderate active BP, to calves with severe active BP. In contrast, retrospective and case-control designs might select cases at the ends of the spectrum, either healthy or calves with severe active BP, resulting in an

overestimation of Se and Sp of investigated tests.[2,37,38] Secondly, the classification of active or nonactive BP was performed blindly and without relying on results of index tests (TUS and TR), limiting incorporation bias. The absence of such bias promoted better accuracy in estimation and comparison of both tests' performances.[37,38] Thirdly, in contrast to studies reporting observational comparisons of TUS and TR,[19-21] this study reports a robust statistical comparison of both tests, with a greater or equivalent number of calves used in previous studies (n = 50 vs n = 56,[19] n = 2,[21] or n = 1[20]). Finally, this study relied on experts' clinical judgment to make a diagnosis of active BP based on a comprehensive set of clinical information and blood work results. This approach differs from studies using the presence of thoracic lesions on diagnostic imaging examinations as a gold standard[18] or latent class analysis.[4]

In this study, calves were classified as having active or nonactive BP based on a PDM. Although imperfect, this method circumvents the absence of generally accepted reference standard procedure for concluding on active BP and allows consideration of multiple sources of information to reach a diagnosis.[22,23] Indeed, 2 limitations resulted from this imperfect gold standard. Firstly, a misclassification error led to an underestimation of Se and Sp of both tests.[37,38] However, this misclassification error was independent of TUS and TR results (nondifferential classification bias) since this error was similar for both tests. Therefore, despite the impact on TUS and TR performance estimations, this classification bias should not have impacted the comparison of both tests. Secondly, PDM relied on experts' interpretation of several imperfect tests, which could carry a certain degree of subjectivity, and therefore be subjected to a classification ruled by chance alone. To minimize this subjectivity, 3 experienced and qualified experts were selected, a number reportedly sufficient to provide reproducible diagnoses in human medicine.[39] Despite high levels of qualification, the inter-expert agreement (Kalpha = 0.58) showed moderate agreement highlighting the current limits and difficulties in classifying a calf as having active or nonactive BP based on clinical variables alone. This emphasizes the need for better definition and incorporation of all diagnostics to improve diagnosis and optimize treatment.

The Se and Sp of TR for the detection of thoracic lesions has recently been described without regard to the active aspect of disease.[18] Despite wide CI of accuracy estimates, our results show that TR is not only sensitive in detecting thoracic lesions, but also in recognizing hospitalized calves with active BP, as supported by a numerical estimate of 89%. Similarly, TUS revealed a numerical estimation of the Se of 84%. This finding is in accordance with a previous study showing an Se of TUS of 89% in detecting active BP in dairy calves in a low prevalence study sample.[4] This result is also similar to the Se of TUS

reported in chronic BP with an Se of 85%,[40] in subclinical BP (ie, calves with lung lesions but without clinical signs) with an Se of 94%,[6] and in detecting thoracic lesions with an Se of 81%.[18]

Interestingly, the Sp of 58% and 74% for TR and TUS, respectively, was relatively low compared with studies reporting an Sp of TR of 89% for the detection of thoracic lesions in calves[18] and an Sp of TUS > 90% in detecting chronic, subclinical or active BP.[4,6,40] Different hypotheses could explain these low Sp. It is likely that tissue repair can only be initiated once the infection is controlled. A delay between resolution of clinical signs and hematologic disturbances and resolution of thoracic lesions is thus inevitable.[7] Moreover, the sample enrolled in this study consisted of calves referred by practitioners and consequently most enrolled calves (80%) had already received a treatment (antimicrobial, NSAIDs, or both treatments) prior to presentation. This treatment could have controlled the infection, but thoracic lesions were still present while TR and TUS were deemed positive, thus lowering their specificity. Moreover, despite clearance of the infection, lung damage could be too extensive for complete tissue repair, leaving permanent scars which in turn could appear as lung opacities and be falsely interpreted as active lesions. Furthermore, prior treatment might have masked clinical signs and reduced the hematological variables below the reference values for an inflammatory process to be considered present, leading to a misclassification of nonactive BP by the experts, while pathogens were still present and lesions detected by TUS or TR.

In this study, the Sp of TUS was 74% vs that of TR of 58%. During healing, the lesions could hypothetically decrease in size, allowing tissue in contact with the pulmonary pleura to heal, preventing detection by TUS,[17] in comparison to TR. The criteria used for defining a positive TUS or TR test results could also have influenced the estimation of performances. Comet-tails were excluded from the analysis along with pleural irregularities and B-lines since these ultrasound abnormalities can be found in calves without any lung lesions on postmortem examination, and consequently, do not have a clear clinical relevance.[6] We also only considered lesions deeper than or equal to 1 cm as clinically relevant.[28] For TR, an alveolar pattern was identified as the most prevalent lesion in calves with active BP, analogous to previous studies conducted in dairy calves.[16,41] Since little information was available on the correlation between the presence of an alveolar pattern and the diagnosis of active BP in dairy hospitalized calves, no specific features (eg, size, well defined, or blurred margins) of the alveolar pattern precluded its classification as positive. As for TUS, it is therefore possible that small focal areas of alveolar pattern on TR might not have necessarily represented pneumonic lesions but instead scarring or another condition (eg, pulmonary contusions) and have overestimated the proportion of calves with active BP on TR, and thus increasing the proportion of false-positive calves.

In contrast to previous studies in cattle, we found no significant difference between TUS and TR in detecting active BP in this study.[19-21] The exact reason for this discrepancy is uncertain, although our study has a few dissimilarities with respect to these previous studies: study sample (newborns[19] and adults[20,21] vs calves) and clinical cases included (including various pneumopathies[19] or only

focusing on severe pleuropneumonia[20,21] vs active BP). Our study design also differs from those of previous studies by evaluating TR and TUS performances on a PDM rather than incorporating TR results in the definition of active BP.

The important width of 95%CI reported for each accuracy estimate (uncertainty by lack of power), could have led to a misinterpretation of equivalent performances of index tests in our study. Since little information was available on performances of TR and TUS in detecting active BP in dairy hospitalized calves, the number of 50 calves was determined in order to achieve a statistical power comparable to previous studies conducted on cattle. A lack of power could have prevented detecting a difference in any of the accuracy estimates and any combinations of tests (TUS alone, TR alone, TUS and TR in series, TUS and TR in parallel). Based on a post hoc sample size calculation, a total of 90 calves without active BP would be needed to detect a difference in Sp of 74 vs 58% of TUS and TR, respectively, with a power of 80% (PROC POWER, SAS 9.4). In addition, a total of 310 calves with active BP would be needed to detect a difference in Se of 84 vs 89% of TUS and TR, respectively, with a power of 80%. In the end, for a study sample with a prevalence of active BP of 78%, a minimum of 400 calves would be needed to detect a difference in the Se and Sp between TUS and TR, with a power of 80%. Despite a heavy case load of dairy calves presented to our institution yearly (around 200 calves), the estimated prevalence of active BP is only around 15%. Consequently, we estimate that 10 years would be needed to reach this number, decreasing our ability to control for all factors such as equipment used for blood tests, imaging equipment and personnel performing the examinations during the entire duration of the study. Unnecessary manipulations on a large number of calves, despite their noninvasiveness, might also raise ethical concerns. Furthermore, since the Se of each combination of tests was >80% in this study, it is unlikely that the small differences between Se in detecting active BP in a dairy hospitalized calf, although obtained with a sufficient power, would have any clinical impact on the selection of TUS vs TR. However, our results do not allow us to exclude that a larger sample size might affect the Sp of a test combination in the future. We could anticipate that if the trend toward greater Sp of TUS than TR is found to be significant when applied to a larger sample size, this could mean that TUS would provide fewer false positives than TR. Clinically, this would mean that TUS would be superior to TR in ruling in active BP. This finding would justify the use of TUS as a first-line test to detect active BP in hospitalized dairy calves. From this perspective, the use of TR in the event of a positive TUS could be justified to support the presence of active BP (Sp of using both tests in series being of 77%) and to obtain a general assessment of thoracic lesions potentially useful for the follow-up and the prognosis of the disease. Importantly, this approach might decrease the average cost of investigation (with the cost of TUS without TR) and limit manipulation of calves in case of TUS negative results. Since the repercussions of this assumption are not negligible, and since based on our results the targeted sample size of 90 enrolled subjects is reasonably achievable, a future study comparing specificities of TUS and TR for the detection of active BP would be justified, and indeed necessary.

BERMAN ET AL.

Journal of Veterinary Internal Medicine ACVIM
Open Access
American College of
Veterinary Internal Medicine

| 2067

4.1 | Conclusion

The major findings of this study include the Se of TUS and TR in detecting active BP in dairy hospitalized calves estimated at 84% and 89%, respectively, and the Sp of TUS and TR estimated at 74% and 58%, respectively. The analysis in series of TUS and TR was more specific than using both tests in parallel. Interestingly, there was no significant difference between either combination of tests and the performances of TUS or TR alone. Clinically, the easier use and accessibility of TUS compared with TR would support recommending its utilization as a first-line test to detect active BP in dairy hospitalized calves. However, we recognize that this statement is premature and further studies with an appropriate sample size are needed to definitively conclude on the best combination of the index tests (TUS and TR).

CONFLICT OF INTEREST DECLARATION
Zoetis was not involved in the study design, data collection, data analysis, and writing of the manuscript. Sebastien M. Buczinksi serves as Consulting Editor for Experimental Design and Statistics for the Journal of Veterinary Internal Medicine. He was not involved in review of this manuscript.

OFF-LABEL ANTIMICROBIAL DECLARATION
Authors declare no off-label use of antimicrobials.

INSTITUTIONAL ANIMAL CARE AND USE COMMITTEE (IACUC) OR OTHER APPROVAL DECLARATION
Approved by the Comité d'éthique de l'utilisation des animaux de l'Université de Montréal (#20160111).

HUMAN ETHICS APPROVAL DECLARATION
Authors declare human ethics approval was not needed for this study.

ORCID
Julie Berman https://orcid.org/0000-0001-7947-760X
Sébastien Buczinski https://orcid.org/0000-0002-8460-4885

REFERENCES
1. Woolums AR. Lower respiratory tract disease. *Smith BP: Large Animal Internal Medicine*. 5th ed. St Louis, MO: Elsevier; 2015:583-617.
2. Buczinski S, Pardon B. Bovine respiratory disease diagnosis: what progress has been made in clinical diagnosis? *Vet Clin North Am Food Anim Pract*. 2020;36:399-423.
3. Reinhold P, Rabeling B, Günther H, Schimmel D. Comparative evaluation of ultrasonography and lung function testing with the clinical signs and pathology of calves inoculated experimentally with Pasteurella muftocida. *Vet Rec*. 2002;150:109-114.
4. Berman J, Francoz D, Dufour S, Buczinski S. Bayesian estimation of sensitivity and specificity of systematic thoracic ultrasound exam for diagnosis of bovine respiratory disease in pre-weaned calves. *Prev Vet Med*. 2019;162:38-45.
5. Tegtmeier C, Arnbjerg J. Evaluation of radiology as a tool to diagnose pulmonic lesions in calves, for example prior to experimental infection studies. *J Vet Med B Infect Dis Vet Public Health*. 2000;47:229-234.
6. Ollivett TL, Caswell JL, Nydam DV, et al. Thoracic ultrasonography and bronchoalveolar lavage fluid analysis in Holstein calves with subclinical lung lesions. *J Vet Intern Med*. 2015;29:1728-1734.
7. Surén P, Try K, Eriksson J, Khoshnewiszadeh B, Wathne KO. Radiographic follow-up of community-acquired pneumonia in children. *Acta Paediatr*. 2008;97:46-50.
8. Teixeira AG, McArt JA, Bicalho RC. Thoracic ultrasound assessment of lung consolidation at weaning in Holstein dairy heifers: reproductive performance and survival. *J Dairy Sci*. 2017;100:2985-2991.
9. Dunn TR, Ollivett TL, Renaud DL, et al. The effect of lung consolidation, as determined by ultrasonography, on first-lactation milk production in Holstein dairy calves. *J Dairy Sci*. 2018;101:5404-5410.
10. Windeyer MC, Leslie KE, Godden SM, Hodgins DC, Lissemore KD, LeBlanc SJ. Factors associated with morbidity, mortality, and growth of dairy heifer calves up to 3 months of age. *Prev Vet Med*. 2014;113:231-240.
11. Buczinski S, Lo T, Dendukuri N. Bayesian estimation of the accuracy of the calf respiratory scoring chart and ultrasonography for the diagnosis of bovine respiratory disease in pre-weaned dairy calves. *Prev Vet Med*. 2015;119:227-231.
12. Love WJ, Lehenbauer TW, Van Eenennaam AL, et al. Sensitivity and specificity of on-farm scoring systems and nasal culture to detect bovine respiratory disease complex in preweaned dairy calves. *J Vet Diagn Invest*. 2016;28:119-128.
13. Maier GU, Rowe JD, Lehenbauer TW, et al. Development of a clinical scoring system for bovine respiratory disease in weaned dairy calves. *J Dairy Sci*. 2019;102:7329-7344.
14. Buczinski S, Menard J, Timsit E. Incremental value (Bayesian framework) of thoracic ultrasonography over thoracic auscultation for diagnosis of bronchopneumonia in preweaned dairy calves. *J Vet Intern Med*. 2016;30:1396-1401.
15. Ollivett TL, Buczinski S. On-farm use of ultrasonography for bovine respiratory disease. *Vet Clin North Am Food Anim Pract*. 2016;32:19-35.
16. Farrow CS. Bovine pneumonia: its radiographic appearance. *Vet Clin N Am Food Anim Pract*. 1999;15:301-358.
17. Babkine M, Blond L. Ultrasonography of the bovine respiratory system and its practical application. *Vet Clin North Am Food Anim Pract*. 2009;25:633-649.
18. Berman J, Masseau I, Fecteau G, et al. Comparison between thoracic ultrasonography and thoracic radiography for the detection of thoracic lesions in dairy calves using a two-stage Bayesian method. *Prev Vet Med*. 2020;184:105-153.
19. Jung C, Bostedt H. Thoracic ultrasonography technique in newborn calves and description of Normal and pathological findings. *Vet Radiol Ultrasound*. 2004;45:331-335.
20. Reef VB, Boy MG, Reid CF, Elser A. Comparison between diagnostic ultrasonography and radiography in the evaluation of horses and cattle with thoracic disease: 56 cases (1984-1985). *J Am Vet Med Assoc*. 1991;198:2112-2118.
21. Braun U, Pusterla N, Fluckiger M. Ultrasonographic findings in cattle with pleuropneumonia. *Vet Rec*. 1997;141:12-17.
22. Rutjes AW, Reitsma JB, Coomarasamy A, et al. Evaluation of diagnostic tests when there is no gold standard. A review of methods. *Health Technol Assess*. 2007;11:1-50.
23. Reitsma JB, Rutjes AW, Khan KS, et al. A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. *J Clin Epidemiol*. 2009;62:797-806.

2068 | Journal of Veterinary Internal Medicine

ACVIM
Open Access
American College of
Veterinary Internal Medicine

BERMAN ET AL.

24. Iorio G, Capasso M, De Luca G, et al. Lung ultrasound in the diagnosis of pneumonia in children: proposal for a new diagnostic algorithm. *PeerJ*. 2015;3:e1374.

25. Smith BP. *Large Animal Internal Medicine*. 5th ed. St Louis, MO: Elsevier; 2015.

26. Buczinski S, Forte G, Belanger AM. Short communication: ultrasonographic assessment of the thorax as a fast technique to assess pulmonary lesions in dairy calves with bovine respiratory disease. *J Dairy Sci*. 2013;96:4523-4528.

27. Bleul U. Respiratory distress syndrome in calves. *Vet Clin North Am Food Anim Pract*. 2009;25:179-193.

28. Buczinski S, Forte G, Francoz D, et al. Comparison of thoracic auscultation, clinical score, and ultrasonography as indicators of bovine respiratory disease in preweaned dairy calves. *J Vet Intern Med*. 2014; 28:234-242.

29. Proutskova P, Gruszczynski M. An r package for performing bootstrap replicates of Krippendorff's alpha on intercoder reliability data. R-package 'Kripp.boot'; 2017. https://github.com/MikeGruz/kripp.boot. Accessed August 23, 2019.

30. Pardon B, Buczinski S, Deprez PR. Accuracy and inter-rater reliability of lung auscultation by bovine practitioners when compared with ultrasonographic findings. *Vet Rec*. 2019;185:109.

31. Zapf A, Castell S, Morawietz L, Karch A. Measuring inter-rater reliability for nominal data–which coefficients and confidence intervals are appropriate? *BMC Med Res Methodol*. 2016;16:93.

32. Gwet KL. Agreement coefficients for nominal ratings: a review. In: Gwet KL, ed. *Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Raters*. 4th ed. Gaithersburg, MD: Advanced Analytics, LLC; 2014:27-73.

33. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*. 1977;33:363-374.

34. Timsit E, Leguillette R, White BJ, et al. Likelihood ratios: an intuitive tool for incorporating diagnostic test results into decision-making. *J Am Vet Med Assoc*. 2018;252:1362-1366.

35. Dohoo IR, Martin SW, Stryhn H, et al. *Methods in Epidemiologic Research*. 2nd ed. Charlottetown, PEI: VER Inc; 2012.

36. Pepe MS. *The Statistical Evaluation of Medical Tests for Classification and Prediction*. Oxford, UK: Oxford University Press; 2003.

37. Buczinski S, O'Connor AM. Specific challenges in conducting and reporting studies on the diagnostic accuracy of ultrasonography in bovine medicine. *Vet Clin North Am Food Anim Pract*. 2016;32:1-18.

38. Schmidt RL, Factor RE. Understanding sources of bias in diagnostic accuracy studies. *Arch Pathol Lab Med*. 2013;137:558-565.

39. Van Houten CB, Naaktgeboren CA, Ashkenazi-Hoffnung L, et al. Expert panel diagnosis demonstrated high reproducibility as reference standard in infectious diseases. *J Clin Epidemiol*. 2019;112:20-27.

40. Rabeling B, Rehage J, Dopfer D, et al. Ultrasonographic findings in calves with respiratory disease. *Vet Rec*. 1998;143:468-471.

41. Fowler J, Stieger-Vanegas SM, Vanegas JA, Bobe G, Poulsen KP. Comparison of thoracic radiography and computed tomography in calves with naturally occurring respiratory disease. *Front Vet Sci*. 2017;4:101.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.